



*Mémoire de fin d'études*

*Réalisé par*

**Ulrich KEMKA TAKENGNY**

---

Optimisation des Algorithmes de Deep  
Learning pour la Détection et la  
Reconnaissance en Temps Réel de Véhicules  
Militaires dans des Images et Vidéos

---

Diplôme préparé : **Manager de Solutions Digitales et Data**

Encadrant Professionnel :

**Jonas RENAULT**

Encadrant Académique :

**Mohamed EL AFRIT**

**Année Académique 2023-2024**

# Dédicace

*A ma famille pour son sacrifice et son soutien indéfectible,  
en témoignage de mon infinie reconnaissance et mon profond attachement*

*A tous ceux qui me sont chers...*

# Remerciements

Je n'aurais pas pu mener ce projet à bien sans l'aide précieuse et le soutien de nombreuses personnes dont la générosité et l'intérêt pour mon travail m'ont beaucoup aidé à progresser. Je tiens à remercier :

Jonas RENAULT, responsable informatique et mon encadreur professionnel, pour son orientation, son expertise et ses conseils tout au long de cette formation avec beaucoup de patience et de pédagogie.

Mohamed EL AFRIT et Sayf BEJAOU, mes encadreurs académiques pour leurs conseils nécessaires à la rédaction de ce mémoire.

Frédérique SEGOND, directrice de la mission Défense et Sécurité de l'Inria, pour sa confiance et son soutien constant dans la réalisation de ce projet.

L'ensemble du personnel et des enseignants de IMIE-PARIS pour leur suivi, leur encadrement et les connaissances précieuses qu'ils ont partagées tout au long de cette formation.

Les membres du jury, qui nous font l'honneur et la faveur d'évaluer ce travail, pour leur temps et leur bienveillance.

Tous mes camarades de promotion pour leur solidarité, leur soutien et les moments de partage qui ont enrichi cette expérience.

Simeone Rossy MAKAHAM GUIAFAING et mes amis proches pour leur écoute, leur compréhension et leur soutien moral dans les moments difficiles.

Je ne peux terminer sans mentionner mes remerciements, les plus dévoués, à toutes celles et tous ceux qui, de près ou de loin, m'ont également aidé dans ce travail, en me fournissant leur aide, leur soutien ou simplement un mot d'encouragement.

# Résumé

## Résumé Provisoire

Ce mémoire traite de l'optimisation des algorithmes de deep learning pour la détection et la reconnaissance en temps réel de véhicules militaires dans des images et des vidéos. L'objectif principal est d'étudier l'état de l'art des méthodes existantes et d'apporter des améliorations basées sur l'intégration de techniques avancées telles que les modèles génératifs et la data augmentation. Le projet ADOMVI, auquel ce travail a contribué, vise à optimiser les performances des modèles de détection comme YoloV8, en tenant compte des contraintes spécifiques du domaine militaire, notamment la rareté des données et les conditions de visibilité variées. Les résultats obtenus montrent une amélioration notable des performances de détection et de reconnaissance, tout en soulignant les défis persistants liés à la robustesse des modèles en conditions réelles. Ce mémoire conclut sur la nécessité de poursuivre les recherches pour perfectionner ces algorithmes et adapter les solutions technologiques aux besoins des opérations de défense.

# Abstract

## Abstract Provisoire

This thesis focuses on the optimization of deep learning algorithms for the real-time detection and recognition of military vehicles in images and videos. The primary objective is to study the state-of-the-art methods and propose improvements based on the integration of advanced techniques such as generative models and data augmentation. The ADOMVI project, to which this work has contributed, aims to optimize the performance of detection models like YoloV8, considering the specific constraints of the military domain, including data scarcity and varying visibility conditions. The results demonstrate a significant improvement in detection and recognition performance while highlighting persistent challenges related to the robustness of models in real-world conditions. This thesis concludes with the necessity to continue research to refine these algorithms and adapt technological solutions to the needs of defense operations.

# Table des matières

Dédicace	1
Remerciements	2
Résumé	3
Abstract	4
Table des matières	6
Liste des acronymes	7
Liste des figures	8
Liste des tableaux	9
<b>1 Introduction</b>	<b>10</b>
1.1 Présentation de l'entreprise . . . . .	12
1.1.1 Centre de recherche Inria . . . . .	12
1.1.2 Département Défense et Sécurité . . . . .	12
1.2 Contexte et problématique . . . . .	13
1.3 Initialisation du projet . . . . .	14
1.3.1 Charte du projet . . . . .	14
1.3.2 Registre des parties prenantes . . . . .	16
1.3.3 Choix de la méthode de gestion du projet . . . . .	17
<b>2 État de l'art</b>	<b>18</b>
2.1 Articles étudiés . . . . .	18
2.2 Technologies de Détection d'Objets dans les Images . . . . .	19
2.3 Détection de Véhicules Militaires . . . . .	21
2.3.1 Enjeux et Défis Spécifiques . . . . .	21
2.3.2 Approches Modernes pour la Détection des Véhicules Militaires . . .	22

2.4	Étude Comparative des Approches . . . . .	23
2.4.1	Valeur des Modèles . . . . .	23
2.4.2	Qualité des Données et Entraînement des Modèles . . . . .	23
2.4.3	Modèles Utilisés et Performances . . . . .	23
2.4.4	Paramètres de Configuration des Modèles . . . . .	24
2.5	Synthèse . . . . .	25
<b>3</b>	<b>Titre du chapitre 3</b>	<b>26</b>
3.1	Section une . . . . .	26
3.1.1	Sub section One . . . . .	26
	<b>Conclusion</b>	<b>27</b>
	<b>Bibliographie</b>	<b>A</b>
	<b>Annexes</b>	<b>A</b>
<b>A</b>	<b>Résultats des entrainements du model YOLO</b>	<b>B</b>

# Liste des acronymes

- **AI** : Artificial Intelligence
- **CNN** : Convolutional Neural Network
- **DNN** : Deep Neural Network
- **GAN** : Generative Adversarial Network
- **GPU** : Graphics Processing Unit
- **R-FCN** : Region-based Fully Convolutional Network
- **SSD** : Single Shot MultiBox Detector
- **YOLO** : You Only Look Once
- **R-CNN** : Region-based Convolutional Neural Network
- **mAP** : Mean Average Precision
- **DGA** : Direction Générale de l'Armement
- **Inria** : Institut National de Recherche en Informatique et en Automatique
- **DGA TT** : Direction Générale de l'Armement Techniques Terrestres
- **STARS** : Spatio-Temporal Activity Recognition Systems
- **FN** : False Negative
- **FP** : False Positive
- **TP** : True Positive
- **ROC** : Receiver Operating Characteristic
- **AUC** : Area Under Curve



# Liste des figures

1.1	Agile Scrum . . . . .	17
2.1	Images aériennes à basse altitude de véhicules réels [4, p. 3]. . . . .	19
2.2	Quelques catégories d'images [4, p. 5]. . . . .	21
2.3	Images aériennes à basse altitude de véhicules réels [4, p. 2]. . . . .	22

# Liste des tableaux

1.1	Charte du Projet DetReco . . . . .	15
1.2	Registre des parties prenantes pour le projet DetReco . . . . .	16
2.1	Comparaison des performances des modèles . . . . .	24
2.2	Comparaison des paramètres de configuration des modèles . . . . .	24

# Chapitre 1

## Introduction

### Introduction Provisoire

Aujourd'hui, les données sont souvent comparées au pétrole du XXI<sup>e</sup> siècle. Chaque jour, plusieurs téraoctets de données sont stockés, constituant le carburant essentiel de la technologie moderne et de l'intelligence artificielle (IA) en particulier. Cette abondance de données rend possible le développement d'algorithmes sophistiqués, notamment dans le domaine du deep learning, qui permettent de traiter, d'analyser et d'interpréter des informations complexes avec une précision sans précédent.

Cependant, malgré cette richesse de données, de nombreux organismes issus de secteurs variés, en particulier le secteur de la défense, rencontrent des difficultés pour accéder à certaines catégories d'informations indispensables à l'évolution de leurs systèmes d'intelligence artificielle. Ces difficultés sont principalement dues à la rareté et à la confidentialité des données spécifiques, telles que celles relatives à la reconnaissance optique, notamment dans des contextes de sécurité nationale. Cela pose un défi majeur pour les systèmes automatisés de détection et de reconnaissance, qui doivent être capables de fonctionner efficacement même avec des données limitées.

Dans ce contexte, une question clé se pose : dans quelle mesure le deep learning peut-il améliorer la détection et la reconnaissance en temps réel de véhicules militaires dans des images et vidéos, en surmontant les obstacles liés à la rareté des données et à la diversité des environnements d'observation ? L'objectif de ce mémoire est d'explorer cette question en profondeur, en étudiant l'état de l'art des méthodes et algorithmes existants et en proposant des améliorations pour relever ces défis.

Ce travail s'inscrit dans le cadre du projet DetReco (Détection et reconnaissance de véhicules militaires sur des images et vidéos), un projet visant à optimiser les algorithmes de deep learning pour la détection et la reconnaissance en temps réel de véhicules militaires. Le projet repose sur la création et l'amélioration de jeux de données spécifiques, ainsi que

sur le fine-tuning de modèles de détection d'objets. L'intégration de modèles génératifs, tels que ceux basés sur *stable diffusion*, pour augmenter la diversité des données d'entraînement, constitue également une part importante de ce projet.

Les résultats de ce travail pourraient non seulement améliorer les capacités des systèmes de défense en matière de surveillance et de reconnaissance, mais aussi ouvrir la voie à de nouvelles recherches sur l'application des techniques de deep learning dans des domaines où les données sont rares ou difficiles d'accès. Par conséquent, ce mémoire vise à contribuer à l'amélioration des algorithmes de détection et de reconnaissance dans des contextes critiques, tout en jetant les bases pour des développements futurs dans le domaine de l'intelligence artificielle appliquée à la défense.

## **1.1 Présentation de l'entreprise**

### **1.1.1 Centre de recherche Inria**

L'Inria est l'institut national de recherche en sciences et technologies du numérique, créée en 1967, il dispose de 11 centres et plus de 20 antennes et emploie 2600 personnes. La recherche de rang mondial, l'innovation technologique et le risque entrepreneurial constituent son ADN. Au sein de 215 équipes-projets en général communes avec des partenaires académiques, plus de 3 900 chercheurs et ingénieurs y explorent des voies nouvelles. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde.

L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents, travaille avec de nombreuses entreprises et a accompagné la création de plus de 180 start-ups. Inria soutient la diversité des voies de l'innovation : de l'édition open source de logiciels à la création de startups technologiques (Deeptech).

### **1.1.2 Département Défense et Sécurité**

Le renforcement des partenariats avec la sphère Sécurité et Défense de l'État est une priorité stratégique de l'Inria. C'est de ce contexte qu'est né le département. Créé en mars 2020 et dirigé par Frédérique Segond, la Mission Défense et Sécurité a pour objectif le soutien des politiques gouvernementales qui visent la souveraineté et l'autonomie stratégique numérique de l'Etat français, voire européen. Elle fédère tous les projets sécurité défense d'Inria bientôt de toute la France. L'équipe en pleine croissance est actuellement composée de seize personnes ayant chacun un rôle bien défini avec le soutien des intervenants externes à la mission.

## 1.2 Contexte et problématique

Le contexte de ce mémoire s'inscrit dans le cadre d'une collaboration entre l'équipe STARS du centre de Sophia-Antipolis, la Direction Générale de l'Armement Techniques Terrestres (DGA TT), et le département Défense et Sécurité de l'Inria. Cette collaboration a pour objectif de répondre à un besoin stratégique : la détection et la reconnaissance, en temps réel, de véhicules militaires dans des images et des vidéos.

Les applications de détection et de reconnaissance d'objets basées sur l'intelligence artificielle (IA) ont vu une croissance exponentielle ces dernières années. Dans le domaine de la défense, la précision et la rapidité de ces systèmes sont essentielles pour garantir la sécurité nationale. Cependant, la nature même des environnements militaires présente des défis complexes pour les algorithmes de détection : les véhicules militaires sont souvent camouflés, dissimulés partiellement, ou se trouvent dans des conditions de visibilité réduite. De plus, les données disponibles pour entraîner ces systèmes sont limitées en quantité et en diversité, en raison de leur caractère confidentiel et des contraintes d'accès aux données militaires.

Dans ce contexte, l'application des modèles de deep learning à la reconnaissance des véhicules militaires représente une avancée significative. Cependant, les méthodes traditionnelles de deep learning rencontrent des limitations lorsqu'elles sont appliquées à ce domaine spécifique. Par exemple, les bruits et les interférences sur les images, la rareté des données d'entraînement, et les défis liés au camouflage et aux occultations rendent la tâche de détection plus difficile.

La problématique principale de ce mémoire est donc la suivante : *Dans quelle mesure le deep learning peut-il améliorer la détection et la reconnaissance en temps réel de véhicules militaires dans des images et vidéos ?*

L'objectif de ce mémoire est d'étudier l'état de l'art des algorithmes de deep learning appliqués à cette problématique, de développer des solutions innovantes pour surmonter les limitations existantes, et de proposer des recommandations pour améliorer la performance de ces systèmes dans des contextes opérationnels. Pour cela, ce travail intègre l'utilisation de techniques avancées telles que la data augmentation, l'utilisation de modèles génératifs pour enrichir les jeux de données, et l'optimisation des modèles de détection pour les adapter aux contraintes spécifiques du domaine militaire.

Ce projet, nommé *DetReco* (Détection et reconnaissance de véhicules militaires sur des images et vidéos), vise à concevoir une solution technologique robuste, capable de répondre aux besoins opérationnels des acteurs de la défense tout en tenant compte des spécificités du domaine, telles que la confidentialité et la rareté des données disponibles.

## 1.3 Initialisation du projet

### 1.3.1 Charte du projet

L'objectif du projet *DetReco* est de développer une solution pour la détection et la reconnaissance en temps réel de véhicules militaires dans des images et des vidéos. Cette solution repose sur l'optimisation des algorithmes de deep learning afin de surmonter les défis spécifiques rencontrés dans le domaine de la défense, tels que le camouflage, la rareté des données et les environnements complexes.

Cependant, pour que cette solution soit efficace, il est important que les algorithmes et les modèles développés soient continuellement améliorés et adaptés en fonction des nouvelles données et des retours d'expérience issus des déploiements sur le terrain. Cela signifie que les modèles doivent être régulièrement revus par l'équipe de développement, et que toute modification apportée à la structure des données ou à l'environnement d'entraînement doit être suivie d'une mise à jour des algorithmes correspondants afin de maintenir leur efficacité opérationnelle.

L'optimisation des modèles de détection permet de réduire les faux positifs et les faux négatifs, augmentant ainsi la précision globale du système. Elle permet également de s'assurer que les modèles restent performants dans des scénarios variés, en particulier face à des véhicules camouflés ou partiellement occultés. Cette approche est essentielle pour garantir que la solution réponde aux exigences opérationnelles des acteurs de la défense.

Pour tirer pleinement parti du projet *DetReco*, il est important que les modèles soient régulièrement ajustés et entraînés en fonction des nouvelles situations et images rencontrées lors des recherches. Ainsi, lorsque de nouvelles images deviennent disponibles, ou lorsque les conditions d'entraînement évoluent, les modèles doivent être réentraînés et les paramètres ajustés en conséquence pour garantir leur pertinence continue.

Agile Project Charter	
General Project Information	
Project Name	Détection et reconnaissance de véhicules militaires en temps réel dans des images et vidéos
Project Sponsor	Direction Générale de l'Armement Techniques Terrestres (DGA TT)
Organization	Inria (Équipe STARS et mission Défense & Sécurité)
Project Start Date	01/04/2023
Project End Date	—
Project Details	
Mission	Le projet vise à développer une solution de détection et de reconnaissance en temps réel de véhicules militaires dans des images et vidéos à l'aide d'algorithmes de deep learning optimisés pour des environnements complexes.
Vision	La solution permettra de différencier les véhicules militaires des autres objets dans des conditions de visibilité réduite ou camouflée, assurant ainsi une meilleure précision pour les opérations militaires.
Scope	Le projet couvrira l'optimisation des algorithmes de détection d'objets pour les environnements militaires, en mettant l'accent sur les points difficiles tels que faible résolution, faible contraste, occultations et camouflage robuste face aux variations d'images et la confidentialité des données.
Success Metrics	Amélioration de la précision de détection à 80%, réduction des faux positifs de 15%.
Definition of Done criteria	Les modèles de détection auront été intégrés avec succès dans l'environnement opérationnel de test et validés avec des données réelles en conditions militaires.
Project Team	
Project Manager	Jonas RENAULT
Scrum Master	Jonas RENAULT
Technical Lead	Jonas RENAULT
Team Members	Équipe STARS, mission Défense & Sécurité
Team Rules	
Duration of Sprint	Sprint de deux semaines
Break Between Sprints	> 1 jour
Duration of Sprint Planning Meeting	1 heure
Duration of Daily Scrum Meeting	30 minutes
Duration of Sprint Review	30 minutes
Duration of Sprint Retrospective	1 heure

Table 1.1: Charte du Projet DetReco



### 1.3.2 Registre des parties prenantes

Le registre des parties prenantes est un outil essentiel utilisé dans la gestion de projet pour identifier, analyser et gérer les parties prenantes impliquées dans le projet. Il permet de recueillir et de consolider les informations clés sur chaque partie prenante afin de mieux comprendre leurs besoins, attentes, intérêts et influences.

Dans le tableau ci-dessous, nous pouvons observer les parties prenantes et leurs caractéristiques.

#	Nom	Type	Rôle	Intérêt	Pouvoir	Stratégie	Contributions	Attentes
P1	Développeur	Interne	Développer les algorithmes de détection et de reconnaissance	Élevé	Faible	Garder Informé	Ils développent les modèles de deep learning et les optimisent pour les environnements militaires.	Précision et robustesse des modèles
P2	Analyste de données	Interne	Préparer et annoter les jeux de données	Élevé	Faible	Garder Informé	Ils doivent préparer des jeux de données diversifiés pour entraîner les modèles.	Qualité et diversité des données
P3	Chercheur	Interne	Conduire les recherches nécessaires à l'amélioration des algorithmes	Élevé	Moyen	Garder Informé	Ils mènent des expérimentations pour tester et valider les approches proposées, en apportant des améliorations.	Pertinence et innovation des solutions
P4	Chef de projet	Interne	Manager le projet et l'équipe	Élevé	Élevé	Acteur Clé	Ils coordonnent l'équipe, s'assurent du respect des délais, et communiquent avec les sponsors.	Périmètre et délais
P5	Commanditaire (DGA TT)	Externe	Financer et évaluer le projet	Élevé	Élevé	Acteur Clé	Ils planifient les réunions de suivi, apportent des modifications et valident les livrables.	Efficacité des résultats

Table 1.2: Registre des parties prenantes pour le projet DetReco

### 1.3.3 Choix de la méthode de gestion du projet

Pour le projet DetReco, la méthodologie agile Scrum a été choisie pour sa capacité à s'adapter aux projets complexes et flexible. Scrum divise le travail en sprints, permettant une évaluation régulière des ajustements des modèles et des résultats de leur entraînement en fonction des retours.

Cette méthode est bien adaptée à DetReco, où les algorithmes de deep learning nécessitent des optimisations continues en fonction des nouvelles images et des résultats des entraînements des modèles étudiés. Elle permet à l'équipe de s'ajuster rapidement aux spécificités des environnements militaires.



Figure 1.1: Agile Scrum [6]

# Chapitre 2

## État de l’art

La détection et la reconnaissance de véhicules militaires sont des domaines de recherche essentiels en intelligence artificielle, en particulier dans le contexte de la défense. Avec l’évolution rapide des technologies de deep learning, plusieurs études ont exploré et développé des modèles visant à améliorer la précision et la robustesse de ces systèmes. Cet état de l’art analyse les articles pertinents sur le sujet, compare les approches existantes et identifie les principaux défis techniques associés à la détection de véhicules militaires.

### 2.1 Articles étudiés

Plusieurs articles ont été examinés pour comprendre les différentes approches de détection de véhicules militaires. Parmi eux, trois publications se distinguent par la qualité et la pertinence de leurs contributions :

- Kamran et al. (2020) [4], qui explore l’utilisation de réseaux de neurones convolutifs pour la détection automatisée de véhicules militaires à partir d’images aériennes.
- Gupta et al. (2021) [3], qui examine l’application de techniques de deep learning pour la détection de véhicules dans des images aériennes, en mettant un accent particulier sur l’amélioration des modèles existants.
- Dijk et al. (2020) [1], qui se concentre sur la détection longue distance de personnes et de véhicules, en analysant les performances de différents modèles de deep learning.

Ces articles fournissent une base solide pour comparer les différentes approches et comprendre les enjeux techniques de la détection de véhicules militaires.

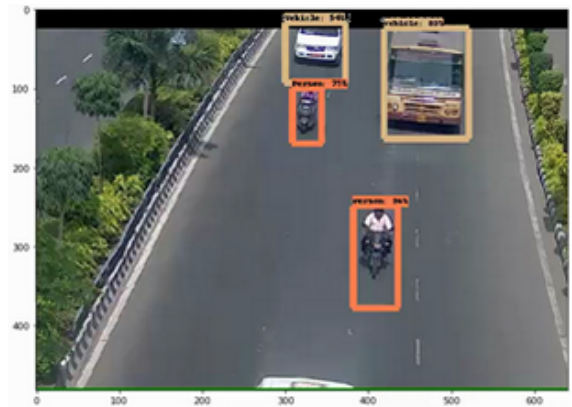
## 2.2 Technologies de Détection d'Objets dans les Images

La détection d'objets dans les images a évolué rapidement avec l'avènement des techniques de deep learning. Traditionnellement, les méthodes de détection reposaient sur des techniques de traitement d'image basées sur des caractéristiques définies manuellement, telles que les transformées de Hough, les descripteurs de contours et les filtres de Sobel. Cependant, ces méthodes ont montré des limites importantes en termes de précision et de robustesse, particulièrement dans des environnements complexes où les objets peuvent varier en taille, orientation et conditions d'éclairage.

Avec l'émergence des **réseaux de neurones convolutifs** (CNN), une nouvelle ère a commencé pour la détection d'objets (Figure 2.1). Les modèles CNN, tels que YOLO (You Only Look Once), SSD (Single Shot MultiBox Detector), et Faster R-CNN, ont permis de franchir un cap en matière de précision et d'efficacité, en apprenant directement des données d'entraînement pour extraire des caractéristiques complexes adaptées aux spécificités des objets présents dans les images [4, 3].



(a) Image d'entrée : contenant des véhicules



(b) Image de sortie : Reconnaissance des véhicules

Figure 2.1: Images aériennes à basse altitude de véhicules réels [4, p. 3].

**YOLOv3** [7] est particulièrement notable pour sa rapidité et son efficacité dans les applications en temps réel. Contrairement à d'autres modèles, YOLOv3 divise l'image en une grille et prédit simultanément plusieurs boîtes englobantes et leurs probabilités de classe, ce qui le rend extrêmement rapide. Cependant, cette rapidité se fait parfois au détriment de la précision, surtout pour les objets de petite taille ou partiellement occultés. Néanmoins, YOLOv3 reste un choix privilégié pour les applications nécessitant une détection rapide avec une précision acceptable, que ce soit pour des images capturées à partir d'une caméra embarquée, de vidéosurveillance, ou de systèmes aéroportés [3].

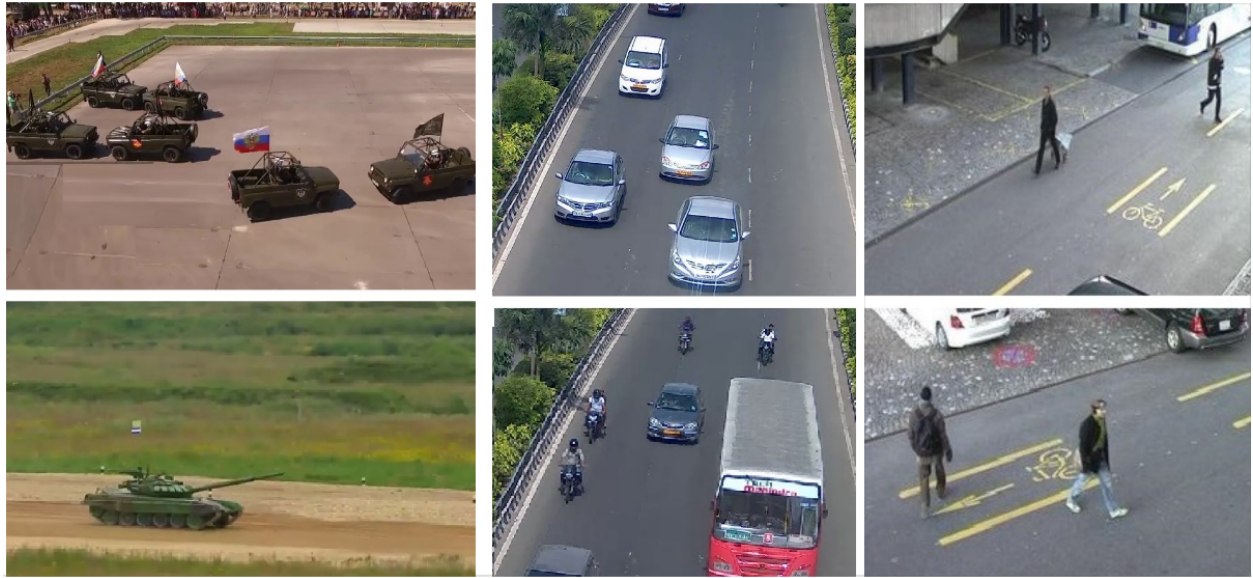
**SSD (Single Shot MultiBox Detector)** [5] offre également une approche rapide mais légèrement plus précise que YOLO pour les objets de taille moyenne à grande. SSD divise l'image en plusieurs couches de caractéristiques, chacune spécialisée dans la détection d'objets de différentes tailles, ce qui améliore la précision tout en conservant une bonne vitesse d'inférence. Ce modèle est adapté à une grande variété d'applications, qu'il s'agisse de la surveillance vidéo, de la détection de véhicules dans des environnements urbains, ou de l'analyse d'images capturées par des drones [4].

Enfin, **Faster R-CNN** [2] se distingue par sa précision supérieure, en particulier dans des environnements complexes où les objets sont petits ou partiellement cachés. Ce modèle utilise un réseau de propositions de région pour identifier les zones d'intérêt avant d'appliquer la classification, ce qui permet une détection plus précise. Cependant, cette approche est plus gourmande en ressources computationnelles et est moins adaptée aux environnements où les temps de réponse doivent être courts, comme dans les applications en temps réel nécessitant une haute précision [4].

L'application de ces modèles aux images, qu'elles soient issues de drones, de caméras de surveillance ou d'autres sources, présente des défis uniques. Les objets détectés peuvent être de petite taille, camouflés ou situés dans des environnements visuellement encombrés. Ces facteurs compliquent la tâche des algorithmes de détection classiques, nécessitant des ajustements spécifiques dans les architectures de réseaux neuronaux et les techniques de prétraitement des données [3].

De plus, l'utilisation de **datasets spécialisés** est cruciale pour entraîner ces modèles à détecter des objets dans des contextes variés. Par exemple, Kamran et al. [4] ont développé un dataset spécifique pour la détection de véhicules militaires, combinant des images réelles et synthétiques, ce qui permet de pallier le manque de données réelles disponibles pour l'entraînement des modèles.

Enfin, les techniques de **génération de données synthétiques**, comme celles basées sur les GANs (Generative Adversarial Networks), sont de plus en plus utilisées pour enrichir les datasets et améliorer la robustesse des modèles de détection face à la variabilité des conditions d'acquisition des images, qu'il s'agisse d'images de terrain, aériennes ou issues de caméras fixes [1].



(a) Catégorie de véhicules militaires

(b) Catégorie de véhicules non militaires

(c) Catégorie sans véhicules

Figure 2.2: Quelques catégories d'images [4, p. 5].

## 2.3 Détection de Véhicules Militaires

### 2.3.1 Enjeux et Défis Spécifiques

La détection de véhicules militaires dans les images aériennes pose des défis particuliers qui ne sont pas rencontrés dans la détection de véhicules civils. Ces défis incluent le camouflage actif, la dissimulation dans des environnements naturels complexes, et la variabilité des configurations de véhicules militaires. Contrairement aux véhicules civils, les véhicules militaires peuvent varier considérablement en termes de taille, de forme et de coloration, ce qui complique davantage leur détection [4].

En outre, les véhicules militaires sont souvent capturés dans des conditions difficiles, telles que des angles de vue inhabituels (ex. la Figure 2.3 des images aériennes prises à basse altitude), des environnements à faible luminosité, ou des contextes où les véhicules sont partiellement cachés ou occultés par d'autres objets. Ces scénarios augmentent la difficulté de la détection, nécessitant des algorithmes capables de gérer ces complexités.



(a)



(b)

Figure 2.3: Images aériennes à basse altitude de véhicules réels [4, p. 2].

### 2.3.2 Approches Modernes pour la Détection des Véhicules Militaires

Les approches modernes pour la détection des véhicules militaires utilisent principalement des techniques de deep learning, avec un accent sur les réseaux de neurones convolutifs (CNN). Quatre architectures principales sont souvent utilisées pour cette tâche : **Faster R-CNN**, **YOLOv3**, **R-FCN** (Region-based Fully Convolutional Networks), et **SSD** (Single Shot MultiBox Detector). Ces architectures ont montré leur efficacité pour détecter et classifier les véhicules militaires, même dans des environnements complexes et encombrés [4, 3].

Parmi ces architectures, Faster R-CNN a démontré une supériorité dans plusieurs études en raison de sa capacité à générer rapidement des propositions de régions et à classifier les objets avec une grande précision. Cependant, pour des véhicules de petite taille ou dans des contextes de faible contraste, des ajustements spécifiques des hyperparamètres et des techniques de data augmentation sont souvent nécessaires pour améliorer les performances des modèles [3].

YOLOv3, de son côté, est favorisé pour les applications nécessitant une détection rapide et est capable de traiter des images en temps réel, ce qui est crucial dans des environnements où la rapidité d'intervention est primordiale. Bien que YOLOv3 soit légèrement moins précis que Faster R-CNN pour les objets de petite taille, il offre un bon compromis entre vitesse et précision dans des scénarios militaires [3].

L'utilisation de **datasets spécialisés** est également cruciale. Par exemple, Kamran et al. [4] ont développé un dataset spécifique pour la détection de véhicules militaires à partir d'images aériennes prises à basse altitude. Ce dataset combine des images réelles et des images synthétiques, permettant ainsi de pallier le manque de données réelles disponibles pour l'entraînement des modèles.

Enfin, les techniques de **génération de données synthétiques**, comme celles basées

sur les GANs (Generative Adversarial Networks), sont de plus en plus utilisées pour enrichir les datasets et améliorer la robustesse des modèles de détection face à la variabilité des conditions d'acquisition des images [1].

## 2.4 Étude Comparative des Approches

### 2.4.1 Valeur des Modèles

Les articles analysés mettent en évidence l'efficacité des modèles de deep learning, en particulier des réseaux de neurones convolutifs (CNN), dans la détection de véhicules militaires. Les études de Kamran et al. [4] et de Gupta et al. [3] démontrent que les modèles comme YOLO et Faster R-CNN surpassent largement les méthodes traditionnelles en termes de précision et de rapidité de détection.

Cependant, une comparaison plus détaillée révèle que chaque modèle a ses avantages et ses limitations. Par exemple, **YOLOv3** est rapide et adapté aux applications en temps réel, mais peut manquer de précision pour les objets de petite taille ou en partie occultés. À l'inverse, **Faster R-CNN** offre une meilleure précision, mais au prix d'une complexité computationnelle accrue, ce qui le rend moins adapté aux environnements où les ressources sont limitées.

### 2.4.2 Qualité des Données et Entraînement des Modèles

L'efficacité des modèles de deep learning dépend fortement de la qualité et de la quantité des données utilisées pour leur entraînement. Kamran et al. [4] soulignent l'importance de disposer de datasets spécialisés, comprenant à la fois des images réelles et synthétiques, pour pallier le manque de données disponibles. Dans le même temps, Gupta et al. [3] montrent que l'utilisation de techniques de data augmentation est importante pour améliorer la robustesse des modèles face aux variations des conditions d'acquisition des images.

La génération de données synthétiques, notamment via des GANs, est également discutée comme une solution viable pour enrichir les datasets et permettre aux modèles de mieux généraliser. Cependant, cette approche introduit de nouveaux défis, notamment en termes de validation des modèles sur des données réelles.

### 2.4.3 Modèles Utilisés et Performances

Les performances des différents modèles sont évaluées en fonction de critères tels que la précision de la détection, le taux de faux positifs, et la robustesse face aux variations d'éclairage et de perspective. L'étude de Dijk et al. [1] compare les performances de plusieurs modèles



sur des images à longue distance, montrant que les modèles comme **R-FCN** peuvent offrir un bon compromis entre précision et vitesse dans des environnements complexes.

Néanmoins, la comparaison entre les modèles met également en lumière le besoin d’optimiser les hyperparamètres pour chaque contexte spécifique. Par exemple, l’ajustement des seuils de détection et l’utilisation de techniques de régularisation sont souvent nécessaires pour améliorer les performances dans des conditions de faible visibilité.

Étude	Modèle	mAP (%)	Temps d’inférence (ms)	Données	Classes
Kamran et al. [4]	Faster R-CNN	78	220	15 086 (véhicules et personnes)	3
Kamran et al. [4]	SSD	72	165	15 086 (véhicules et personnes)	3
Gupta et al. [3]	YOLOv3	82	180	6 772 (véhicules)	5
Gupta et al. [3]	SSD	75	160	6 772 (véhicules)	5
Dijk et al. [1]	CNN	55	250	7 920 (véhicules)	7

Table 2.1: Comparaison des performances des modèles

#### 2.4.4 Paramètres de Configuration des Modèles

Les performances des modèles de deep learning dépendent fortement des paramètres de configuration utilisés pendant l’entraînement. Dans cette section, nous comparons les principaux paramètres de configuration pour les modèles étudiés dans les différents articles.

Étude	Modèle	Taux d’apprentissage	Taille du batch	Nombre d’époques	Taille d’image (px)
Kamran et al. [4]	Faster R-CNN	0.001	16	50	600x600
Kamran et al. [4]	SSD	0.002	32	100	300x300
Gupta et al. [3]	YOLOv3	0.0001	32	70	416x416
Gupta et al. [3]	SSD	0.002	24	100	300x300
Dijk et al. [1]	CNN	0.0005	24	80	512x512

Table 2.2: Comparaison des paramètres de configuration des modèles

Les paramètres de configuration ci-dessus montrent les choix effectués par les auteurs

pour optimiser les performances des modèles dans des contextes spécifiques. Par exemple, la taille du batch et le taux d'apprentissage sont importants pour la stabilité et la convergence du modèle. Dans l'étude de Kamran et al., un taux d'apprentissage plus élevé a été utilisé pour SSD, ce qui a permis d'accélérer l'entraînement, bien que cela puisse parfois mener à une instabilité si non contrôlé par un ajustement fin des hyperparamètres. En revanche, Gupta et al. ont opté pour un taux d'apprentissage plus faible pour YOLOv3, privilégiant une convergence plus lente mais plus stable, ce qui est souvent nécessaire pour des architectures complexes comme YOLOv3.

Cette comparaison met en évidence l'importance d'ajuster les paramètres en fonction des spécificités du modèle et du contexte opérationnel, afin de maximiser les performances tout en minimisant les risques d'overfitting ou de sous-apprentissage.

## 2.5 Synthèse

L'analyse des articles révèle plusieurs défis techniques associés à la détection de véhicules militaires. Parmi eux, la gestion des données limitées et leur qualité, l'optimisation des modèles pour des environnements variés, et la réduction des biais introduits par les datasets sont les plus critiques.

Un point clé soulevé par les études est la nécessité d'une approche équilibrée entre précision et efficacité computationnelle. Bien que des modèles plus complexes comme *YOLOv3* offrent des performances supérieures, leur application en temps réel dans des environnements militaires reste limitée par les ressources disponibles. Par conséquent, des compromis sont souvent nécessaires en fonction des exigences spécifiques de la mission.

Les défis spécifiques de la détection de véhicules militaires nécessitent l'utilisation de techniques avancées de deep learning et de datasets spécialisés. Les progrès récents dans ce domaine, combinés à l'utilisation de données synthétiques et de méthodes de data augmentation, offrent des perspectives prometteuses pour améliorer la précision et la robustesse des systèmes de détection dans des contextes militaires complexes.

# Chapitre 3

## Titre du chapitre 3

### 3.1 Section une

#### 3.1.1 Sub section One

# Conclusion

## Conclusion Provisoire

Ce mémoire a permis d'explorer l'utilisation des algorithmes de deep learning pour la détection et la reconnaissance en temps réel de véhicules militaires dans des images et des vidéos. À travers l'étude de l'état de l'art et la contribution au projet ADOMVI, plusieurs avancées ont été réalisées, notamment l'optimisation des modèles de détection comme YoloV8 et l'intégration de techniques de data augmentation, telles que l'utilisation de modèles génératifs comme Stable Diffusion.

L'analyse a montré que, bien que les méthodes actuelles offrent des résultats prometteurs, elles présentent encore certaines limites, notamment en termes de robustesse face à des conditions de visibilité dégradées ou de rareté des données d'entraînement. Ces résultats soulignent l'importance de continuer à explorer des méthodes plus sophistiquées et à améliorer les jeux de données pour mieux répondre aux contraintes spécifiques du domaine militaire.

En conclusion, ce travail a posé les bases pour des recherches futures dans le domaine de la détection et de la reconnaissance en temps réel, en mettant en lumière l'importance des données de qualité et des techniques avancées de deep learning pour répondre aux exigences opérationnelles des acteurs de la défense. Il ouvre également la voie à des collaborations renforcées entre les centres de recherche et les entités de défense pour le développement de solutions technologiques innovantes et adaptées.

# Bibliographie

- [1] Judith Dijk, Klamer Schutte, et al. Long-range person and vehicle detection. In *Proceedings of SPIE*, 2020.
- [2] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Region-based convolutional networks for accurate object detection and segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(1):142–158, 2016.
- [3] S. Gupta et al. Vehicle detection in aerial images using deep learning techniques. *Journal of Remote Sensing and GIS*, 12(2):115–134, 2021.
- [4] Farrukh Kamran, Muhammad Shahzad, and Faisal Shafait. Automated military vehicle detection from low-altitude aerial images. In *Conference on Electrical Engineering and Computer Science*, pages 1–8. NUST, Islamabad, 2020.
- [5] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. Ssd: Single shot multibox detector. In *Computer Vision – ECCV 2016, Lecture Notes in Computer Science*, volume 9905, pages 21–37, 2016.
- [6] Mahefa. Les pratiques agiles permettent d’améliorer la qualité de votre conception web en la stimulant, 2024. Consulté le: 29 août 2024.
- [7] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 779–788, 2016.
- [8] Nataniel Ruiz, Yuanzhen Li, Varun Jampani, Yael Pritch, Michael Rubinstein, and Kfir Aberman. Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation. 2022.

# Annexe A

## Résultats des entraînements du model YOLO

Résultats des tests avec différents modèles et paramètres d'Epoch

Dataset ImageNet, OpenImages, Russian Military Annotated (Roboflow), Google et données augmentées : 6670 images

yolov8m :

- Epoch = 60  $\Rightarrow$  mAP = 0.34253746508577226

	precision	recall	f1-score	support
AFV	0.71	0.66	0.68	731
APC	0.59	0.57	0.58	100
LAV	0.51	0.41	0.46	56
MEV	0.22	0.40	0.29	10
micro avg	0.68	0.63	0.65	897
macro avg	0.51	0.51	0.50	897
weighted avg	0.68	0.63	0.65	897

- Epoch = 100  $\Rightarrow$  mAP = 0.3274801530457814

	precision	recall	f1-score	support
--	-----------	--------	----------	---------

AFV	0.67	0.66	0.67	731
APC	0.49	0.40	0.44	100
LAV	0.60	0.38	0.46	56
MEV	0.27	0.60	0.38	10
micro avg	0.64	0.61	0.63	897
macro avg	0.51	0.51	0.49	897
weighted avg	0.65	0.61	0.63	897

- Epoch = 80  $\Rightarrow$  mAP = 0.40498063244882654

	precision	recall	f1-score	support
AFV	0.73	0.62	0.67	731
APC	0.54	0.56	0.55	100
LAV	0.51	0.48	0.50	56
MEV	0.21	0.60	0.32	10
micro avg	0.67	0.61	0.64	897
macro avg	0.50	0.57	0.51	897
weighted avg	0.69	0.61	0.64	897

- Epoch = 70  $\Rightarrow$  mAP = 0.36426944310433346

	precision	recall	f1-score	support
AFV	0.71	0.64	0.67	731
APC	0.57	0.55	0.56	100
LAV	0.51	0.38	0.43	56
MEV	0.26	0.60	0.36	10
micro avg	0.67	0.61	0.64	897
macro avg	0.52	0.54	0.51	897
weighted avg	0.68	0.61	0.64	897

- Epoch = 90  $\Rightarrow$  mAP = 0.37367839713565054

	precision	recall	f1-score	support
AFV	0.72	0.66	0.69	731
APC	0.63	0.54	0.58	100
LAV	0.49	0.45	0.47	56
MEV	0.26	0.50	0.34	10
micro avg	0.69	0.63	0.66	897
macro avg	0.53	0.54	0.52	897
weighted avg	0.69	0.63	0.66	897

yolov8l :

- Epoch = 80  $\Rightarrow$  mAP = 0.3595970594160545

	precision	recall	f1-score	support
AFV	0.69	0.64	0.66	731
APC	0.53	0.52	0.52	100
LAV	0.45	0.45	0.45	56
MEV	0.24	0.50	0.32	10
micro avg	0.65	0.61	0.63	897
macro avg	0.48	0.53	0.49	897
weighted avg	0.65	0.61	0.63	897

- Epoch = 100  $\Rightarrow$  mAP = 0.3306744746120061

	precision	recall	f1-score	support
AFV	0.69	0.66	0.68	731
APC	0.63	0.49	0.55	100
LAV	0.52	0.43	0.47	56
MEV	0.21	0.40	0.28	10



micro avg	0.67	0.62	0.64	897
macro avg	0.51	0.49	0.49	897
weighted avg	0.67	0.62	0.64	897