

UNIVERSITÉ DU QUÉBEC

**MÉMOIRE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES**

**COMME EXIGENCE PARTIELLE DE LA MAÎTRISE EN
MATHÉMATIQUES ET INFORMATIQUE APPLIQUÉES**

PAR

TANKAM NGNEUPI AXEL-FRANK

**ÉVALUATION AUTOMATISÉE DES DOMMAGES AUTOMOBILES À
L'AIDE DE DETR ET DE LA LOGIQUE FLOUE**

AOÛT 2026

UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

SERVICE DE LA BIBLIOTHÈQUE

Avertissement

L'auteur de ce mémoire a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de ce mémoire.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce document requiert son autorisation.

Dédicace

À ma famille, pour son soutien constant, sa patience et la confiance qu'elle m'a toujours accordée.

Je dédie ce mémoire à toutes les personnes qui ont cru en moi pendant ce parcours, et qui m'ont encouragé à persévérer jusqu'à l'aboutissement de ce travail.

Remerciements

Avant toute chose, je rends grâce à Dieu, source de toute sagesse et de toute force, pour m'avoir guidé et soutenu tout au long de ce parcours. Dans les moments de doute comme dans les instants de progrès, Sa présence m'a permis de trouver la persévérance, la lucidité et la détermination nécessaires à l'aboutissement de ce travail.

Je tiens à exprimer ma profonde reconnaissance à mes parents, dont le soutien indéfectible a été le socle de mon engagement. Leur amour, leurs sacrifices et leur foi en mes capacités ont façonné mon parcours et nourri mon ambition. Ce travail est, en grande partie, le reflet de ce qu'ils m'ont transmis.

J'adresse mes sincères remerciements à mes encadrants, le Dr Usef FAGHIHI et le Dr Fadel TOURE, pour la qualité exceptionnelle de leur accompagnement. Leur rigueur scientifique, leur disponibilité constante et la pertinence de leurs orientations ont été déterminantes dans la construction et la maturation de cette recherche. Leur exigence intellectuelle a constitué un véritable levier d'excellence.

Je souhaite également remercier mes collègues, partenaires de recherche et amis, pour la richesse des échanges, la qualité des discussions et l'esprit de collaboration qui ont jalonné ce travail. Leur contribution, tant humaine que scientifique, a permis d'enrichir ma réflexion et d'élargir mes perspectives.

Mes remerciements s'adressent enfin à toutes les personnes qui, de près ou de loin, ont apporté leur soutien à ce projet. Chaque encouragement, chaque conseil et chaque geste, aussi discret soit-il, ont contribué à sa réalisation.

Au-delà de sa portée académique, ce mémoire constitue une étape importante de mon développement intellectuel et personnel. Il témoigne d'un engagement fondé sur la rigueur, la persévérance et la quête constante de compréhension. Il marque également le point de départ d'une ambition plus large : celle de contribuer activement, avec exigence et intégrité, à l'avancement des connaissances dans mon domaine.

Résumé

Dans un contexte où les compagnies d'assurance et les systèmes d'expertise recherchent des solutions capables d'automatiser l'analyse des dommages automobiles de manière fiable et interprétable, ce mémoire présente DIRF-R (Detection–Intersection–Ratio–Fuzzy Reasoning), un pipeline d'évaluation automatisée des dommages automobiles fondé sur quatre étapes explicitement séparées : Détection, Intersection, Ratio et Fuzzy Reasoning. Il peut être précédé d'une phase d'isolation du véhicule principal afin de limiter les interférences causées par la présence de plusieurs véhicules dans une même scène. L'approche repose sur deux détecteurs DETR entraînés séparément, l'un pour les pièces du véhicule et l'autre pour les zones de dommages. Les sorties des deux détecteurs sont ensuite associées par une mesure géométrique réelle d'intersection, calculée à l'aide de l'Intersection over Union (IoU), afin de déterminer quelles pièces sont touchées. Cette association permet d'estimer, pour chaque pièce, un Damage Ratio représentant la proportion de surface atteinte. Enfin, un module de logique floue interprète ce ratio de manière graduelle pour attribuer un niveau de gravité plus proche d'un raisonnement d'expert que d'une décision binaire rigide.

Le jeu de données final utilisé dans ce travail comprend 1 503 images après augmentation par retournement vertical (flip) dans Roboflow, dont 805 images de véhicules intacts et 698 images de véhicules endommagés. Le corpus initial, constitué de plus de 700 images annotées, provient principalement d'images acquises au Canada en collaboration avec Progi et d'un complément public issu de l'Université Stanford. Chaque image a été annotée manuellement via la plateforme Roboflow selon une taxonomie structurée des pièces automobiles et des zones de dommages. Les annotations ont été exportées sous forme de boîtes englobantes et utilisées pour entraîner séparément les deux branches de détection.

Les expérimentations quantitatives mettent en évidence les performances des deux détecteurs utilisés dans DIRF-R. La branche dommages obtient un mAP@0,5 de 0,759, un mAP@50:95 de 0,452 et un score F1 de 0,505, tandis que la branche pièces atteint respectivement 0,558, 0,289 et 0,312. Au-delà de ces métriques de détection, l'apport principal de DIRF-R réside dans l'enchaînement cohérent entre détection, association géométrique, quantification locale et interprétation floue. L'architecture proposée améliore ainsi l'explicabilité du système et le rend plus proche des besoins de l'assurance automobile et de l'expertise technique.

Mots-clés : DETR, logique floue, dommages automobiles, vision par ordinateur, IoU, Damage Ratio, raisonnement flou, évaluation automatisée.

Abstract

In a context where insurance companies and appraisal systems are seeking solutions capable of automating vehicle damage analysis in a reliable and interpretable manner, this thesis presents DIRF-R, an automated vehicle damage assessment pipeline built around four explicitly separated stages: Detection, Intersection, Ratio, and Fuzzy Reasoning. It may be preceded by a main-vehicle isolation stage designed to limit interference when several vehicles appear in the same scene. The method relies on two independently trained DETR detectors, one dedicated to vehicle parts and the other to damage regions. Their predictions are associated through a geometric overlap measure based on Intersection over Union in order to determine which parts are affected. This association makes it possible to compute, for each part, a Damage Ratio representing the proportion of impacted surface. A fuzzy logic module then interprets this ratio progressively and assigns a severity level that is closer to expert reasoning than to a rigid binary decision.

The final dataset used in this work contains 1,503 images after vertical flip augmentation in Roboflow, including 805 images of intact vehicles and 698 images of damaged vehicles. The initial corpus, composed of more than 700 annotated images, was assembled mainly from images acquired in Canada in collaboration with Progi and complemented with a public repository from Stanford University. Each image was manually annotated in Roboflow using a structured taxonomy of vehicle parts and damage regions. Bounding-box annotations were then exported and used to train the two detection branches separately.

Quantitative experiments highlight the performance of the two detectors used in DIRF-R. The damage branch achieves a $\text{mAP}@0.5$ of 0.759, a $\text{mAP}@50:95$ of 0.452, and an F1-score of 0.505, while the parts branch reaches 0.558, 0.289, and 0.312 respectively. Beyond detection metrics, the main contribution of DIRF-R lies in the coherent chaining of detection, geometric association, local quantification, and fuzzy interpretation. The proposed architecture therefore improves interpretability and makes the system more relevant for insurance workflows and technical appraisal.

Keywords: DETR, fuzzy logic, vehicle damage assessment, computer vision, IoU, Damage Ratio, fuzzy reasoning, automated evaluation.

Acronymes

- AP: Average Precision
- CBNet: Composite Backbone Network
- COCO: Common Objects in COntext
- CNN: Convolutional Neural Network
- DETR: DETection TRansformer
- DAB-DETR: Dynamic Anchor Boxes for DETR
- DIoU: Distance Intersection over Union
- DN-DETR: DeNoising DETR
- DIRF-R: Detection–Intersection–Ratio–Fuzzy Reasoning
- F1 : score harmonique entre la précision et le rappel
- FN: False Negative
- FP: False Positive
- GCNet: Global Context Network
- GPU: Graphics Processing Unit
- HTC: Hybrid Task Cascade
- IoU: Intersection over Union
- GIoU: Generalized Intersection over Union
- CIoU: Complete Intersection over Union
- mAP@0,5 : mean Average Precision à un seuil d'IoU de 0,5
- mAP@50 :95 : mean Average Precision calculé sur plusieurs seuils d'IoU de 0,50 à 0,95
- PANet: Path Aggregation Network

- R-CNN: Region-based Convolutional Neural Network
- ROI: Region of Interest
- TP: True Positive

Table des matières

<i>Avertissement</i>	<i>I</i>
<i>Dédicace.....</i>	<i>II</i>
<i>Remerciements.....</i>	<i>III</i>
<i>Résumé.....</i>	<i>IV</i>
<i>Abstract.....</i>	<i>V</i>
<i>Acronymes.....</i>	<i>VI</i>
<i>Table des matières.....</i>	<i>VIII</i>
<i>Liste des figures.....</i>	<i>XII</i>
<i>Liste des tableaux.....</i>	<i>XIII</i>
CHAPITRE 1 – Introduction	1
1.1 Contexte général.....	1
1.2 Problématique de recherche	2
1.3 Objectifs du mémoire	2
1.4 Contributions principales	3
1.5 Organisation du mémoire	4
1.6 Délimitation du travail.....	4
CHAPITRE 2 – État de l’art.....	5
2.1 Introduction générale sur l’état de l’art.....	5
2.2 Travaux connexes sur la détection et la classification des dommages.....	5
2.3 Travaux connexes sur la segmentation et la localisation des pièces et des dommages	6
2.4 Travaux connexes pour la robustesse et les performances en conditions réelles	7

2.5 Travaux connexes concernant les modèles à base de Transformers et les approches efficaces.....	8
2.6 Apport de la logique floue et synthèse des lacunes	9
2.7 Positionnement spécifique de DIRF-R.....	10
CHAPITRE 3 – Méthodologie proposée.....	11
3.1 DETR comme base de détection.....	11
3.1.1 Architecture interne de DETR	11
3.1.2 Deux détecteurs spécialisés dans DIRF-R.....	12
3.2 Logique floue : concepts, intérêt et place dans le pipeline.....	13
3.3 Vue d'ensemble du pipeline DIRF-R.....	14
3.4 Détection : sortie des détecteurs de pièces et de dommages.....	15
3.5 Intersection : association géométrique explicite par IoU	16
3.6 Ratio : calcul du Damage Ratio local.....	17
3.7 Raisonnement flou : interprétation graduelle de la gravité.....	18
3.8 Évaluation locale et globale.....	21
3.9 Robustesse structurelle : isolation du véhicule principal et stabilisation	22
3.10 Distinction gauche/droite des pièces.....	23
CHAPITRE 4 – Expérimentations	25
4.1 Description du jeu de données.....	25
4.1.1 Prétraitement et filtrage des images.....	26
4.1.2 Annotation manuelle avec Roboflow	26
4.1.3 Taxonomie des classes	27
4.2 Outils et environnement de travail	30
4.3 Protocole expérimental	31
4.4 Métriques d'évaluation	32

4.5 Lecture expérimentale spécifique à DIRF-R.....	33
4.5.1 Préparation des annotations pour l'entraînement.....	34
4.5.2 Schéma d'entraînement des deux branches.....	34
4.5.3 Du résultat brut au pipeline complet.....	35
CHAPITRE 5 – Résultats.....	36
5.1 Présentation synthétique des résultats.....	36
5.2 Interprétation détaillée des performances	36
5.3 Ce que les métriques disent, et ce qu'elles ne disent pas	37
5.4 Évaluation de bout en bout du pipeline DIRF-R.....	37
5.5 Analyse qualitative des meilleurs cas et des cas difficiles	39
5.6 Analyse conceptuelle de l'apport des composantes de DIRF-R.....	40
5.7 Synthèse.....	41
5.8 Discussion qualitative et intérêt applicatif.....	41
CHAPITRE 6 – Menaces à la validité de l'étude	44
6.1 Validité interne et propagation des erreurs	44
6.2 Validité externe et changement de domaine	44
6.3 Validité de construit.....	45
6.4 Limites pratiques	46
CHAPITRE 7 – Conclusion.....	47
7.1 Bilan général.....	47
7.2 Apports scientifiques et méthodologiques	47
7.3 Limites et perspectives	47
7.4 Conclusion finale	48
Références	49

<i>Annexes</i>	52
ANNEXE A : Liste des pièces détectées de la voiture dans la Figure 5	52
ANNEXE B : Détail des pièces endommagées, ratios de dommage et niveaux de gravité associés à la Figure 5	52
ANNEXE C : Liste des pièces détectées de la voiture dans la Figure 8	52

Liste des figures

<i>Figure 1 : Architecture de DETR.</i>	<i>12</i>
<i>Figure 2 : Vue d'ensemble de l'architecture DIRF-R.....</i>	<i>15</i>
<i>Figure 3 : Fonctions d'appartenance des niveaux de gravité.....</i>	<i>19</i>
<i>Figure 4 : Projection d'une valeur de sévérité floue.....</i>	<i>20</i>
<i>Figure 5 : Sortie brute des détecteurs de pièces et de dommages.....</i>	<i>20</i>
<i>Figure 6 : Pourcentages de dommages par pièce détectée.</i>	<i>21</i>
<i>Figure 7 : Distinction exploratoire gauche/droite sur une image réelle.....</i>	<i>24</i>
<i>Figure 8 : Exemples de résultats obtenus avec DIRF-R.....</i>	<i>39</i>

Liste des tableaux

<i>Tableau 1 – Composition du corpus et répartition entraînement/validation/test</i>	<i>25</i>
<i>Tableau 2 – Classes annotées pour les branches pièces et dommages</i>	<i>28</i>
<i>Tableau 3 – Répartition des annotations avant augmentation par retournement vertical dans Roboflow</i>	<i>29</i>
<i>Tableau 4 – Hyperparamètres et environnement d'entraînement des modèles DETR</i>	<i>31</i>
<i>Tableau 5 – Évaluation qualitative de bout en bout du pipeline DIRF-R</i>	<i>38</i>
<i>Tableau 6 – Analyse conceptuelle de l'apport des composantes du pipeline DIRF-R</i>	<i>40</i>
<i>Tableau 7 – Résultats quantitatifs et comparaison avec YOLOv8/FuzzyYOLOv8</i>	<i>43</i>

CHAPITRE 1 – Introduction

1.1 Contexte général

L'évaluation automatisée des dommages automobiles constitue aujourd'hui un enjeu stratégique pour les compagnies d'assurance, les ateliers de carrosserie, les experts techniques et, plus largement, tous les acteurs engagés dans la gestion numérique des sinistres. Dans les pratiques traditionnelles, l'état d'un véhicule est encore très souvent estimé à partir d'une inspection humaine, réalisée sur place ou à distance à partir d'images. Cette méthode reste pertinente sur le plan métier, mais elle présente plusieurs limites : elle dépend fortement de l'expérience de l'expert, elle nécessite du temps, elle engendre des coûts d'opération élevés et elle peut varier d'un évaluateur à un autre. À grande échelle, cette variabilité devient problématique, notamment lorsque les entreprises souhaitent uniformiser les processus d'estimation, réduire les délais de traitement et documenter les décisions de manière cohérente.

Dans le même temps, la disponibilité massive d'images issues de déclarations de sinistres, d'applications mobiles, d'inspections en atelier ou de bases de données internes a ouvert la voie à des méthodes d'analyse automatisée par vision par ordinateur. Les réseaux de neurones profonds ont permis des avancées majeures dans la détection d'objets [1], l'extraction de caractéristiques visuelles [2], la modélisation par attention [3], la segmentation des pièces automobiles [7] et la classification visuelle des dommages [4]. Toutefois, les dommages automobiles représentent un cas d'usage particulièrement exigeant. Les pièces visibles varient selon l'angle de vue, les reflets de la carrosserie peuvent masquer une altération, les dommages peuvent être partiels, subtils ou très localisés, et plusieurs véhicules peuvent parfois apparaître dans la même image. En pratique, un système utile ne doit pas seulement repérer des boîtes englobantes ; il doit aussi être capable de relier les dommages détectés aux pièces réellement touchées et d'en proposer une interprétation intelligible.

Dans ce contexte, les détecteurs à base de Transformers, et plus particulièrement DETR (DEtection TRansformer), offrent une perspective intéressante. DETR reformule la détection d'objets comme une prédiction d'ensemble de bout en bout, fondée sur l'attention globale et l'appariement bipartite optimal. Cette architecture est particulièrement adaptée aux scènes où plusieurs éléments doivent être détectés simultanément sans dépendre d'ancres prédéfinies ni de nombreuses heuristiques de post-traitement. Néanmoins, même une bonne détection ne suffit pas à produire une expertise exploitable. Il faut encore déterminer l'intersection entre pièces et

dommages, quantifier la proportion de surface affectée et transformer cette quantité en un niveau de gravité interprétable.

1.2 Problématique de recherche

La problématique centrale de ce mémoire peut être formulée de la manière suivante : comment concevoir un système d'analyse d'images automobiles capable non seulement de détecter les pièces et les dommages, mais aussi d'associer correctement ces deux niveaux d'information afin de produire une évaluation graduelle, cohérente et compréhensible de l'état du véhicule ?

Cette question comporte plusieurs difficultés scientifiques et méthodologiques. Premièrement, la détection des pièces et la détection des dommages ne relèvent pas exactement du même problème. Les pièces du véhicule correspondent à des composants structurels ayant des formes relativement connues, tandis que les dommages sont des altérations visuelles plus irrégulières, parfois peu contrastées ou de faible taille. Deuxièmement, la simple présence simultanée d'une pièce et d'un dommage dans une image ne permet pas d'affirmer automatiquement que cette pièce est endommagée. Il faut une étape explicite d'association géométrique. Troisièmement, une fois cette association établie, le système doit encore traduire un recouvrement spatial, c'est-à-dire la zone commune partagée entre une pièce détectée et une zone de dommage détectée, en une notion de sévérité. Or, les seuils rigides de type « présent / absent » ou « faible / fort » ne rendent pas correctement compte de l'incertitude visuelle et de la gradation des dommages.

Un mémoire orienté exclusivement vers la détection aurait donc laissé de côté la partie la plus proche du raisonnement d'expertise. À l'inverse, un raisonnement flou appliqué sans base de détection suffisamment robuste aurait produit des décisions peu fiables. Le défi est donc double : disposer d'une base de détection cohérente et construire, au-dessus d'elle, un mécanisme d'interprétation qui reste simple, explicable et compatible avec les exigences de l'expertise en assurance automobile.

1.3 Objectifs du mémoire

L'objectif général de ce travail est de proposer un pipeline complet, pour l'évaluation automatisée des dommages automobiles à partir d'images. Ce pipeline repose sur deux détecteurs DETR entraînés séparément, sur une association géométrique fondée sur l'Intersection over Union, sur le calcul d'un Damage Ratio local par pièce, puis sur un module de logique floue destiné à interpréter progressivement la gravité.

Plus précisément, les objectifs spécifiques sont les suivants. D’abord, il s’agit de construire une branche de détection dédiée aux pièces du véhicule, capable d’identifier les principaux composants visibles d’un véhicule dans une image. Ensuite, de construire une seconde branche, indépendante, dédiée à la localisation des zones de dommages. Le troisième objectif consiste à relier les sorties de ces deux détecteurs par une mesure d’intersection réelle, afin de savoir quelles pièces sont touchées et dans quelle proportion. Le quatrième objectif est de convertir cette information géométrique en une représentation quantitative simple, le Damage Ratio, puis en une interprétation sémantique graduelle via la logique floue. Enfin, le dernier objectif consiste à évaluer la pertinence de cette architecture sur un jeu de données réel et à montrer en quoi elle améliore l’explicabilité du processus d’évaluation. Ces choix sont motivés par la nécessité de distinguer clairement la localisation visuelle, l’association géométrique et l’interprétation de la gravité, afin de rendre le système à la fois mesurable, traçable et compréhensible.

1.4 Contributions principales

Les contributions de ce mémoire se situent à l’interface entre détection d’objets, analyse géométrique et raisonnement flou. La première contribution est la formalisation d’un pipeline clair en quatre étapes (détection, intersection, ratio et raisonnement flou) qui rend explicite le rôle de chaque composant. La deuxième contribution est l’utilisation de deux détecteurs DETR spécialisés et entraînés séparément, ce qui évite de confondre la structure du véhicule et la localisation des dommages. La troisième contribution est l’introduction d’un Damage Ratio comme variable intermédiaire simple et interprétable, directement reliée à la proportion de surface d’une pièce touchée par un dommage. La quatrième contribution réside dans l’usage de la logique floue comme couche terminale d’interprétation, sans modification de l’apprentissage des détecteurs eux-mêmes. Une contribution complémentaire concerne l’étude d’une phase d’isolation du véhicule principal en amont des deux branches DETR, afin de réduire les ambiguïtés dans les scènes contenant plusieurs véhicules et de rendre le pipeline plus robuste en contexte réel.

Au-delà de ces éléments techniques, le travail apporte aussi une contribution méthodologique. Il montre qu’il est possible de construire une chaîne de traitement cohérente sans recourir à une accumulation d’heuristiques opaques. Le système ne se limite pas à une liste de détections ; il produit une lecture structurée de l’état du véhicule, pièce par pièce, puis au niveau global. Cette propriété est importante dans les contextes où la confiance dans la décision et la possibilité d’expliquer le résultat importent autant que la performance brute. L’apport de DIRF-R se distingue ainsi des travaux qui se concentrent uniquement sur la détection ou sur

l'intégration directe de la logique floue dans le modèle : ici, l'interprétation floue intervient après une association pièce-dommage explicite et vérifiable.

1.5 Organisation du mémoire

Le reste du document est structuré de la manière suivante. Le chapitre 2 présente l'état de l'art. Il traite des approches de détection des dommages, des méthodes de segmentation et de localisation, des limites en conditions réelles, de l'apport des modèles à base de Transformers et de l'intérêt de la logique floue, en cohérence avec le positionnement scientifique de ce travail.

Le chapitre 3 détaille la méthodologie proposée et présente le pipeline DIRF-R. Il expose d'abord les principes de DETR, puis la place exacte du module flou, avant de décrire successivement les quatre étapes du pipeline : Détection, Intersection, Ratio et Fuzzy Reasoning. Ce chapitre introduit également les mécanismes complémentaires utilisés pour l'isolation du véhicule principal, la stabilisation géométrique des détections et l'attribution gauche/droite des pièces.

Le chapitre 4 décrit le jeu de données, les choix d'annotation, les outils mobilisés, le protocole expérimental et les métriques retenues. Le chapitre 5 présente les résultats quantitatifs obtenus par les deux détecteurs DETR et les interprète à la lumière du pipeline complet. Le chapitre 6 discute les principales menaces à la validité de l'étude. Enfin, le chapitre 7 conclut le mémoire et présente plusieurs pistes d'amélioration et d'extension.

1.6 Délimitation du travail

Ce mémoire se concentre sur l'évaluation visuelle des dommages externes à partir d'images 2D annotées par boîtes englobantes. Il ne vise pas la reconstruction 3D du véhicule, l'estimation du coût de réparation, la détection de dommages internes ou la classification fine des types de dommages selon des catégories expertes détaillées. La branche dommages utilise, par choix méthodologique, une seule classe, afin de préserver la cohérence entre l'objectif de localisation et le rôle du module flou dans l'évaluation de gravité.

Cette délimitation renforce la lisibilité du travail. Elle évite d'attribuer au système des prétentions qu'il n'a pas. Le mémoire ne cherche pas à remplacer l'expert humain dans toute la chaîne de décision ; il propose une architecture d'aide à l'analyse, capable de structurer les informations visuelles et d'en fournir une lecture graduelle. Ce cadrage est essentiel pour interpréter correctement les résultats et pour situer les perspectives futures.

CHAPITRE 2 – État de l’art

2.1 Introduction générale sur l’état de l’art

L’automatisation de l’estimation des dommages sur les véhicules mobilise plusieurs objectifs complémentaires : détecter les zones endommagées à partir d’images [4], localiser les altérations par des boîtes englobantes [5], identifier les pièces concernées [7], relier correctement les dommages aux composants touchés [5], puis produire une estimation suffisamment fiable pour être utile en contexte applicatif [4]. La littérature dans ce domaine est abondante, mais elle reste fragmentée. Certains travaux se concentrent sur la classification de dommages [4], d’autres sur la segmentation de pièces [7], d’autres encore sur la robustesse en conditions réelles [8] ou sur l’interprétation des sorties du modèle par des approches hybrides [17]. Pour situer correctement l’approche proposée dans ce mémoire, il est nécessaire de parcourir ces contributions en distinguant ce qu’elles apportent et leurs limites.

Ce chapitre présente les familles d’approches les plus pertinentes pour DIRF-R : les travaux de classification des dommages [4], les méthodes de localisation des dommages par boîtes englobantes [5], les réseaux convolutionnels appliqués aux dommages automobiles [6], les approches de segmentation et de localisation des pièces [7], les enjeux de robustesse en conditions réelles [8], l’émergence des détecteurs à base de Transformers [1], puis l’apport de la logique floue pour l’interprétation [14]. Le positionnement retenu ici s’appuie sur DETR [1] et sur le raisonnement flou [14] comme base méthodologique de la solution proposée.

2.2 Travaux connexes sur la détection et la classification des dommages

Les premières recherches sur les dommages automobiles se sont principalement appuyées sur des réseaux convolutionnels de classification ou sur des détecteurs d’objets conventionnels. Les travaux de van Ruitenbeek et Bhulai [6] étudient l’usage des réseaux convolutionnels pour la détection des dommages automobiles à partir d’images. Mohammed et al. [4], dans une étude publiée en 2023, proposent une approche fondée sur l’apprentissage profond pour la classification des dommages et l’estimation des coûts. Qaddour et Siddiq [5] explorent également une méthode améliorée d’estimation automatique des dommages de véhicules. Ces travaux montrent que l’apprentissage profond peut capturer des motifs visuels pertinents, comme les rayures, les bosses, les fissures ou les déformations localisées [4]. Ils rappellent aussi que cette capacité dépend fortement de la disponibilité de jeux de données annotés de taille suffisante [6].

Un premier ensemble de travaux a porté sur la classification des dommages en plusieurs catégories. Mohammed et al. [4] montrent que des réseaux de type CNN (Convolutional Neural Network), entraînés à partir d'images de véhicules accidentés ou non accidentés, peuvent reconnaître des motifs visuels liés à des défauts spécifiques. Ces approches atteignent de bons scores de précision dans des environnements relativement contrôlés, notamment lorsqu'un transfert d'apprentissage depuis ImageNet est utilisé [4]. L'augmentation de données, combinant rotation, zoom, symétrie, variation de luminosité ou flou modéré, apparaît également comme un moyen d'améliorer la généralisation [6]. Néanmoins, leur principale limite réside dans le fait qu'elles apprennent souvent à reconnaître un état global de l'image, plutôt qu'à relier explicitement un dommage à une pièce du véhicule [5].

Un second ensemble de travaux s'est tourné vers la détection des dommages à l'aide de boîtes englobantes. Qaddour et Siddiqua [5] montrent l'intérêt d'une localisation plus opérationnelle que la simple classification. Les résultats restent toutefois dépendants de la qualité des annotations, de la variété des angles de prise de vue et de la distribution des classes [6]. Les dommages discrets, peu contrastés ou partiellement visibles restent souvent sous-détectés [6]. De plus, lorsque la tâche implique plusieurs types de dommages ou plusieurs composants dans une même image, les faux positifs et les ambiguïtés spatiales augmentent rapidement [5].

La littérature met donc en évidence un premier constat : localiser une zone de dommage ne suffit pas pour une expertise utile [5]. Un système pertinent doit aussi comprendre la structure du véhicule, autrement dit savoir sur quelle pièce le dommage apparaît [7]. Cette exigence explique pourquoi les travaux sur la segmentation et la localisation des pièces occupent une place importante dans l'état de l'art [7].

2.3 Travaux connexes sur la segmentation et la localisation des pièces et des dommages

L'identification des pièces du véhicule a donné lieu à plusieurs travaux fondés sur la segmentation sémantique et l'instance segmentation. L'étude de Pasupa et al. [7], publiée en 2022, évalue des architectures comme Mask R-CNN (Region-based Convolutional Neural Network), HTC (Hybrid Task Cascade), PANet (Path Aggregation Network), CBNet (Composite Backbone Network) ou GCNet (Global Context Network) sur des ensembles d'images annotées contenant des pièces telles que les portières, les ailes, les vitres, les phares, les rétroviseurs ou les pare-chocs. Ces travaux montrent qu'une modélisation plus fine du contour des objets améliore la compréhension structurale du véhicule [7], notamment lorsque

plusieurs composants se chevauchent partiellement ou que les conditions de prise de vue sont complexes.

Les résultats les plus intéressants de cette famille de travaux concernent la possibilité d'annoter les pièces avec une granularité assez fine, y compris la distinction entre côté gauche et côté droit [7]. Cette précision topologique devient précieuse lorsqu'on cherche à associer un dommage à un composant précis. En revanche, ces méthodes ont souvent un coût computationnel élevé, et leurs performances se dégradent lorsque le jeu de données est limité ou très déséquilibré [7]. En outre, beaucoup de travaux se concentrent sur la structure visible du véhicule, sans relier explicitement cette structure à des zones de dommages [7].

D'autres recherches ont proposé des systèmes à double branche, une première pour repérer les côtés ou les composants du véhicule, et une seconde pour localiser les altérations visuelles [5]. L'intérêt de cette séparation est méthodologiquement proche de celui retenu dans DIRF-R : au lieu d'entraîner un système à tout faire en une seule étape, on distingue la structuration du véhicule de la détection des dommages. Cette dissociation facilite ensuite l'interprétation, car elle permet de relier explicitement une zone endommagée à une pièce donnée [5]. Toutefois, dans la majorité de ces travaux, l'étape d'association reste peu formalisée. Elle repose souvent sur des heuristiques implicites, sur une proximité spatiale non justifiée ou sur un simple recouvrement observé à l'œil [5].

Pour notre problématique, la leçon essentielle de ces travaux est la suivante : une bonne évaluation des dommages automobiles exige de combiner la localisation des altérations [5] avec une représentation structurée des pièces [7]. Le simple repérage de zones anormales ne suffit pas ; il faut encore construire une relation explicite entre ces zones et les composants du véhicule [7].

2.4 Travaux connexes pour la robustesse et les performances en conditions réelles

Un autre axe majeur de la littérature concerne la robustesse des systèmes en conditions réelles. Les travaux de van Ruitenbeek et Bhulai [6] rappellent que les variations d'angle et de qualité visuelle peuvent affecter la détection des dommages. Les études portant sur la segmentation des pièces soulignent aussi la difficulté liée aux arrière-plans complexes et aux scènes multi-véhicules [7]. De leur côté, Ullah et Yow [8] insistent sur l'effet des reflets, des ombres et des conditions d'éclairage variables. Dans ce type de contexte, plusieurs méthodes qui fonctionnent bien sur un jeu de données propre perdent rapidement en précision [8].

Ullah et Yow [8], dans une étude publiée en 2025, montrent que les réflexions spéculaires sur la carrosserie peuvent tromper les modèles, en particulier lorsque ceux-ci apprennent des corrélations visuelles superficielles plutôt qu'une vraie géométrie des dommages. Van Ruitenbeek et Bhulai [6] mettent aussi en évidence la fragilité des systèmes face aux petits objets ou aux dommages partiels. Une légère variation d'angle suffit parfois à modifier fortement la détection [6]. Enfin, la présence simultanée de plusieurs véhicules dans une image peut conduire le modèle à attribuer des pièces secondaires au véhicule principal, ce qui perturbe l'évaluation globale [7].

Ces observations sont importantes pour le présent travail, car elles justifient l'ajout de mécanismes complémentaires autour des détecteurs. L'isolation du véhicule principal répond aux difficultés liées aux scènes multi-véhicules [7]. La consolidation des détections redondantes et la cohérence gauche/droite prolongent également des préoccupations méthodologiques déjà présentes dans les travaux sur l'interprétation structurée des dommages [17]. Un modèle utile doit donc être pensé comme une chaîne complète de traitement et non comme un seul algorithme de détection isolé [5].

2.5 Travaux connexes concernant les modèles à base de Transformers et les approches efficaces

L'introduction des Transformers dans la vision par ordinateur a marqué une évolution importante dans la détection d'objets. La revue de Li et al. [9], publiée en 2023, présente un panorama des détecteurs d'objets fondés sur les Transformers. Le modèle DETR de Carion et al. [1], publié en 2020, reformule la détection comme une tâche de prédiction d'ensemble de bout en bout, associant les prédictions aux annotations par appariement bipartite optimal. Cette formulation simplifie le pipeline en supprimant plusieurs heuristiques classiques, notamment la dépendance aux ancres prédéfinies et la multiplication des étapes de post-traitement [1]. L'un des principaux avantages de DETR tient à sa capacité de modélisation globale : l'attention introduite dans les architectures Transformer par Vaswani et al. [3] lui permet d'exploiter le contexte complet de l'image, ce qui est particulièrement utile lorsque plusieurs pièces et plusieurs dommages coexistent.

La première version de DETR présentait néanmoins deux limites bien connues : une convergence lente et une sensibilité aux petits objets [1]. Pour répondre à ces difficultés, Zhu et al. [10] proposent Deformable DETR, qui concentre l'attention sur un nombre restreint de points pertinents à différentes échelles et améliore la prise en compte des objets de petite taille. DAB-DETR, introduit par Liu et al. [11], utilise des requêtes dynamiques guidées par des boîtes et

fournit au décodeur des repères spatiaux plus explicites. DN-DETR, proposé par Li et al. [12], accélère l'apprentissage en ajoutant une stratégie de débruitage des requêtes. Conditional DETR, présenté par Meng et al. [13], permet quant à lui à chaque requête de se focaliser plus efficacement sur une région spécifique de l'image.

Ces variantes n'ont pas pour effet d'annuler l'intérêt du modèle original ; elles montrent plutôt qu'un cadre Transformer est suffisamment riche pour être adapté à des cas d'usage exigeants [9]. Pour notre mémoire, elles jouent deux rôles. D'une part, elles confirment la pertinence d'un positionnement DETR pour analyser la structure d'un véhicule et les zones de dommages [1]. D'autre part, elles rappellent que les métriques de détection ne suffisent pas, à elles seules, à fournir une expertise. Même un excellent détecteur produit des sorties numériques qui demandent ensuite à être interprétées [8].

2.6 Apport de la logique floue et synthèse des lacunes

La logique floue apporte précisément ce niveau d'interprétation. Les travaux fondateurs de Zadeh [14], publiés en 1965, introduisent la théorie des ensembles flous pour modéliser des phénomènes graduels, ambigus ou partiellement vrais. Dans un système flou, une variable ne bascule pas brutalement d'une catégorie à une autre ; elle peut appartenir à plusieurs catégories avec des degrés différents [15]. Yager et Filev [16], dans leurs travaux sur les systèmes flous, montrent l'intérêt de cette approche pour la modélisation et le contrôle de systèmes où l'information est incertaine. Cette propriété est particulièrement utile lorsque l'on cherche à traduire un indicateur continu, par exemple la proportion de surface d'une pièce affectée, en niveaux de gravité compréhensibles [16].

Dans le domaine de la vision par ordinateur, la logique floue a déjà été exploitée pour améliorer la segmentation, fusionner plusieurs indices ou interpréter des sorties incertaines [15]. La méthode de Ullah et Yow [8], publiée en 2025, applique une prise de décision floue à la détection de bosses automobiles afin de mieux traiter des situations où les indices visuels sont faibles. Son intérêt principal n'est pas de remplacer la détection, mais de compléter celle-ci par une étape de décision plus progressive et plus explicable [8]. Dans cette continuité, Kogbeti-Wongo [17] propose FuzzyYOLOv8, une variante de YOLOv8 enrichie par la logique floue afin d'assouplir les décisions de détection, d'intégrer la classification gauche/droite et de mieux interpréter la gravité des dommages. Ce travail confirme l'intérêt d'une hybridation entre détection visuelle et raisonnement flou, tout en restant méthodologiquement distinct de DIRF-R, qui sépare explicitement la détection par DETR et l'interprétation floue [17].

La synthèse de la littérature met donc en évidence quatre lacunes majeures. Premièrement, beaucoup de travaux détectent les dommages sans représenter clairement la structure du véhicule. Deuxièmement, lorsqu'une structuration existe, l'association entre pièce et dommage reste souvent implicite ou insuffisamment justifiée. Troisièmement, les conditions réelles de prise de vue dégradent rapidement la fiabilité des prédictions. Quatrièmement, la gravité du dommage est encore fréquemment abordée comme une décision rigide, alors qu'elle relève davantage d'une interprétation graduelle. C'est précisément pour répondre à ces limites que nous proposons DIRF-R : un pipeline structuré autour de deux détecteurs DETR spécialisés, d'une association géométrique explicite basée sur l'IoU, du calcul d'un Damage Ratio, puis d'une interprétation floue progressive de la gravité du dommage.

2.7 Positionnement spécifique de DIRF-R

Au regard de cette revue, le positionnement de DIRF-R peut être résumé en trois choix. Le premier est un choix de structure : plutôt qu'un modèle unique supposé résoudre simultanément toutes les dimensions du problème, nous retenons deux détecteurs spécialisés et une chaîne d'interprétation explicite. Le deuxième est un choix de représentation : la relation entre une pièce et un dommage est fondée sur une mesure géométrique simple, calculable et vérifiable, à savoir l'IoU. Le troisième est un choix de décision : la gravité n'est pas fixée par une règle dure, mais par une interprétation floue du Damage Ratio.

Ce positionnement s'inscrit dans la continuité de l'état de l'art tout en s'en distinguant clairement. Il s'inspire des travaux sur la structuration des pièces, des travaux sur la détection de dommages et des travaux sur la gestion de l'incertitude, mais il les articule dans un pipeline unique. C'est cette articulation qui constitue la singularité du mémoire : au lieu d'ajouter un module flou directement dans le détecteur, ou d'utiliser un score global difficile à interpréter, DIRF-R conserve une chaîne d'étapes simples, chacune justifiable indépendamment.

CHAPITRE 3 – Méthodologie proposée

3.1 DETR comme base de détection

Dans le cadre de ce travail, la base de détection repose sur DETR, proposé par Carion et al. [1] comme modèle de détection d’objets de bout en bout fondé sur l’architecture Transformer introduite par Vaswani et al. [3]. L’intérêt de DETR, pour notre problématique, tient à sa capacité à traiter l’image comme une scène globale plutôt que comme une simple juxtaposition de régions locales. Là où des architectures plus anciennes s’appuient fortement sur des ancres prédéfinies et des heuristiques de suppression des doublons, DETR apprend directement à produire un ensemble cohérent de prédictions.

Le principe général peut être résumé simplement. Une image est d’abord transformée en une représentation visuelle compacte à l’aide d’un backbone convolutionnel. Cette représentation est ensuite transmise à un encodeur Transformer, qui modélise les relations globales entre toutes les régions de l’image. Un décodeur Transformer manipule ensuite un ensemble fixe de requêtes d’objets apprises. Chaque requête correspond à une hypothèse d’objet potentiel. À la sortie, le modèle produit des classes et des boîtes englobantes normalisées. Pendant l’apprentissage, un appariement bipartite optimal garantit qu’une prédiction donnée correspond au plus à un objet annoté, ce qui réduit naturellement les doublons.

Dans le contexte de l’analyse automobile, cette propriété est précieuse. Une image de véhicule contient souvent plusieurs composants visuellement proches, parfois partiellement occultés ou déformés par la perspective. Un modèle qui raisonne avec une vision globale de la scène peut mieux prendre en compte la cohérence structurelle entre les pièces. C’est pourquoi DETR constitue ici le socle de la phase Détection de DIRF-R.

3.1.1 Architecture interne de DETR

L’architecture interne de DETR peut être décrite à travers quatre blocs principaux : le backbone convolutionnel, l’encodeur Transformer, le décodeur Transformer et la tête de prédiction. Le backbone, de type ResNet-50 dans cette implémentation, extrait des caractéristiques visuelles à partir de l’image brute. L’encodeur Transformer reçoit ensuite ces caractéristiques, enrichies par un codage positionnel, et apprend des dépendances globales entre les différentes régions de l’image. Le décodeur manipule un nombre fixe de requêtes d’objets, qui sont affinées couche après couche. Enfin, la tête de prédiction produit, pour chaque requête, une classe et une boîte englobante.

Afin de mieux comprendre le fonctionnement général du modèle, trois propriétés méritent d'être soulignées. Premièrement, DETR ne parcourt pas l'image de manière purement locale ; il construit une représentation de scène. Deuxièmement, le modèle apprend à gérer un ensemble d'objets plutôt qu'une séquence ordonnée. Troisièmement, l'apprentissage par appariement bipartite lui permet de simplifier la phase d'inférence en limitant le recours à des heuristiques externes.

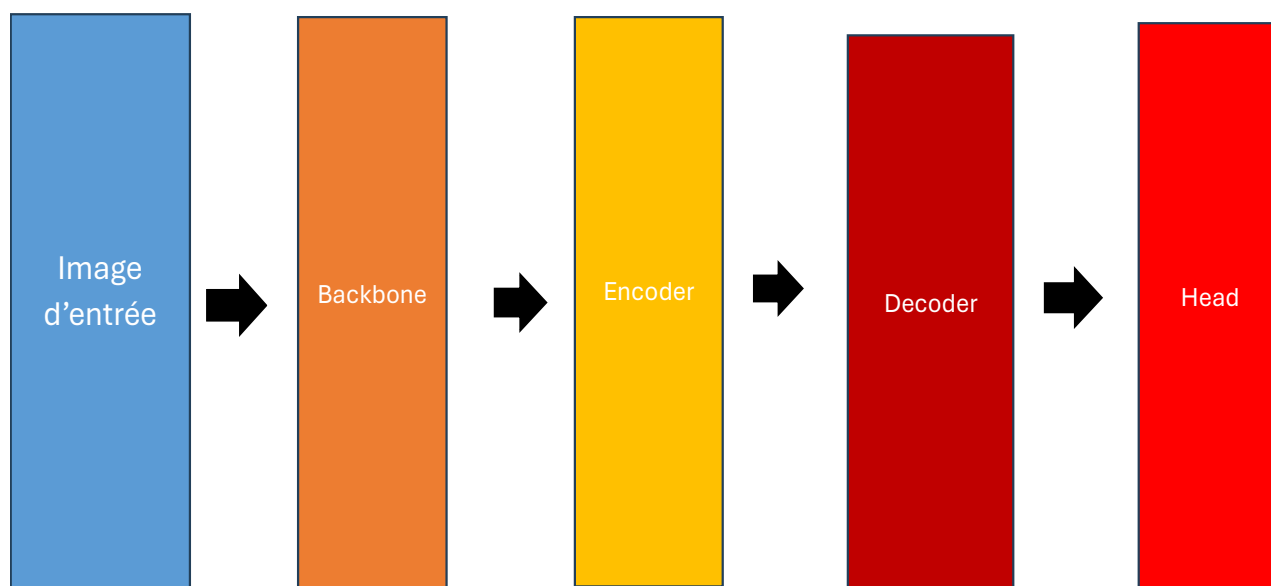


Figure 1 : Architecture de DETR.

3.1.2 Deux détecteurs spécialisés dans DIRF-R

DIRF-R n'utilise pas un seul détecteur pour toutes les tâches. Le pipeline repose sur deux détecteurs DETR distincts et entraînés séparément. Le premier est spécialisé dans la détection des pièces du véhicule. Son rôle est d'identifier les composants structurels visibles, tels que les portières, les ailes, le capot, les phares, le pare-brise ou les rétroviseurs. Le second détecteur est spécialisé dans la localisation des dommages et ne travaille que sur une seule classe : « Endommagée ». Cette approche permet de concentrer la branche dommages sur l'identification des zones affectées, sans imposer une catégorisation détaillée des types de dommages. L'évaluation de la gravité est ensuite assurée par le calcul du Damage Ratio et par le module de logique floue, qui fournissent une interprétation graduelle des dommages détectés.

Cette séparation est essentielle pour préserver la clarté méthodologique du système. Les pièces et les dommages ne possèdent ni la même stabilité géométrique ni la même régularité

visuelle. Les pièces ont des formes plus récurrentes et une structure relativement stable ; les dommages, eux, peuvent être diffus, très localisés, partiellement visibles ou de forme irrégulière. En dissociant les deux tâches, on laisse à chaque détecteur la possibilité d'apprendre sa propre distribution sans le forcer à résoudre plusieurs problèmes hétérogènes simultanément.

Ainsi, la phase Détection de DIRF-R ne doit pas être comprise comme une détection globale unique, mais comme la combinaison ordonnée de deux inférences indépendantes, dont les sorties seront ensuite rapprochées par l'étape d'Intersection.

3.2 Logique floue : concepts, intérêt et place dans le pipeline

La logique floue, introduite par Zadeh [14], constitue un cadre adapté à la modélisation de situations où la frontière entre les catégories n'est pas rigide [15]. Dans une logique classique, une proposition est vraie ou fausse ; dans une logique floue, elle peut être vraie à un certain degré. Concrètement, cela signifie qu'une pièce peut être considérée à la fois « légèrement » et « modérément » endommagée avec des degrés d'appartenance différents, plutôt que d'être assignée brutalement à une seule catégorie.

Sur le plan formel, un ensemble flou A défini sur un univers U est caractérisé par une fonction d'appartenance qui associe à chaque élément x un degré compris entre 0 et 1 [14]. Cette écriture résume l'idée centrale de la logique floue : représenter progressivement l'appartenance plutôt que de l'imposer de manière binaire [15].

$$\mu_A : U \rightarrow [0,1] \quad \text{et} \quad A = \{(x, \mu_A(x)) \mid x \in U\} \quad (1)$$

À titre de rappel théorique, une règle floue peut ensuite s'écrire sous la forme : « Si x est A_i , alors y est B_i ». En pratique, plusieurs règles sont agrégées avant une étape éventuelle de défuzzification. Dans ce mémoire, cette présentation reste volontairement générale et pédagogique.

Dans le cas des dommages automobiles, cette souplesse est particulièrement utile. Une faible variation de recouvrement entre une pièce et une zone de dommage ne devrait pas provoquer un saut brutal de décision. De plus, les conditions visuelles réelles introduisent de l'incertitude : un dommage peut être visible seulement sous un certain angle, une partie de la pièce peut être masquée, ou le contour exact du dommage peut rester imprécis. Une logique de seuils stricts reste alors trop pauvre pour traduire le raisonnement d'un expert.

Il faut toutefois souligner un point central pour la cohérence scientifique du mémoire : dans DIRF-R, la logique floue n'intervient ni dans l'apprentissage des deux détecteurs DETR,

ni dans le calcul géométrique de l'intersection. Elle agit exclusivement après ces étapes, au moment de l'interprétation. Son rôle n'est donc pas de corriger un détecteur ou de remplacer l'IoU ; il consiste à transformer un indicateur continu, le Damage Ratio, en niveaux de gravité progressifs et plus explicables.

3.3 Vue d'ensemble du pipeline DIRF-R

Le nom DIRF-R résume exactement l'enchaînement retenu dans ce mémoire : Détection, Intersection, Ratio et Fuzzy Reasoning. Ce découpage a une valeur plus forte qu'un simple acronyme. Il matérialise une architecture où chaque étape repose sur une entrée claire, produit une sortie identifiable et joue un rôle méthodologiquement distinct.

Avant les étapes principales du pipeline, une phase optionnelle d'isolation du véhicule principal peut être appliquée afin de limiter les interférences causées par la présence de plusieurs véhicules dans une même scène. Cette étape repose sur un détecteur de véhicules chargé d'identifier l'ensemble des véhicules visibles dans l'image. Le véhicule associé à la plus grande boîte englobante est alors considéré comme le véhicule principal et transmis au pipeline DIRF-R pour les étapes d'analyse détaillée.

La première étape, Détection, consiste à obtenir les prédictions des deux détecteurs DETR spécialisés : une liste de pièces et une liste de zones de dommages. La deuxième étape, Intersection, relie ces deux ensembles à l'aide d'une mesure géométrique réelle, l'Intersection over Union. La troisième étape, Ratio, transforme cette relation spatiale en une proportion de surface touchée par pièce. Enfin, la quatrième étape, Fuzzy Reasoning, interprète le ratio obtenu au moyen de règles et de fonctions d'appartenance floues pour produire des niveaux de gravité. Une étape finale de consolidation géométrique et de filtrage permet ensuite de stabiliser les détections avant la sortie finale du pipeline.

La grande force de cette organisation est la traçabilité. À tout moment, on peut expliquer le passage d'une image brute à une conclusion. On sait quel véhicule a été isolé, quelles pièces ont été détectées, quels dommages ont été localisés, quelle intersection a été calculée, quel ratio en découle et pourquoi le système aboutit à un niveau de gravité donné.

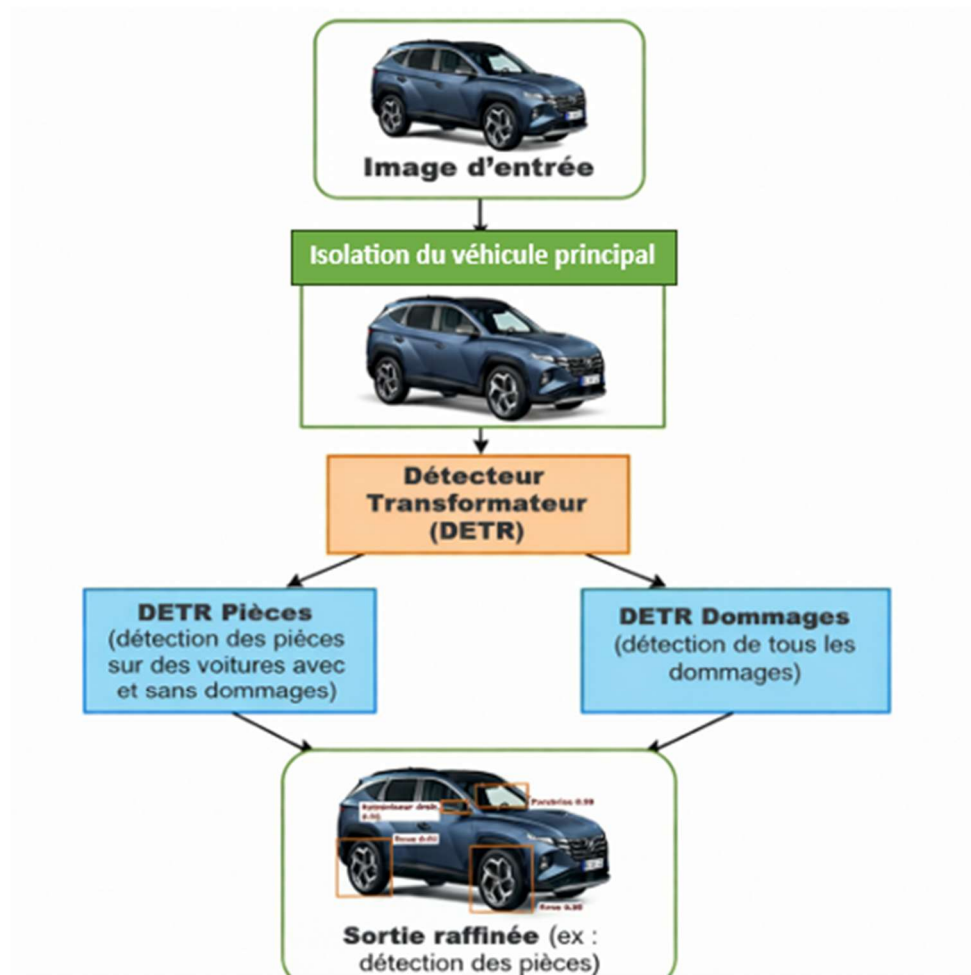


Figure 2 : Vue d'ensemble de l'architecture DIRF-R

3.4 Détection : sortie des détecteurs de pièces et de dommages

Pour une image d'entrée donnée, le détecteur des pièces retourne un ensemble de prédictions de la forme : (pièce_i, boîte_i, confiance_i)

Le détecteur des dommages retourne un ensemble plus simple : (dommage_j, boîte_j, confiance_j)

Chaque boîte est définie par ses coordonnées dans l'image. À ce stade, le système sait où se trouvent les pièces et où se trouvent les dommages, mais il ne sait pas encore quelles pièces sont réellement affectées. C'est précisément ce manque d'information relationnelle qui justifie l'étape suivante.

La séparation des détecteurs apporte aussi un avantage analytique. Elle permet d'évaluer différemment les performances des deux branches. Une mauvaise détection de pièces n'a pas la même signification qu'une mauvaise détection de dommages. En distinguant les deux, on obtient une lecture plus fine des forces et des limites du pipeline.

3.5 Intersection : association géométrique explicite par IoU

L'Intersection over Union (IoU) est une mesure géométrique de recouvrement entre deux régions, largement utilisée en vision par ordinateur pour comparer une prédiction à une annotation de référence. Son usage s'est imposé progressivement comme critère central d'évaluation dans les benchmarks modernes de détection d'objets, notamment avec la diffusion des protocoles d'évaluation de type PASCAL [18].

Au fil de l'évolution des travaux, plusieurs variantes de l'IoU ont été proposées pour mieux traiter certaines limites de la version classique. La GIoU (Generalized IoU) introduit une pénalisation lorsque deux boîtes ne se recouvrent pas [19], alors que la DIoU (Distance IoU) et la CIoU (Complete IoU) ajoutent respectivement des informations sur la distance entre centres et sur le rapport d'aspect des boîtes [20]. Ces variantes sont particulièrement utiles lorsqu'on cherche à optimiser l'apprentissage ou à améliorer la régression des boîtes englobantes.

Dans ce mémoire, la version retenue est toutefois l'IoU classique. Ce choix est volontaire : ici, l'objectif n'est pas de définir une nouvelle fonction de perte ni d'améliorer l'entraînement des détecteurs, mais d'établir une relation géométrique simple, stable et interprétable entre une pièce détectée et une zone de dommage. L'IoU standard est donc la plus cohérente avec le rôle d'association explicite que joue l'étape « Intersection » dans DIRF-R.

L'étape d'association repose sur l'Intersection over Union, notée IoU. Cette mesure compare la surface d'intersection de deux boîtes à la surface de leur union. Pour une pièce P_i et un dommage D_j , on définit :

$$IoU(P_i, D_j) = \frac{Aire(P_i \cap D_j)}{Aire(P_i \cup D_j)} \quad (2)$$

Où P_i représente la boîte englobante de la pièce i , et D_j la boîte englobante du dommage j

Cette formule est volontairement simple. Elle ne cherche pas à produire un raisonnement flou ; elle mesure objectivement le recouvrement spatial entre une pièce et une zone de

dommage. Si l'IoU est nulle, la pièce et le dommage ne se recouvrent pas. Si elle est positive, on dispose d'un indice clair qu'un lien spatial existe entre eux.

Dans DIRF-R, l'IoU n'est donc pas un critère d'évaluation externe uniquement utilisé pour calculer des métriques ; elle joue un rôle interne dans le pipeline. Elle sert à établir la relation entre structure et altération. Cette distinction est importante pour éviter une confusion fréquente. Ici, l'IoU intervient comme mécanisme d'association pièce-dommage, alors que la logique floue n'intervient qu'après cette association.

3.6 Ratio : calcul du Damage Ratio local

Une fois l'association géométrique établie, il faut quantifier la proportion de la pièce réellement touchée. C'est le rôle du Damage Ratio. Pour une pièce P_i associée à une ou plusieurs zones de dommage, on définit :

$$DR_i = \frac{\text{Aire}(U_{j=1}^n(P_i \cap D_j))}{\text{Aire}(P_i)} \quad (3)$$

Où $(U_{j=1}^n(P_i \cap D_j))$ représente la zone de dommage ou l'union des zones de dommages retenues pour la pièce P_i .

Le Damage Ratio varie entre 0 et 1. Une valeur proche de 0 indique qu'une très petite portion de la pièce est affectée ; une valeur proche de 1 signifie qu'une part importante de sa surface visible est endommagée. Cette variable joue un rôle central dans DIRF-R, car elle assure la transition entre la géométrie et l'interprétation. Tant que l'on reste au niveau de l'IoU, on manipule une relation spatiale. Dès que l'on calcule le Damage Ratio, on dispose d'un indicateur directement exploitable pour raisonner sur la gravité du dommage.

Pour rendre le système plus compréhensible, on peut associer ce ratio à des exemples simples. Un Damage Ratio de 0,05 correspond à une atteinte très localisée, souvent visuellement faible. Un ratio autour de 0,25 peut correspondre à un dommage visible mais encore limité à une portion de la pièce. Un ratio supérieur à 0,50 traduit généralement une atteinte importante qui affecte une large zone de la pièce. Ces exemples ne remplacent pas le raisonnement flou ; ils en préparent l'interprétation.

Dans l'implémentation utilisée, le calcul du Damage Ratio suit directement la logique du code expérimental. Pour chaque boîte englobante, l'aire de la pièce est calculée à partir de ses coordonnées par le produit de sa largeur et de sa hauteur. L'intersection entre une pièce détectée et une boîte de dommage est obtenue en conservant les coordonnées communes aux

deux boîtes ; lorsqu'il n'existe aucun recouvrement, l'aire d'intersection est fixée à zéro. Le pourcentage de dommage est ensuite calculé uniquement si l'aire de la pièce est strictement positive. Une pièce est considérée comme touchée lorsque l'aire d'intersection entre sa boîte et celle du dommage est supérieure à zéro. Cette procédure conserve la formule théorique du Damage Ratio tout en précisant son passage opérationnel dans le pipeline. Lorsque le résultat est exprimé en pourcentage, on utilise $DR_i(\%) = 100 \times DR_i$

Il convient toutefois de préciser la limite principale de ce calcul. Le Damage Ratio est estimé à partir de boîtes englobantes et non de masques de segmentation. Par conséquent, la surface utilisée pour le calcul peut inclure des pixels non endommagés situés à l'intérieur de la boîte de dommage ou de la boîte de la pièce. Cette approximation est particulièrement sensible dans le cas des rayures fines, des dommages diagonaux, des bosses localisées ou des déformations irrégulières. Le ratio obtenu doit donc être interprété comme une mesure géométrique approximative du recouvrement visible entre une zone de dommage détectée et une pièce, et non comme une estimation pixel à pixel de la surface réellement abîmée. Une extension naturelle du travail consisterait à remplacer ou compléter les boîtes englobantes par des masques de segmentation, afin d'obtenir une mesure de surface plus fidèle.

3.7 Raisonnement flou : interprétation graduelle de la gravité

Le Damage Ratio calculé à l'étape précédente constitue l'entrée principale du module flou. Dans sa forme la plus simple, ce module définit plusieurs catégories linguistiques de gravité, par exemple : mineure, modérée, sévère et très sévère. À chacune de ces catégories correspond une fonction d'appartenance. Une pièce donnée peut donc appartenir partiellement à plusieurs niveaux de gravité.

Le choix de fonctions d'appartenance triangulaires ou trapézoïdales permet de conserver une interprétation simple. Dans le cadre de ce mémoire, le raisonnement flou reste volontairement lisible. Il ne s'agit pas de construire un système d'inférence trop lourd, mais de fournir une couche de décision progressive. Par exemple, un Damage Ratio proche de 0,20 peut appartenir faiblement à la catégorie « mineur » et plus fortement à la catégorie « modéré ». Le système évite ainsi l'effet de seuil abrupt.

Dans le code expérimental, les fonctions d'appartenance retenues sont triangulaires et définies sur un univers de 0 à 100. La catégorie « mineur » utilise les points [0, 0, 30], la catégorie « modéré » les points [20, 40, 60], la catégorie « sévère » les points [50, 70, 90] et la catégorie « très sévère » les points [70, 100, 100]. Les mêmes fonctions sont utilisées pour la variable de sortie « severity ». Les règles floues sont directes : si le pourcentage de dommage est mineur, alors la sévérité est mineure ; s'il est modéré, alors la sévérité est moyenne ; s'il est

sévère, alors la sévérité est sévère ; et s'il est très sévère, alors la sévérité est très sévère. Ce choix reste simple, mais il rend le raisonnement vérifiable et cohérent avec les sorties du programme.

Pour l'affichage local des résultats, une catégorisation complémentaire est également utilisée dans le script : un pourcentage inférieur ou égal à 25 % est indiqué comme dommage mineur, un pourcentage compris entre 25 % et 50 % comme dommage modéré, un pourcentage compris entre 50 % et 75 % comme dommage sévère, et un pourcentage supérieur à 75 % comme dommage très sévère. Cette catégorisation sert surtout à rendre les sorties lisibles pour l'utilisateur, tandis que l'évaluation globale repose sur l'inférence floue appliquée à la moyenne des pourcentages de dommages détectés.

Afin de synthétiser dans une même séquence visuelle les éléments clés du raisonnement flou et de l'exploitation du pipeline, les figures suivantes sont regroupées dans cette section.

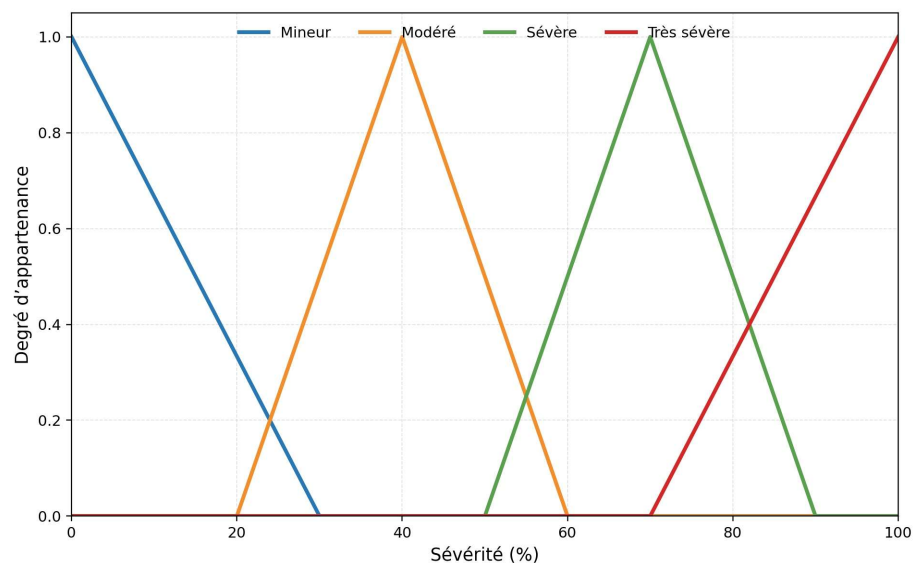


Figure 3 : Fonctions d'appartenance des niveaux de gravité

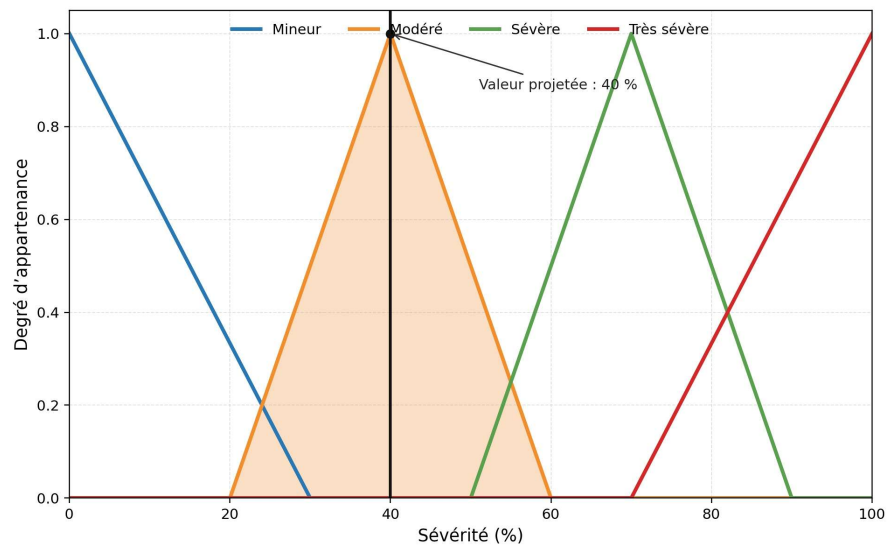


Figure 4 : Projection d'une valeur de sévérité floue



Figure 5 : Sortie brute des détecteurs de pièces et de dommages — vert : pièces ; rouge : dommages.

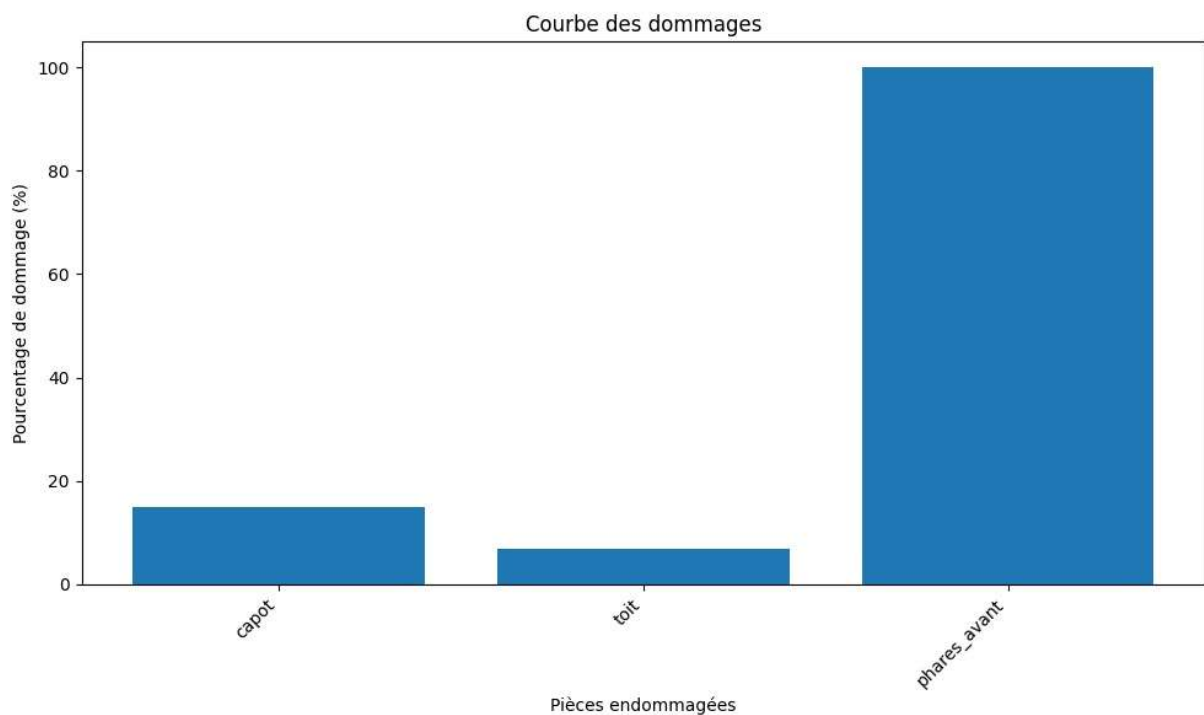


Figure 6 : Pourcentages de dommages par pièce détectée.

À titre d'approximation descriptive, on peut lire les niveaux de gravité selon les repères suivants :

- Capot 14,94 % endommagé Dommage mineur
- Toit 6,78 % endommagé Dommage mineur
- Phare avant gauche 100,00 % endommagé Dommage très sévère

Ces repères n'ont pas vocation à remplacer les fonctions d'appartenance elles-mêmes. Ils servent à rendre le raisonnement lisible pour un lecteur humain. Le point important, scientifiquement, est que la décision finale ne repose pas sur une coupure binaire stricte, mais sur une interprétation progressive d'un indicateur continu.

3.8 Évaluation locale et globale

L'évaluation locale est produite pièce par pièce. Chaque pièce détectée reçoit un Damage Ratio et une interprétation floue de gravité. On obtient ainsi une description structurée du véhicule, du type : « Porte avant gauche modérément endommagée », « Pare-chocs avant sévèrement endommagé », « Aile droite intacte », etc. Cette granularité est particulièrement utile

dans les contextes d'expertise, car elle permet de discuter des réparations composant par composant.

Pour obtenir une lecture globale de l'état du véhicule, il est possible d'agréger les ratios locaux. Une forme simple d'agrégation consiste à utiliser une moyenne pondérée :

$$ScoreGlobal = \frac{\sum_{i=1}^N w_i \times DR_i}{\sum_{i=1}^N w_i} \quad (4)$$

Où w_i représente le poids attribué à la pièce i , et DR_i le Damage Ratio associé à cette pièce. Dans ce mémoire, cette pondération est présentée comme une extension naturelle du pipeline. Elle permettrait, par exemple, de donner plus d'importance à des pièces structurales ou coûteuses à réparer qu'à des éléments secondaires. Même lorsque l'agrégation reste simple, elle rend possible une synthèse finale de l'état du véhicule sans perdre entièrement l'information locale.

3.9 Robustesse structurelle : isolation du véhicule principal et stabilisation

Dans des images réelles, le véhicule étudié n'est pas toujours seul. Il peut apparaître dans un stationnement, sur la voie publique ou devant un atelier, avec d'autres véhicules visibles en arrière-plan. Si le pipeline exploitait directement toutes les détections produites, des pièces appartenant à un véhicule secondaire pourraient être associées à des dommages du véhicule principal, ou inversement. Pour éviter ce problème, DIRF-R intègre un mécanisme d'isolation du véhicule principal.

Le premier mécanisme étudié repose sur une hypothèse de taille apparente. Les pièces du véhicule principal occupent généralement une aire plus importante dans l'image que celles des véhicules situés à l'arrière-plan. Un regroupement par aire, fondé sur un seuil relatif entre l'aire de la boîte et l'aire de l'image, permet alors d'identifier l'ensemble dominant [17]. Le second mécanisme repose sur une région d'intérêt (ROI – Region of Interest) centrée. Dans cette version du pipeline, la ROI est définie de manière relative à l'image : elle couvre la zone centrale principale du cadre et peut être ajustée selon la résolution et le cadrage. Dans les prises de vue standards, le véhicule à expertiser se trouve le plus souvent au centre du cadre. En combinant ces deux critères, on écarte une partie significative des détections parasites. Cette approche reste toutefois fondée sur des hypothèses de cadrage et de perspective. Elle devient moins fiable lorsque le véhicule principal est fortement décentré ou lorsqu'un autre véhicule occupe une surface visuelle dominante dans l'image.

Afin d'améliorer la robustesse de cette étape, une seconde stratégie plus structurée a également été étudiée. Cette approche consiste à utiliser un détecteur de véhicules en amont du pipeline principal. Le détecteur identifie d'abord tous les véhicules présents dans l'image à l'aide de boîtes englobantes. Le système conserve ensuite uniquement le véhicule associé à la plus grande boîte englobante, considéré comme le véhicule principal de la scène. La région correspondante est alors recadrée puis transmise au pipeline DIRF-R pour les étapes de détection des pièces, d'association géométrique et d'interprétation floue. Cette stratégie permet de réduire les ambiguïtés dans les scènes contenant plusieurs véhicules et améliore la robustesse globale du pipeline en conditions réelles.

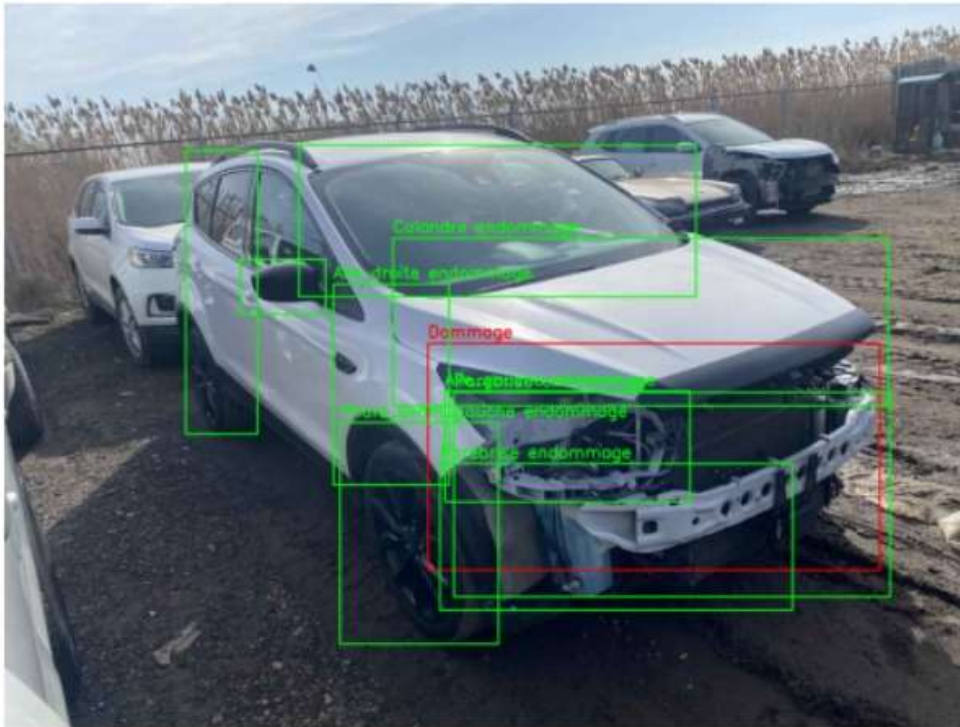
Une fois le véhicule principal isolé, il reste nécessaire de stabiliser les détections. Certaines pièces peuvent être détectées plusieurs fois ou partiellement fragmentées. Une consolidation géométrique, inspirée du principe de Non-Maximum Suppression (NMS), permet de fusionner les boîtes très redondantes lorsque leur IoU dépasse un seuil fixé, et d'éliminer des régions trop petites ou peu plausibles. Le filtrage des petites régions est appliqué de façon relative, à partir du rapport entre l'aire de la boîte et l'aire de l'image, afin d'éviter de conserver des détections visuellement négligeables. Cette étape ne modifie pas le principe de DIRF-R ; elle assure simplement que l'étape d'intersection repose sur une représentation plus propre et plus cohérente de la scène.

3.10 Distinction gauche/droite des pièces

La capacité à distinguer les côtés gauche et droit constitue un élément pratique important du pipeline. Dans un contexte d'assurance ou de réparation, il ne suffit pas de dire qu'une portière est endommagée ; il faut savoir s'il s'agit de la portière avant gauche, arrière droite, du rétroviseur droit ou de l'aile gauche. DIRF-R introduit donc une règle de classification latérale fondée sur la position horizontale du centre d'une pièce.

Soit x_c la coordonnée horizontale du centre de la pièce à classer et x_{ref} celle d'un point de référence central. Lorsque le pare-chocs avant ou un autre composant central robuste est disponible, son centre sert de référence. À défaut, on peut utiliser le centre de l'image. La décision s'écrit :

$$\text{Si } x_c < x_{ref}, \text{ alors la pièce est classée à gauche ; sinon, elle est classée à droite} \quad (5)$$



Pièces endommagées détectées, pourcentages de dommages et catégories :

Phare_avant_gauche:	29.66% endommagé	- Catégorie: Dommages moyens
Aile_gauche:	100.00% endommagé	- Catégorie: Dommages très graves
Pare-brise:	72.73% endommagé	- Catégorie: Dommages graves
Aile_droite:	12.85% endommagé	- Catégorie: Dommages mineurs
Calandre:	33.93% endommagé	- Catégorie: Dommages moyens
Pare-brise:	84.94% endommagé	- Catégorie: Dommages très graves

Figure 7 : Distinction exploratoire gauche/droite sur une image réelle.

Cette règle reste simple, mais elle suffit à renforcer la cohérence topologique du pipeline lorsque le véhicule est photographié dans des conditions standards. Sa valeur principale n'est pas la sophistication mathématique ; elle réside dans l'utilité métier de l'information produite.

Cette composante doit toutefois être lue comme une étape exploratoire. Les essais montrent qu'elle fournit souvent une indication latérale utile, mais qu'elle reste sensible aux perspectives obliques, aux occlusions et aux erreurs de détection des pièces de référence. Une perspective future consistera donc à améliorer cette distinction gauche/droite à l'aide d'indices géométriques plus robustes.

CHAPITRE 4 – Expérimentations

4.1 Description du jeu de données

Le corpus final utilisé dans ce mémoire comprend 1 503 images après augmentation par retournement vertical (flip) dans Roboflow. Avant cette augmentation, le corpus initial comprenait plus de 700 images annotées de véhicules intacts et endommagés. Il a été constitué au Canada dans un contexte de collaboration avec Progi [21], puis enrichi par des images publiques provenant principalement de l'Université Stanford [22]. Environ 80 % des images correspondent à des cas réels collectés au Canada en collaboration avec Progi, tandis que les 20 % restants proviennent de sources publiques mobilisées pour accroître la diversité visuelle du corpus.

Après augmentation, le corpus final se compose de 805 images de véhicules intacts et de 698 images de véhicules endommagés. Cette répartition montre que les deux types de situations sont représentés dans le corpus, tout en conservant une part importante d'images endommagées, indispensable pour l'apprentissage et l'évaluation de la branche dédiée aux dommages. La distinction entre le corpus initial et le corpus augmenté est importante : les distributions d'annotations présentées ci-dessous correspondent au corpus avant application du flip dans Roboflow.

Tableau 1 – Composition du corpus et répartition entraînement/validation/test

Sous-ensemble	Total	Entraînement	Validation	Test
Corpus final après augmentation par flip	1 503 images	—	—	—
Véhicules intacts	805 images	—	—	—
Véhicules endommagés	698 images	—	—	—
Branche pièces	363 images	277	57	29
Branche dommages	698 images	516	135	47

La répartition entraînement/validation/test concerne les branches effectivement utilisées pour l'apprentissage des détecteurs. Le corpus final global indique la taille disponible après augmentation par retournement vertical, tandis que les branches peuvent utiliser des sous-ensembles distincts selon la nature des annotations disponibles.

Le jeu de données contient plusieurs catégories de véhicules, notamment des berlines, des véhicules utilitaires sport (SUV), des véhicules compacts et certaines camionnettes. Plusieurs marques automobiles sont représentées dans le corpus, parmi lesquelles Toyota, Hyundai, Honda, Ford, Nissan et Kia. Cette diversité permet de limiter un apprentissage trop spécifique à une seule structure de véhicule ou à une seule géométrie de carrosserie.

Les conditions de prise de vue présentent également une forte variabilité. Certaines images ont été capturées en extérieur, dans des stationnements ou sur la voie publique, tandis que d'autres proviennent d'environnements d'atelier ou d'inspection technique. Les angles de vue, les distances de prise de vue et les conditions d'éclairage varient fortement d'une image à l'autre. Plusieurs images contiennent également des reflets importants, des zones d'ombre, des arrière-plans complexes ou la présence simultanée de plusieurs véhicules. Cette diversité structurelle et visuelle est importante pour l'évaluation de DIRF-R, car elle rapproche les expérimentations des conditions réelles d'utilisation en expertise automobile.

4.1.1 Prétraitement et filtrage des images

Avant l'entraînement, un travail de nettoyage a été effectué sur le corpus. Les images trop floues, de résolution insuffisante, très fortement occultées ou comportant trop peu d'information exploitable ont été écartées. Les doublons et les images dont les dommages étaient difficilement discernables en raison de l'angle de vue ou de la qualité visuelle ont également été supprimés.

Le filtrage a aussi visé à exclure les cas où les pièces du véhicule n'étaient plus suffisamment identifiables, par exemple des véhicules trop sévèrement endommagés pour permettre une lecture structurelle fiable. L'objectif était de constituer un jeu de données compatible avec une tâche réaliste d'association pièce-dommage, et non de conserver des cas extrêmes peu interprétables.

4.1.2 Annotation manuelle avec Roboflow

Chaque image a été annotée manuellement via la plateforme Roboflow [23]. L'annotation a suivi une taxonomie structurée des pièces du véhicule et des zones de dommages associées. Des boîtes englobantes ont été attribuées à chaque objet annoté. Ce choix méthodologique est cohérent avec l'architecture proposée, puisque les deux détecteurs DETR manipulent précisément ce type d'annotation en entrée.

L'utilisation de Roboflow a présenté plusieurs avantages pratiques. D'une part, la plateforme facilite la gestion centralisée du corpus, la vérification des annotations et l'export

dans différents formats. D'autre part, elle permet d'appliquer des augmentations de données de manière contrôlée et traçable. Dans le présent travail, elle a surtout joué un rôle dans la phase de constitution et d'organisation du corpus, en rendant le processus d'annotation plus homogène.

Le jeu de données a été organisé en ensembles distincts d'entraînement, de validation et de test afin de limiter les risques de fuite d'information entre l'apprentissage et l'évaluation. Pour la branche pièces, les données ont été réparties en 277 images d'entraînement, 57 images de validation et 29 images de test. Pour la branche dommages, la répartition utilisée est de 516 images d'entraînement, 135 images de validation et 47 images de test. Cette séparation permet d'évaluer les deux détecteurs sur des images non vues tout en conservant suffisamment d'exemples pour l'apprentissage.

Le découpage en ensembles d'entraînement, de validation et de test a été réalisé avant l'application du retournement vertical. Le flip a ensuite été appliqué séparément à chacun des trois sous-ensembles, de sorte que chaque image transformée demeure dans le même sous-ensemble que son image originale. Roboflow a automatiquement recalculé les coordonnées des boîtes englobantes afin qu'elles restent alignées avec les pièces annotées. Le retournement étant vertical, il ne modifie pas la latéralité horizontale des pièces et ne nécessite donc aucun échange entre les étiquettes gauche et droite. Les métriques rapportées ont ainsi été calculées sur un ensemble de test comprenant également des images retournées verticalement.

L'application du retournement vertical aux ensembles de validation et de test constitue néanmoins une limite du protocole. Les images transformées sont directement dérivées des images originales et ne représentent donc pas des observations entièrement indépendantes. De plus, le retournement vertical produit des orientations de véhicules peu fréquentes dans les conditions réelles d'utilisation. Les métriques obtenues évaluent ainsi en partie la robustesse du détecteur à cette transformation, sans refléter exclusivement ses performances sur des photographies en orientation naturelle.

4.1.3 Taxonomie des classes

Le détecteur des pièces utilise une taxonomie de composants automobiles structurée. Le détecteur des dommages, en revanche, n'apprend qu'une seule classe, notée « Endommagée ». Ce choix est volontaire. Le mémoire ne cherche pas à classer les dommages par type visuel détaillé ; il cherche d'abord à les localiser, puis à en interpréter la gravité à partir de leur recouvrement avec les pièces.

Le tableau ci-dessous présente les classes utilisées. La classe « Endommagée » correspond à la branche dommages, tandis que toutes les autres classes appartiennent à la branche pièces. Dans le cas des roues, la logique métier retient également la quantité de roues touchées lorsqu'une ou plusieurs roues sont détectées comme endommagées.

Tableau 2 – Classes annotées pour les branches pièces et dommages

ID	Classe
0	Pare-chocs avant
1	Pare-chocs arrière
2	Calandre
3	Aile droite
4	Aile gauche
5	Capot
6	Pare-brise
7	Toit
8	Phare avant droit
9	Phare avant gauche
10	Phare arrière droit
11	Phare arrière gauche
12	Porte avant droite
13	Porte avant gauche
14	Porte arrière droite
15	Porte arrière gauche
16	Rétroviseur droit
17	Rétroviseur gauche
18	Roues
19	Panneau latéral droit
20	Panneau latéral gauche
21	Coffre
22	Vitre arrière

ID	Classe
23	Endommagée

La distribution des annotations n'est pas uniforme. La classe Roues est la plus représentée, avec 1 440 annotations, suivie des Rétroviseurs avec 795 annotations, des Portes avant avec 722 annotations et des Portes arrière avec 686 annotations. À l'inverse, les classes Calandre avec 300 annotations, Coffre avec 321 annotations et Vitre arrière avec 321 annotations sont les moins représentées. Pour la branche de détection des dommages, une seule classe, Damage, regroupe 1 459 annotations. Cette distribution traduit à la fois la fréquence d'apparition des différentes pièces dans les images du corpus et leur niveau de visibilité selon les angles de prise de vue.

Tableau 3 – Répartition des annotations avant augmentation par retournement vertical dans Roboflow

Classe ou groupe de classes	Nombre d'annotations avant retournement vertical	Lecture expérimentale
Roues	1 440	Classe la plus représentée, visible sur plusieurs angles de vue
Rétroviseurs	795	Forte présence latérale et bonne visibilité dans le corpus
Portes avant	722	Composant fréquent dans les vues latérales
Portes arrière	686	Composant fréquent, mais plus dépendant de l'angle
Vitre arrière	321	Classe moins visible, surtout présente dans les vues arrière
Coffre	321	Classe moins représentée, dépendante des vues arrière

Classe ou groupe de classes	Nombre d'annotations avant retournement vertical	Lecture expérimentale
Calandre	300	Classe la moins représentée parmi les classes détaillées
Damage	1 459	Classe unique de la branche dommages

Ces valeurs correspondent à la distribution communiquée avant l'application du retournement vertical. Elles mettent surtout en évidence le déséquilibre entre les pièces très fréquentes et les pièces plus rarement visibles.

Ce déséquilibre entre les classes constitue un défi supplémentaire pour l'entraînement du détecteur de pièces. Les catégories les moins représentées peuvent être plus difficiles à apprendre, ce qui peut entraîner une baisse de précision, une augmentation des faux négatifs ou une localisation moins stable pour ces classes. Cette observation permet d'expliquer en partie les performances plus limitées de la branche dédiée aux pièces automobiles et doit être prise en compte dans l'interprétation des résultats expérimentaux.

4.2 Outils et environnement de travail

Plusieurs outils ont été mobilisés dans la réalisation du pipeline DIRF-R. Python a constitué le langage de programmation principal pour le développement des scripts, la préparation des données, l'entraînement des modèles et l'évaluation [24]. Son écosystème scientifique, en particulier PyTorch, NumPy, OpenCV et les bibliothèques de manipulation de données, a permis de structurer l'ensemble du travail expérimental.

Google Drive a servi d'espace centralisé pour le stockage et l'organisation des fichiers, facilitant le partage des données, des annotations et des sorties intermédiaires [25]. Google Colaboratory a fourni l'environnement de calcul utilisé pour les entraînements et les expérimentations [26]. L'accès à des GPU (Graphics Processing Unit) a permis de réduire significativement les temps de calcul, ce qui est particulièrement utile dans le cas de détecteurs à base de Transformers, souvent plus coûteux à entraîner que des modèles plus légers.

Roboflow a enfin joué un rôle clé dans la gestion du corpus annoté. Au-delà de l'annotation, la plateforme a facilité l'export des données, la vérification des classes et l'application de l'augmentation principale retenue dans ce travail, à savoir le retournement vertical (flip). Cette augmentation a permis d'enrichir le corpus final tout en préservant la cohérence géométrique des annotations utilisées par le pipeline.

4.3 Protocole expérimental

Le protocole expérimental a été conçu de façon à rester cohérent avec la structure du pipeline. Les deux détecteurs DETR ont été entraînés séparément sur leurs jeux d'annotations respectifs. Cette séparation garantit que le détecteur des pièces apprend la structure du véhicule, tandis que le détecteur des dommages apprend uniquement la localisation des zones altérées. Une fois les deux branches entraînées, leurs sorties sont combinées uniquement en phase d'inférence au sein de DIRF-R.

Le processus expérimental comporte donc deux niveaux. Le premier niveau est purement quantitatif : il consiste à évaluer la qualité de chaque détecteur à l'aide de métriques standard de détection. Le second niveau est fonctionnel : il consiste à examiner comment ces sorties peuvent être exploitées par le pipeline complet pour produire une lecture structurée des dommages. Le mémoire met l'accent sur cette articulation, car l'objectif n'est pas seulement d'obtenir des scores, mais de montrer que les scores alimentent un raisonnement utile.

L'augmentation de données retenue dans la version finale du corpus repose principalement sur le retournement vertical (flip) appliqué dans Roboflow. Cette transformation a été appliquée séparément dans chacun des sous-ensembles afin de préserver le découpage initial et la cohérence géométrique des annotations.

Tableau 4 – Hyperparamètres et environnement d'entraînement des modèles DETR

Paramètre	Valeur retenue	Commentaire
Architecture de base	Facebook/detr-resnet-50	Même base pour les deux branches
Backbone	ResNet-50	Extraction des caractéristiques visuelles
Nombre de branches	2	Une branche pièces et une branche dommages
Nombre d'époques	100	Entraînement séparé des deux détecteurs
Taille des images	640 × 640	Redimensionnement appliqué lors de la préparation

Paramètre	Valeur retenue	Commentaire
Batch size	4	Choix adapté aux contraintes GPU
Taux d'apprentissage	1e-4	Paramètre principal de l'optimisation
Taux d'apprentissage du backbone	1e-5	Apprentissage plus prudent des couches préentraînées
Optimiseur	AdamW	Optimiseur utilisé pour le fine-tuning
Weight decay	1e-4	Régularisation des poids
Augmentation principale	Flip vertical Roboflow	Appliqué séparément dans chaque sous-ensemble
Seuil de confiance en inférence	0,7	Valeur utilisée dans le code de post-traitement DETR
Association pièce-dommage	Intersection positive	Une pièce est retenue si l'aire d'intersection avec un dommage est > 0
Environnement	Google Colab / GPU Tesla T4	Environnement de calcul utilisé pour l'entraînement et l'inférence

4.4 Métriques d'évaluation

Les performances des deux détecteurs sont évaluées à l'aide de trois indicateurs principaux : le $\text{mAP}@0,5$ (mean Average Precision à un seuil d'IoU de 0,5), le $\text{mAP}@50:95$ (mean Average Precision calculé sur plusieurs seuils d'IoU de 0,50 à 0,95) et le score F1 (score harmonique entre la précision et le rappel). En notation standard, le mAP correspond à la moyenne des Average Precisions (AP) calculées sur l'ensemble des classes [28]. Le $\text{mAP}@0,5$ mesure la précision moyenne de détection à un seuil d'IoU de 0,5. Le $\text{mAP}@50:95$ est plus exigeant, car il agrège la précision sur plusieurs seuils allant de 0,50 à 0,95, selon une logique d'évaluation popularisée par le protocole COCO (Common Objects in Context) [27]. Il fournit

donc une lecture plus stricte de la qualité de localisation. Le score F1 résume pour sa part l'équilibre entre précision et rappel [28].

Les définitions usuelles sont rappelées ci-dessous pour garder le document lisible :

$$\text{Précision} = \text{TP (True Positive)} / (\text{TP} + \text{FP (False Positive)}) \quad (6)$$

$$\text{Rappel} = \text{TP} / (\text{TP} + \text{FN (False Negative)}) \quad (7)$$

$$\text{F1} = 2 \times (\text{Précision} \times \text{Rappel}) / (\text{Précision} + \text{Rappel}) \quad (8)$$

Ces métriques n'évaluent pas directement la qualité du raisonnement flou. Elles évaluent la qualité de la base visuelle qui alimente le pipeline. L'apport du module flou se situe en aval, au niveau de l'interprétation des Damage Ratios. C'est pourquoi les résultats de détection doivent être lus comme une condition nécessaire à la pertinence de DIRF-R, mais non comme son unique finalité.

Le protocole d'évaluation effectivement exécuté et conservé pour ces expérimentations a fourni le mAP@0,5, le mAP@50:95 et le score F1 pour chacune des deux branches DETR. Il n'a pas généré automatiquement de matrices de confusion, de courbes précision-rappel ni de valeurs d'AP détaillées par classe. Ces analyses sont techniquement calculables à partir des prédictions, mais elles nécessiteraient une nouvelle exécution sur les ensembles de test avec un module d'évaluation complémentaire et la conservation des sorties par image. Les résultats présentés dans ce mémoire sont donc volontairement limités aux métriques effectivement obtenues et archivées au cours des expérimentations.

4.5 Lecture expérimentale spécifique à DIRF-R

Une erreur classique, dans les systèmes hybrides, consiste à vouloir résumer toute l'architecture par une seule métrique. Ce n'est pas l'option retenue ici. DIRF-R est une chaîne où les détecteurs fournissent la base des étapes suivantes. Ainsi, une lecture correcte des résultats doit distinguer les performances de détection des performances d'interprétation. Un détecteur plus précis produit une meilleure base pour l'association pièce-dommage, mais l'intérêt du pipeline ne s'arrête pas à ce constat.

Le rôle de la métrique est ici de quantifier la qualité des sorties des deux branches DETR. Le rôle du reste du pipeline est de transformer ces sorties en connaissance interprétable : quelles pièces sont touchées, dans quelle proportion et avec quel degré de gravité. C'est précisément cette hiérarchie entre mesure, association et interprétation qui donne au mémoire sa cohérence.

4.5.1 Préparation des annotations pour l’entraînement

Les annotations produites dans Roboflow ont été exportées dans un format compatible avec le pipeline d’apprentissage de détection d’objets. Dans un cadre DETR, cela suppose de disposer, pour chaque image, des coordonnées des boîtes, des identifiants de classes et des métadonnées minimales nécessaires au chargement du jeu de données. Une partie du travail expérimental a donc consisté à vérifier la cohérence des noms de classes, l’absence d’étiquettes manquantes et la validité des coordonnées.

Cette étape de préparation est souvent sous-estimée dans les travaux appliqués, alors qu’elle conditionne directement la qualité des résultats. Une boîte mal définie, une classe mal étiquetée ou une incohérence entre les versions du corpus peuvent dégrader l’apprentissage de manière silencieuse. Le choix d’un pipeline propre, avec une séparation claire entre annotations pièces et annotations dommages, a donc été déterminant pour la suite des expérimentations.

4.5.2 Schéma d’entraînement des deux branches

Les deux branches DETR ont été entraînées séparément, chacune sur sa propre cible. Cette décision évite que la branche pièces n’apprenne des signaux propres aux dommages et que la branche dommages ne soit pénalisée par la forte diversité des composants structurels. Dans la pratique, l’entraînement séparé rend également le diagnostic expérimental plus simple : si une erreur apparaît, il est plus facile d’identifier si elle provient de la localisation des pièces, de la localisation des dommages ou de l’étape d’association.

Le suivi expérimental a privilégié une logique de stabilisation plutôt qu’une recherche exhaustive d’hyperparamètres. Le but de ce mémoire n’est pas de présenter une optimisation extrême des détecteurs DETR, mais de montrer qu’ils peuvent alimenter correctement le pipeline DIRF-R. Cette précision méthodologique est importante. Elle rappelle que le cœur de la contribution n’est pas un nouveau détecteur, mais une manière cohérente de transformer des détections en interprétations d’expertise.

Pour renforcer la reproductibilité, les entraînements ont été menés avec les paramètres expérimentaux suivants : 100 époques, batch size de 4, taux d’apprentissage principal de $1e-4$, taux d’apprentissage du backbone de $1e-5$, optimiseur AdamW et weight decay de $1e-4$. Les images sont ramenées à une taille compatible avec le pipeline de détection, notamment 640×640 dans la préparation Roboflow. Les expérimentations ont été réalisées dans Google Colaboratory, avec accès à un GPU de type Tesla T4 lorsque disponible.

Les deux branches ont été entraînées à partir du modèle pré entraîné Facebook/DETR-resnet-50, avec un backbone ResNet-50. La branche pièces utilise 23 classes sémantiques. La branche dommages utilise une seule classe sémantique, « Endommagée ». La classe no_object est une classe interne au mécanisme de prédiction de DETR et ne constitue pas une classe annotée du corpus. Les poids entraînés sont ensuite chargés séparément pour chaque branche avant l'inférence. Les prédictions sont post-traitées avec le processeur DETR et un seuil de confiance fixé à 0,7 dans le script d'expérimentation. L'exécution se fait sur GPU lorsque CUDA est disponible, sinon sur CPU.

4.5.3 Du résultat brut au pipeline complet

Une fois les deux détecteurs entraînés, les sorties brutes sont exploitées en phase d'inférence. Pour chaque image test, la branche pièces et la branche dommages produisent leurs boîtes et leurs scores. Ces sorties sont d'abord filtrées pour ne conserver que les prédictions les plus pertinentes. Elles sont ensuite transmises au module d'association géométrique. Le calcul de l'IoU permet d'identifier quelles pièces et quels dommages se recouvrent réellement. Le Damage Ratio est alors calculé pour chaque pièce concernée, puis transmis au module flou.

Dans le script d'inférence, les étapes sont exécutées de manière séquentielle. Les boîtes des pièces sont d'abord affichées en vert pour vérification visuelle, puis les boîtes des dommages sont affichées en rouge. Chaque dommage détecté est comparé à chaque pièce détectée. Lorsqu'une intersection positive existe, la pièce est ajoutée à la liste des pièces endommagées avec son pourcentage de dommage. Les résultats locaux sont ensuite imprimés sous forme de triplets : pièce endommagée, pourcentage de dommage et catégorie. Enfin, une moyenne des pourcentages de dommage est calculée pour produire une classification globale par logique floue.

Cette procédure permet de passer d'un résultat brut de détection à une représentation organisée. Le lecteur peut alors suivre, de manière séquentielle, le chemin d'une image dans le pipeline complet. Cette traçabilité est l'un des avantages majeurs de DIRF-R dans une perspective de validation scientifique et d'usage applicatif.

CHAPITRE 5 – Résultats

5.1 Présentation synthétique des résultats

Les résultats quantitatifs sont rapportés sous forme de synthèse afin d'évaluer les deux détecteurs qui constituent la base du pipeline DIRF-R : le détecteur des pièces et le détecteur des dommages. Une comparaison indicative avec YOLOv8 et FuzzyYOLOv8, issue du mémoire de Kogbeti-Wongo [17], est également ajoutée pour situer les performances obtenues par DETR par rapport à une méthode de référence utilisée sur une problématique proche. Cette comparaison doit toutefois être interprétée avec prudence, car les protocoles et les objectifs exacts ne sont pas strictement identiques.

Le détecteur des dommages obtient les meilleures valeurs sur l'ensemble des métriques. Cette observation est cohérente avec la structure du problème : la branche dommages ne travaille que sur une seule classe, alors que la branche pièces doit distinguer plusieurs composants visuellement proches. La diversité des pièces, la présence d'occlusions partielles et la variabilité des angles rendent la tâche intrinsèquement plus difficile.

5.2 Interprétation détaillée des performances

Les scores du détecteur des pièces montrent que la structuration du véhicule reste une tâche exigeante. Un $mAP@0,5$ de 0,558 indique que le modèle parvient à localiser correctement une part importante des composants, mais que des marges de progression subsistent, en particulier lorsque l'on considère des seuils de localisation plus stricts. Le $mAP@50:95$ de 0,289 confirme cette difficulté. Il révèle que certaines boîtes détectées restent approximatives lorsque le critère de localisation devient plus sévère. Le score F1 de 0,312 montre enfin que le compromis entre précision et rappel reste perfectible.

À l'inverse, la branche dommages atteint des scores plus élevés : 0,759 en $mAP@0,5$, 0,452 en $mAP@50:95$ et 0,505 en F1. Ces résultats montrent que le détecteur localise relativement bien les zones de dommages dans le cadre du corpus étudié. Il faut cependant éviter une lecture naïve. Un bon score de détection des dommages ne signifie pas automatiquement que toutes les évaluations locales seront correctes. La qualité finale du pipeline dépend aussi de l'étape d'association. Un dommage peut être bien détecté, mais mal relié à une pièce si cette dernière est mal localisée.

5.3 Ce que les métriques disent, et ce qu'elles ne disent pas

Les métriques rapportées dans ce chapitre évaluent exclusivement les sorties des détecteurs DETR. Elles ne mesurent pas directement le comportement du module flou. Ce point doit être explicite pour éviter tout malentendu. Le rôle du module flou n'est pas d'augmenter artificiellement le mAP ou le F1 des détecteurs ; il intervient plus tard, une fois les Damage Ratios calculés. Son apport est de transformer une information continue en une décision plus progressive.

Pour compléter cette lecture, le pipeline complet est évalué fonctionnellement à partir des sorties présentées dans le mémoire. Les figures de résultats et les annexes montrent le passage des boîtes de détection vers les pièces réellement associées à un dommage, puis vers le Damage Ratio et la catégorie de gravité. Cette évaluation ne remplace pas un benchmark end-to-end annoté manuellement, mais elle démontre que les étapes Détection, Intersection, Ratio et Fuzzy Reasoning sont bien connectées dans le programme et produisent une sortie interprétable pièce par pièce.

Autrement dit, les métriques disent si la base visuelle du pipeline est suffisamment fiable pour alimenter les étapes suivantes. Elles ne disent pas, à elles seules, si le système produit une expertise de qualité. Pour répondre à cette seconde question, il faut regarder la cohérence du pipeline complet : les pièces détectées, les dommages détectés, l'IoU calculée, le ratio obtenu, puis l'interprétation floue. C'est la combinaison de ces éléments qui fait la valeur de DIRF-R.

5.4 Évaluation de bout en bout du pipeline DIRF-R

L'évaluation end-to-end vise à vérifier le comportement du pipeline complet, depuis les sorties des deux détecteurs jusqu'à la décision finale de gravité. Contrairement aux métriques de détection, qui évaluent séparément les boîtes produites par chaque modèle, cette lecture suit la chaîne complète : détection des pièces, détection des dommages, association géométrique, calcul du Damage Ratio, puis interprétation par logique floue. Les exemples ci-dessous sont issus des sorties présentées dans le mémoire et les annexes ; ils servent à montrer la cohérence fonctionnelle du pipeline sur des cas représentatifs.

Tableau 5 – Évaluation qualitative de bout en bout du pipeline DIRF-R

Cas	Pièces associées au dommage	Damage Ratios observés	Gravité obtenue	Lecture end-to-end
Image A	Capot, pare-chocs avant, phare avant droit, roues	5,66 % ; 14,81 % ; 100,00 % ; 9,29 %	Mineur à très sévère	Association cohérente entre dommage frontal et pièces touchées
Image B	Capot, capot, phare avant gauche, phare avant droit, pare-chocs avant	72,73 % ; 100,00 % ; 72,58 % ; 100,00 % ; 14,08 %	Mineur à très sévère	Le pipeline distingue plusieurs niveaux locaux sur une même scène
Image C	Capot, pare-chocs avant, phare avant gauche, phare avant droit, roues	60,36 % ; 30,20 % ; 95,93 % ; 88,75 % ; 28,89 %	Modéré à très sévère	La sortie combine localisation, ratio et interprétation graduelle
Image D	Capot, pare-chocs avant, phare avant gauche, roues	41,72 % ; 18,86 % ; 95,43 % ; 4,87 %	Mineur à très sévère	Les dommages importants et faibles coexistent dans une même lecture structurée

Cette évaluation reste qualitative, car elle ne repose pas sur une annotation humaine complète des associations pièce-dommage pour toutes les images de test. Elle permet néanmoins de vérifier que les composantes du pipeline sont effectivement connectées et que la sortie finale est interprétable au niveau des pièces.

Sur le plan fonctionnel, DIRF-R permet de passer d'une image brute à une lecture structurée de l'état du véhicule. Lorsqu'une pièce est détectée et qu'une zone de dommage lui est associée par un recouvrement positif, le Damage Ratio fournit immédiatement une mesure de l'atteinte locale. Le module flou transforme alors cette quantité en un niveau de gravité qui peut être compris par un lecteur non spécialiste. Cette chaîne est particulièrement utile dans des scénarios où l'on souhaite expliquer pourquoi les dommages d'une pièce sont considérés comme mineurs, modérés, sévères ou très sévères.

L'étude fonctionnelle permet également d'identifier les points de propagation des erreurs. Si une pièce n'est pas détectée, le système ne peut pas lui associer de dommage. Si une boîte de pièce est approximative, le Damage Ratio peut être sous-estimé ou surestimé. Si un dommage est correctement détecté mais recouvre plusieurs pièces voisines, l'association peut

devenir ambiguë. Ainsi, DIRF-R est explicable parce que chaque décision peut être suivie, mais cette traçabilité montre aussi que la qualité finale dépend de l'ensemble de la chaîne.

L'intérêt pratique du pipeline est donc moins de produire un unique score global que d'offrir une interprétation traçable. C'est ce qui distingue une architecture d'aide à l'expertise d'un simple système de détection. Dans un contexte d'assurance, par exemple, on ne cherche pas seulement à savoir si un dommage existe ; on cherche à documenter où il se trouve, quelle pièce il affecte et de quelle ampleur il est.

5.5 Analyse qualitative des meilleurs cas et des cas difficiles

Au-delà des résultats numériques, l'analyse visuelle des sorties du pipeline permet de mieux comprendre les situations dans lesquelles DIRF-R produit une interprétation cohérente, ainsi que les cas où certaines limites apparaissent. La Figure 8 regroupe quatre exemples représentatifs issus des sorties du système. Ces exemples ne remplacent pas les métriques quantitatives, mais ils permettent d'explicitier le comportement du pipeline face à des configurations visuelles concrètes.

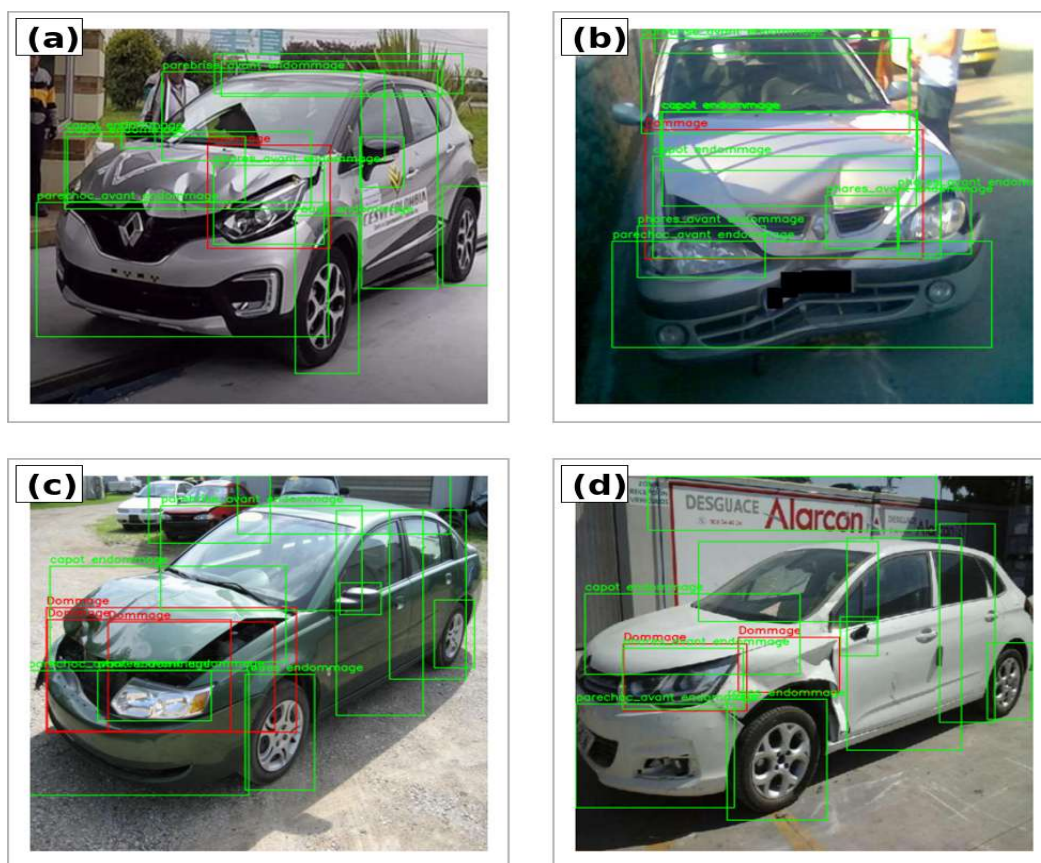


Figure 8 : Exemples de résultats obtenus avec DIRF-R.

Les cas (a) et (b) illustrent des situations relativement favorables pour le pipeline. Le véhicule principal est clairement visible, plusieurs pièces structurelles sont détectées, et les zones de dommages apparaissent dans des régions cohérentes avec les composants concernés. Dans ce type de configuration, l'association géométrique entre pièces et dommages peut être exploitée pour alimenter le calcul du Damage Ratio, puis l'interprétation floue de la gravité.

Les cas (c) et (d) mettent davantage en évidence les difficultés rencontrées en conditions réelles. On observe une densité importante de boîtes englobantes, des superpositions entre étiquettes, des pièces partiellement visibles et des perspectives parfois obliques. Ces éléments peuvent rendre l'association pièce-dommage plus délicate et influencer la précision du ratio calculé. Ils montrent que la qualité de la sortie finale dépend non seulement de la détection des dommages, mais aussi de la localisation correcte des pièces et de la stabilité géométrique des prédictions.

Cette analyse qualitative confirme donc l'intérêt d'un pipeline explicable, dans lequel chaque étape peut être examinée séparément. Les exemples favorables montrent la capacité du système à produire une lecture structurée de l'état du véhicule, tandis que les cas plus difficiles soulignent l'importance des mécanismes de filtrage, de consolidation géométrique et, à terme, d'approches plus fines telles que la segmentation ou un modèle d'estimation spécialisé pour améliorer l'évaluation locale des dommages.

5.6 Analyse conceptuelle de l'apport des composantes de DIRF-R

Cette analyse conceptuelle ne repose pas sur le réentraînement de plusieurs variantes complètes du modèle. Elle examine l'effet fonctionnel attendu de la suppression de chaque composante du pipeline et met ainsi en évidence pourquoi DIRF-R ne se limite pas à un détecteur, mais correspond à une chaîne explicable dans laquelle chaque étape transforme l'information précédente.

Tableau 6 – Analyse conceptuelle de l'apport des composantes du pipeline DIRF-R

Composante retirée	Effet attendu	Limite observée	Contribution conservée dans DIRF-R
Deux détecteurs spécialisés	Un modèle unique devrait apprendre simultanément pièces et dommages	Risque de confusion entre composants stables et altérations irrégulières	Séparation claire des tâches visuelles
Isolation du véhicule principal	Toutes les détections de la scène seraient conservées	Risque d'associer des pièces de véhicules secondaires au dommage principal	Réduction des distracteurs et meilleure cohérence de scène

Composante retirée	Effet attendu	Limite observée	Contribution conservée dans DIRF-R
Association par intersection	Les pièces et les dommages resteraient deux listes indépendantes	Impossible d'indiquer quelle pièce est réellement touchée	Lien spatial explicite entre dommage et composant
Damage Ratio	La sortie resterait une simple présence/absence de dommage	Absence de quantification locale de l'atteinte	Mesure interprétable de la proportion visible touchée
Logique floue	La gravité serait obtenue par seuils rigides	Décisions abruptes autour des frontières de classes	Interprétation progressive et plus proche du raisonnement expert

Cette analyse conceptuelle met en évidence le rôle attendu de chaque composante du pipeline.

5.7 Synthèse

Les résultats confirment que les deux détecteurs fournissent une base exploitable pour le pipeline proposé. Ils montrent aussi que la branche pièces reste le maillon le plus difficile, ce qui est cohérent avec la nature multi classes et structurelle du problème. Malgré des performances inférieures à celles de la branche dommages, la branche pièces atteint un $mAP@0,5$ de 0,558, un $mAP@50:95$ de 0,289 et un score F1 de 0,312. Ces résultats restent suffisants pour soutenir les étapes subséquentes du pipeline DIRF-R, dans lesquelles la détection constitue une base d'analyse destinée à être enrichie par l'association géométrique, la quantification des dommages et le raisonnement flou.

En ce sens, les résultats soutiennent l'hypothèse centrale du mémoire : une architecture comme DIRF-R, qui sépare détection, association, quantification et interprétation, constitue une manière cohérente et compréhensible de traiter l'évaluation automatisée des dommages automobiles.

5.8 Discussion qualitative et intérêt applicatif

Les résultats obtenus mettent en évidence une différence notable entre les performances de la branche de détection des dommages et celles de la branche de détection des pièces. La branche dommages atteint un $mAP@0,5$ de 0,759, ce qui indique que le détecteur parvient à localiser correctement une proportion importante des zones endommagées lorsque le seuil d'acceptation est relativement permissif. Ce résultat montre que les dommages visibles

possèdent généralement des caractéristiques suffisamment discriminantes pour être identifiés par le modèle.

Lorsque l'évaluation est réalisée sur une plage plus stricte de seuils d'IoU, le $mAP@50:95$ diminue à 0,452. Cette diminution est attendue, car elle traduit la difficulté à positionner avec précision les boîtes englobantes autour de dommages dont les contours sont souvent irréguliers, diffus ou partiellement visibles. Malgré cette baisse, le score obtenu demeure cohérent avec la complexité du problème étudié.

Le score F1 de la branche dommages atteint 0,505. Cette valeur reflète un compromis raisonnable entre précision et rappel. Elle indique que le modèle parvient à détecter une part importante des dommages réels tout en limitant les fausses détections. Dans le contexte du pipeline DIRF-R, cette performance est particulièrement importante puisque les dommages détectés constituent l'entrée principale des étapes d'association géométrique et d'interprétation floue.

La branche dédiée aux pièces obtient des résultats plus modestes, avec un $mAP@0,5$ de 0,558, un $mAP@50:95$ de 0,289 et un score F1 de 0,312. Plusieurs facteurs peuvent expliquer cette différence. Tout d'abord, le nombre de classes est significativement plus élevé pour les pièces que pour les dommages. Ensuite, certaines pièces possèdent des formes similaires, apparaissent sous des angles variés ou peuvent être partiellement occultées. Enfin, les dimensions relatives des composants varient fortement d'une image à l'autre, ce qui augmente la difficulté de détection.

Malgré ces performances plus faibles, les résultats restent suffisants pour soutenir l'objectif principal de DIRF-R. En effet, le pipeline ne repose pas uniquement sur la détection brute des pièces, mais également sur l'étape d'association géométrique par IoU, sur le calcul du Damage Ratio et sur l'interprétation progressive assurée par la logique floue. Ainsi, même lorsque certaines détections demeurent imparfaites, le système conserve une capacité d'analyse structurée et explicable.

L'analyse conceptuelle détaillée à la section 5.6 montre que chaque composante de DIRF-R joue un rôle spécifique dans la transformation d'une détection brute en information exploitable. Cette lecture renforce l'idée que la contribution principale du mémoire ne réside pas uniquement dans la performance de détection, mais dans la structuration explicable de la décision.

La comparaison indicative avec YOLOv8 et FuzzyYOLOv8 montre que les approches de type YOLO obtiennent de meilleurs scores sur la branche pièces dans le mémoire de

référence [17]. Cette différence est cohérente avec la littérature : YOLO est souvent plus efficace sur des jeux de données de taille limitée et sur des tâches de détection rapide. En revanche, le choix de DETR reste défendable dans DIRF-R, car l'objectif principal est de construire une chaîne explicable et structurée, où les sorties de détection sont reliées à une association géométrique et à un raisonnement flou. Le gain attendu n'est donc pas uniquement un score de détection, mais une meilleure traçabilité de la décision.

D'un point de vue applicatif, ces résultats montrent que DIRF-R peut constituer un outil d'aide à l'expertise automobile. L'association explicite entre les dommages détectés et les pièces concernées fournit une information plus riche qu'une simple localisation de défauts. De plus, l'utilisation du Damage Ratio et de la logique floue permet de produire des évaluations graduelles plus proches du raisonnement d'un expert humain. Cette propriété est particulièrement intéressante pour les compagnies d'assurance et les ateliers de réparation qui recherchent des systèmes à la fois performants et interprétables.

Tableau 7 – Résultats quantitatifs et comparaison avec YOLOv8/FuzzyYOLOv8

Modèle	mAP@0,5	mAP@50:95	F1
DETR – Pièces (DIRF-R)	0,558	0,289	0,312
DETR – Dommages (DIRF-R)	0,759	0,452	0,505
YOLOv8 – Pièces intactes [17]	0,899	0,440	0,87
FuzzyYOLOv8 – Pièces intactes [17]	0,920	0,580	0,88
YOLOv8 – Dommages [17]	0,637	0,315	0,64
FuzzyYOLOv8 – Dommages [17]	0,680	0,340	0,69

CHAPITRE 6 – Menaces à la validité de l'étude

6.1 Validité interne et propagation des erreurs

La validité interne de l'étude dépend d'abord de la qualité du corpus et des annotations. Même si le jeu de données final comprend 1 503 images après augmentation par retournement vertical, il reste modeste au regard de la diversité réelle des véhicules, des dommages et des contextes visuels. Une erreur d'annotation sur une pièce ou sur une zone de dommage peut affecter à la fois l'apprentissage des détecteurs et le calcul des Damage Ratios. De plus, comme les deux détecteurs sont entraînés séparément, une faiblesse dans l'un d'eux peut se propager à l'ensemble du pipeline. Si une pièce est absente, mal classée ou mal localisée, l'association par intersection, le Damage Ratio et la catégorie finale peuvent être affectés.

La propagation des erreurs suit la structure même du pipeline. Une pièce non détectée ne peut recevoir aucun Damage Ratio, même si un dommage est correctement localisé dans la région correspondante. Une pièce mal localisée peut au contraire produire un ratio sous-estimé ou surestimé, car l'aire de référence utilisée au dénominateur devient imprécise. De la même manière, un faux positif dans la branche dommages peut conduire à attribuer une gravité injustifiée à une pièce saine, tandis qu'un faux négatif peut empêcher l'identification d'un dommage réel.

L'étape d'association par intersection constitue donc un point de transmission critique : elle ne crée pas l'information visuelle, mais elle amplifie les conséquences des erreurs produites en amont. Cette propriété explique pourquoi la branche pièces, dont les performances sont plus faibles que celles de la branche dommages, représente le maillon le plus sensible du système. Pour limiter cette vulnérabilité, les travaux futurs devraient intégrer une validation explicite des associations pièce-dommage, par exemple à l'aide d'annotations humaines end-to-end ou de masques de segmentation.

Une autre limite interne concerne le choix des fonctions d'appartenance floues. Les fonctions triangulaires retenues dans le code sont simples et interprétables, mais elles traduisent un compromis empirique entre souplesse et lisibilité. Un autre réglage pourrait déplacer certaines interprétations de gravité, notamment autour des zones de transition entre dommage mineur, modéré, sévère et très sévère. Cette remarque ne remet pas en cause l'intérêt du module flou, mais elle invite à considérer ses paramètres comme des choix à justifier et à tester plus systématiquement dans des travaux futurs.

6.2 Validité externe et changement de domaine

La validité externe renvoie à la capacité de généralisation du système. Le corpus utilisé a été constitué dans un contexte précis, au Canada, en collaboration avec Progi et incluant des

données issues de l'Université Stanford. Même si les images couvrent plusieurs angles et plusieurs types de scènes, un domaine shift peut exister entre les sources, les conditions de prise de vue et les types de véhicules. Des véhicules de marchés différents, des prises de vue dans des conditions météorologiques extrêmes, des reflets importants ou des scénarios avec de très fortes occlusions pourraient dégrader les performances.

Le domaine shift entre les images issues de Progi et celles issues de Stanford constitue une limite importante. Les deux sources peuvent différer par la résolution, le cadrage, les arrière-plans, les conditions d'éclairage, les types de véhicules, la qualité de compression et la manière dont les dommages apparaissent. Un modèle entraîné sur un mélange de ces sources peut apprendre des régularités propres à une source plutôt que des caractéristiques générales des dommages et des pièces automobiles.

Pour évaluer plus précisément cette généralisation, une expérience complémentaire pourrait consister à entraîner le modèle sur une source et à le tester sur l'autre, ou à construire un découpage de test strictement séparé par origine des images. Une telle validation permettrait de mesurer la robustesse de DIRF-R face à des images provenant d'un autre contexte d'acquisition, ce qui est essentiel pour une application réelle en assurance automobile.

Il faut également rappeler que le pipeline travaille à partir d'images bidimensionnelles. Des dommages internes, des déformations peu visibles en façade ou des atteintes situées sur des zones non photographiées ne peuvent pas être correctement évalués par le système. L'architecture reste donc pertinente pour une expertise visuelle à partir d'images, mais non pour une inspection exhaustive de tout le véhicule.

6.3 Validité de construit

La validité de construit concerne l'adéquation entre les concepts étudiés et les indicateurs utilisés. Dans ce mémoire, les métriques de détection évaluent bien la qualité de localisation des pièces et des dommages, mais elles ne capturent pas directement la pertinence du raisonnement flou. Le Damage Ratio, de son côté, constitue une approximation géométrique de la surface atteinte, calculée à partir des boîtes englobantes visibles. Une boîte englobante peut contenir une quantité importante de pixels non endommagés, surtout dans le cas de dommages fins, diagonaux ou irréguliers. Le ratio obtenu ne correspond donc pas toujours à la surface réelle du dommage.

En ce sens, le mémoire assume une définition opératoire de la gravité : il s'agit de la gravité visible, telle qu'elle peut être inférée à partir d'un recouvrement image. Cette définition est cohérente avec le cadre du travail, mais elle doit être explicitement reconnue comme une simplification du concept métier de dommage.

6.4 Limites pratiques

Sur le plan pratique, l'approche reste tributaire de la qualité des images d'entrée. Des images très compressées, trop sombres, prises de trop loin ou comportant une partie seulement du véhicule peuvent réduire l'utilité du pipeline. La méthode suppose également que la surface visible d'une pièce est un proxy raisonnable de son atteinte, ce qui n'est pas toujours vrai lorsque le dommage se prolonge hors champ ou se manifeste surtout par une déformation interne. À terme, une segmentation par masques pourrait améliorer cette limite en estimant plus directement la surface endommagée.

Ces limites ne rendent pas le système inutilisable ; elles rappellent simplement qu'il s'agit d'un outil d'aide à l'évaluation fondé sur l'analyse visuelle de photographies. Le jugement expert garde donc sa place, surtout dans les dossiers complexes.

CHAPITRE 7 – Conclusion

7.1 Bilan général

Ce mémoire a proposé DIRF-R, un pipeline d'évaluation automatisée des dommages automobiles articulé autour de quatre étapes clairement séparées : Détection, Intersection, Ratio et Fuzzy Reasoning. L'architecture repose sur deux détecteurs DETR indépendants, l'un pour les pièces du véhicule et l'autre pour les zones de dommages. Leurs sorties sont mises en relation par une mesure géométrique réelle d'intersection, puis transformées en Damage Ratios interprétés à l'aide d'un module de logique floue. Les performances des deux détecteurs ont été évaluées quantitativement, tandis que le fonctionnement du pipeline complet a été illustré qualitativement sur plusieurs cas représentatifs. Une validation quantitative de bout en bout reste à réaliser.

La principale force de cette approche tient à sa lisibilité méthodologique. Le système ne cherche pas à cacher la décision dans une seule boîte noire. Il explicite au contraire les différentes transformations qui mènent d'une image à une évaluation. Cette structuration rend le pipeline plus compréhensible pour un lecteur généraliste, plus défendable dans un cadre académique, et plus pertinent dans un contexte applicatif où l'explication du résultat est importante.

7.2 Apports scientifiques et méthodologiques

Sur le plan scientifique, le mémoire montre qu'une approche fondée sur DETR et sur la logique floue est pertinente pour traiter l'évaluation des dommages automobiles. Le recours à deux détecteurs spécialisés évite de mélanger la structure du véhicule et les zones d'altération. L'association par IoU fournit un mécanisme simple et objectivable pour relier les deux branches. Le Damage Ratio introduit une variable intermédiaire facile à interpréter. Enfin, le module flou permet d'éviter une décision trop rigide et rapproche le système d'un raisonnement d'expertise.

Sur le plan méthodologique, DIRF-R apporte aussi une réponse à un problème fréquent dans les mémoires techniques : la confusion entre performance algorithmique et utilité du système. Ici, les métriques de détection sont clairement distinguées du raisonnement final. Cette séparation renforce la cohérence de l'ensemble et limite les interprétations abusives.

7.3 Limites et perspectives

Plusieurs prolongements sont envisageables. Le premier concerne l'augmentation du corpus et la diversification des véhicules et des contextes de prise de vue. Un corpus plus vaste permettrait de renforcer la robustesse des détecteurs, en particulier pour la branche pièces. Le deuxième prolongement concerne l'amélioration de la représentation géométrique. Le calcul du

Damage Ratio à partir de boîtes englobantes constitue une approche simple, interprétable et cohérente avec les objectifs de DIRF-R. Toutefois, une piste d'amélioration future consisterait à remplacer cette estimation déterministe par un troisième modèle spécialisé chargé d'apprendre directement la relation entre les sorties du détecteur de pièces et du détecteur de dommages. Une telle approche permettrait d'exploiter automatiquement des relations spatiales plus complexes que le simple recouvrement géométrique, tout en offrant une meilleure capacité d'adaptation à mesure que de nouvelles données d'apprentissage deviennent disponibles. L'estimation du niveau de dommage ne reposerait alors plus uniquement sur des règles définies manuellement, mais sur un mécanisme d'apprentissage susceptible d'être continuellement amélioré. Le troisième prolongement porte sur la validation métier. Une confrontation systématique des sorties du pipeline avec des évaluations humaines expertes permettrait de mieux calibrer les fonctions d'appartenance floues et les règles d'agrégation globale.

Enfin, il serait intéressant d'explorer des variantes DETR plus avancées, telles que Deformable DETR, Conditional DETR ou d'autres modèles spécialisés pour les petits objets, afin d'améliorer la qualité de la branche pièces sans rompre la philosophie générale du pipeline. Ces perspectives confirment que DIRF-R constitue une base solide, mais encore perfectible, pour des systèmes futurs d'aide à l'expertise automobile.

7.4 Conclusion finale

En conclusion, l'étude montre qu'un pipeline combinant deux détecteurs DETR, une association géométrique explicite par IoU, un Damage Ratio local et une interprétation floue graduelle constitue une approche crédible pour l'évaluation automatisée des dommages automobiles. Le nom DIRF-R ne désigne pas seulement une chaîne technique ; il résume une manière cohérente de penser le problème. Détecter, relier, quantifier puis interpréter : telle est la logique générale qui structure ce mémoire.

Cette organisation améliore la traçabilité méthodologique et permet d'identifier plus clairement l'origine des erreurs produites par le pipeline, car chaque étape est justifiée, limitée et reliée à un besoin précis. C'est aussi ce qui rend le travail accessible à un lecteur non spécialiste : le système n'est pas présenté comme une simple collection de modèles, mais comme un raisonnement progressif appliqué aux images automobiles.

Références

- [1] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov et S. Zagoruyko, « End-to-End Object Detection with Transformers », dans Proc. European Conference on Computer Vision (ECCV), 2020, p. 213-229.
- [2] K. He, X. Zhang, S. Ren et J. Sun, « Deep Residual Learning for Image Recognition », dans Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, p. 770-778.
- [3] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser et I. Polosukhin, « Attention Is All You Need », dans Advances in Neural Information Processing Systems, vol. 30, 2017, p. 5998-6008.
- [4] N. A. Mohammed, M. Y. Potrus et A. M. Ali, « Deep Learning Based Car Damage Classification and Cost Estimation », Zanco Journal of Pure and Applied Sciences, vol. 35, no 1, p. 1-9, 2023.
- [5] J. Qaddour et S. A. Siddiqua, « Automatic Damaged Vehicle Estimator Using Enhanced Deep Learning Algorithm », Intelligent Systems with Applications, vol. 18, art. 200192, 2023.
- [6] R. E. van Ruitenbeek et S. Bhulai, « Convolutional Neural Networks for Vehicle Damage Detection », Machine Learning with Applications, vol. 9, art. 100332, 2022.
- [7] K. Pasupa, P. Kittiworapanya, N. Hongngern et K. Woraratpanya, « Evaluation of Deep Learning Algorithms for Semantic Segmentation of Car Parts », Complex & Intelligent Systems, vol. 8, p. 3613-3625, 2022.
- [8] I. Ullah et K. S. Yow, « Automated Vehicle Dent Detection Using Hybrid Image Processing and Fuzzy Decision Making », Mechatronics and Intelligent Transportation Systems, vol. 4, no 1, p. 49-60, 2025.
- [9] Y. Li, N. Miao, L. Ma, F. Shuang et X. Huang, « Transformer for Object Detection: Review and Benchmark », Engineering Applications of Artificial Intelligence, vol. 126, art. 107021, 2023.
- [10] X. Zhu, W. Su, L. Lu, B. Li, X. Wang et J. Dai, « Deformable DETR: Deformable Transformers for End-to-End Object Detection », dans Proc. International Conference on Learning Representations (ICLR), 2021.

- [11] S. Liu, F. Li, H. Zhang, X. Yang, X. Qi, H. Su, J. Zhu et L. Zhang, « DAB-DETR: Dynamic Anchor Boxes are Better Queries for DETR », dans Proc. International Conference on Learning Representations (ICLR), 2022.
- [12] F. Li, H. Zhang, S. Liu, J. Guo, L. M. Ni et L. Zhang, « DN-DETR: Accelerate DETR Training by Introducing Query DeNoising », dans Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022.
- [13] D. Meng, X. Chen, Z. Fan, G. Zeng, H. Li, Y. Yuan, L. Sun et J. Wang, « Conditional DETR for Fast Training Convergence », dans Proc. IEEE/CVF International Conference on Computer Vision (ICCV), 2021.
- [14] L. A. Zadeh, « Fuzzy Sets », *Information and Control*, vol. 8, no 3, p. 338-353, 1965.
- [15] D. Dubois et H. Prade, *Fuzzy Sets and Systems: Theory and Applications*. New York, NY, USA : Academic Press, 1980.
- [16] R. R. Yager et D. P. Filev, *Essentials of Fuzzy Modeling and Control*. New York, NY, USA : John Wiley & Sons, 1994.
- [17] S. J. Kogbeti-Wongo, « Détection des pièces de véhicules intacts et endommagées à l'aide d'algorithmes d'apprentissage profond (YOLOv8) améliorés par la logique floue », mémoire de maîtrise, Université du Québec à Trois-Rivières, Trois-Rivières, QC, Canada, septembre 2025.
- [18] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn et A. Zisserman, « The Pascal Visual Object Classes (VOC) Challenge », *International Journal of Computer Vision*, vol. 88, no 2, p. 303-338, 2010.
- [19] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid et S. Savarese, « Generalized Intersection over Union: A Metric and a Loss for Bounding Box Regression », dans Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, p. 658-666.
- [20] Z. Zheng, P. Wang, W. Liu, J. Li, R. Ye et D. Ren, « Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression », *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no 7, p. 12993-13000, 2020.

[21] Progi, « Solutions numériques pour l'automobile et l'assurance », documentation institutionnelle officielle. [En ligne]. Disponible : <https://www.progi.com/>. Consulté le 24 mars 2026.

[22] J. Krause, J. Deng, M. Stark et L. Fei-Fei, « Collecting a Large-Scale Dataset of Fine-Grained Cars », dans Proc. Second Workshop on Fine-Grained Visual Categorization (FGVC), 2013.

[23] Roboflow, « Roboflow Documentation », documentation officielle. [En ligne]. Disponible : <https://docs.roboflow.com/>. Consulté le 24 mars 2026.

[24] Python Software Foundation, « Python Documentation », documentation officielle. [En ligne]. Disponible : <https://docs.python.org/>. Consulté le 24 mars 2026.

[25] Google, « Google Drive Help », documentation officielle. [En ligne]. Disponible : <https://support.google.com/drive/>. Consulté le 24 mars 2026.

[26] Google, « Google Colaboratory », documentation officielle. [En ligne]. Disponible : <https://colab.research.google.com/>. Consulté le 24 mars 2026.

[27] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar et C. L. Zitnick, « Microsoft COCO: Common Objects in Context », dans Computer Vision - ECCV 2014, 2014, p. 740-755.

[28] R. Padilla, S. L. Netto et E. A. B. da Silva, « A Survey on Performance Metrics for Object-Detection Algorithms », dans Proc. International Conference on Systems, Signals and Image Processing (IWSSIP), 2020, p. 237-242.

Annexes

ANNEXE A : Liste des pièces détectées de la voiture dans la Figure 5

Élément détecté	Score	Boîte
Rétroviseur gauche	0.9830425977706909	[30 52 48 76]
Phare avant droit	0.9857574701309204	[160 105 208 147]
Capot	0.9955295920372009	[26 60 198 124]
Phare avant droit	0.879694938659668	[165 101 196 120]
Pare-brise	0.9170448184013367	[38 25 187 74]
Toit	0.771710216999054	[10 175 184 195]
Rétroviseur droit	0.9635394215583801	[182 51 198 72]
Toit	0.9921765923500061	[62 19 170 40]
Phare avant gauche	0.9606732130050659	[24 105 60 134]

ANNEXE B : Détail des pièces endommagées, ratios de dommage et niveaux de gravité associés à la Figure 5

Pièces endommagées détectées	Pourcentages de dommages	Catégories
Capot	14,94 % endommagé	Domage mineur
Toit	6,78 % endommagé	Domage mineur
Phare avant gauche	100,00 % endommagé	Domage très sévère

ANNEXE C : Liste des pièces détectées de la voiture dans la Figure 8

Image A

Élément détecté	Score	Boîte
Toit	0.9017642385061947	[144, 20, 338, 54]
Pare-brise	0.8549273816502947	[103, 35, 278, 110]
Capot	0.9415827390648271	[28, 85, 220, 145]
Pare-chocs avant	0.9136805274190638	[5, 145, 234, 258]
Phare avant droit	0.8362047158936271	[144, 110, 230, 180]
Rétroviseur droit	0.8126749305812046	[257, 90, 293, 133]
Porte avant droite	0.7819364250736148	[259, 35, 319, 217]
Roues	0.8937159406281746	[207, 150, 258, 289]
Roues	0.7218463059274183	[323, 132, 357, 216]

Pièces endommagées détectées	Pourcentages de dommages	Catégories
Capot	5,66 % endommagé	Domage mineur
Pare-chocs avant	14,81 % endommagé	Domage mineur
Phare avant droit	100,00 % endommagé	Domage très sévère
Roues	9,29 % endommagé	Domage mineur

Image B

Élément détecté	Score	Boîte
Pare-brise	0.8619472058364917	[50, 7, 260, 87]
Capot	0.9246817395062814	[66, 69, 265, 118]
Capot	0.8815739462071384	[59, 106, 285, 191]
Phare avant gauche	0.8463095728163947	[48, 164, 148, 208]
Phare avant droit	0.8074516293847502	[250, 132, 309, 189]
Pare-chocs avant	0.8927350164827395	[28, 177, 325, 267]

Pièces endommagées détectées	Pourcentages de dommages	Catégories
Capot	72,73 % endommagé	Domage sévère
Capot	100,00 % endommagé	Domage très sévère
Phare avant gauche	72,58 % endommagé	Domage sévère
Phare avant droit	100,00 % endommagé	Domage très sévère
Pare-chocs avant	14,08 % endommagé	Domage mineur

Image C

Élément détecté	Score	Boîte
Toit	0.8037485162093746	[162, 1, 329, 49]
Pare-brise	0.8264517390285641	[102, 25, 260, 113]
Capot	0.9326847150938274	[16, 76, 201, 163]
Pare-chocs avant	0.9051738462071935	[0, 163, 172, 268]
Phare avant gauche	0.8163927405819632	[53, 156, 142, 206]
Phare avant droit	0.7946281739502846	[142, 113, 201, 163]
Rétroviseur droit	0.7716382947502164	[239, 89, 275, 117]
Porte avant droite	0.7358294610372849	[239, 29, 307, 201]
Porte arrière droite	0.7142058396174823	[281, 29, 341, 171]
Roues	0.8649173506827419	[168, 156, 223, 264]
Roues	0.7085362917462058	[316, 104, 348, 161]

Pièces endommagées détectées	Pourcentages de dommages	Catégories
Capot	60,36 % endommagé	Domage sévère
Pare-chocs avant	30,20 % endommagé	Domage modéré
Phare avant gauche	95,93 % endommagé	Domage très sévère

Phare avant droit	88,75 % endommagé	Dommmage très sévère
Roues	28,89 % endommagé	Dommmage modéré

Image D

Élément détecté	Score	Boîte
Pare-brise	0.8319472058364917	[95, 55, 236, 122]
Capot	0.9148265738294612	[6, 99, 175, 167]
Pare-chocs avant	0.8873649205713846	[0, 191, 123, 278]
Phare avant gauche	0.8427156384902731	[36, 145, 131, 196]
Rétroviseur droit	0.7835394215583801	[206, 118, 236, 151]
Porte avant droite	0.7629153846072189	[211, 52, 301, 230]
Porte arrière droite	0.7314928371058274	[283, 40, 327, 207]
Roues	0.8916047283519625	[118, 184, 196, 289]
Roues	0.7246819357204681	[320, 140, 356, 205]

Pièces endommagées détectées	Pourcentages de dommages	Catégories
Capot	41,72 % endommagé	Dommmage modéré
Pare-chocs avant	18,86 % endommagé	Dommmage mineur
Phare avant gauche	95,43 % endommagé	Dommmage très sévère
Roues	4,87 % endommagé	Dommmage mineur