

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

KHADIM SY

**IDENTIFICATION DES ÎLOTS DE CHALEUR VIA DES
ALGORITHMES D'APPRENTISSAGE PROFOND.**

AVRIL 2025

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire, de cette thèse ou de cet essai a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire, de sa thèse ou de son essai.

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, cette thèse ou cet essai. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire, de cette thèse et de son essai requiert son autorisation.

PRÉSENTATION DU JURY

CE MÉMOIRE A ÉTÉ ÉVALUÉ PAR UN JURY COMPOSÉ DE :

M. Jean-Sébastien Dessureault, directeur de recherche, Professeur au
département de Mathématiques et Informatique Université du Québec à Trois-
Rivières

Usef Faghihi, Professeur au département de Mathématiques et Informatique.
Université du Québec à Trois-Rivières

Michel Lemaire, Professeur au département de génie électrique/informatique
Université du Québec à Trois-Rivières

Résumé

Ce mémoire vise à répondre aux défis posés par les îlots de chaleur urbains dans le cadre de la revitalisation urbaine. Les îlots de chaleur représentent un phénomène où certaines zones urbaines accumulent plus de chaleur que les zones environnantes, principalement en raison de la densité des constructions, de l'absence de végétation et des surfaces imperméables. Cette problématique, exacerbée par le changement climatique, a des conséquences importantes sur la santé publique et le confort thermique des habitants.

Pour identifier et évaluer les zones à risque d'îlots de chaleur, ce mémoire des approches de pointe en IA, avec un accent sur la segmentation et la classification d'images. Les réseaux de neurones profonds, tels que U-Net pour la segmentation et ResNet pour la classification, sont utilisés pour analyser des images satellites de zones urbaines ciblées. Ces modèles permettent de détecter des caractéristiques spécifiques liées aux zones de chaleur, comme les matériaux urbains, l'ombrage, la couverture végétale et les plans d'eau.

L'approche U-Net est employée pour segmenter et cartographier les zones critiques, tandis que le modèle ResNet améliore la classification des secteurs présentant des risques potentiels. L'utilisation de Colab facilite l'entraînement de ces modèles sur des ensembles de données étendus sans nécessiter d'infrastructures matérielles coûteuses, garantissant ainsi des performances optimales et une meilleure généralisation des résultats.

Les analyses réalisées apportent des informations essentielles pour la planification urbaine et l'élaboration de stratégies visant à limiter les îlots de chaleur. Enfin, des recommandations sont proposées pour atténuer ces effets, notamment par l'intégration accrue de végétation, l'emploi de matériaux réfléchissants et l'augmentation des surfaces perméables. Cette étude contribue ainsi à une meilleure compréhension des enjeux environnementaux liés aux îlots de chaleur et soutient le développement de solutions durables pour une revitalisation urbaine résiliente.

Abstract

This thesis aims to address the challenges posed by urban heat islands (UHI) in the context of urban revitalization. Urban heat islands refer to areas within cities that retain more heat than their surroundings, primarily due to high building density, lack of vegetation, and the prevalence of impermeable surfaces. This phenomenon, intensified by climate change, has significant implications for public health and urban thermal comfort.

To identify and assess high-risk UHI zones, this research leverages cutting-edge artificial intelligence techniques, with a focus on image segmentation and classification. Deep learning models, such as U-Net for segmentation and ResNet for classification, are applied to satellite images of urban environments. These models enable the detection of key heat-related characteristics, including urban materials, shading, vegetation cover, and water bodies.

The U-Net model is employed to segment and map critical zones, while ResNet enhances the classification of areas with potential heat risks. The use of Colab facilitates the training of these models on extensive datasets without requiring costly computing infrastructure, ensuring optimal performance and improved generalization of results.

The findings provide essential insights for urban planning and the development of strategies to mitigate UHI effects. Finally, several recommendations are proposed to reduce urban heat islands, including increasing vegetation cover, using reflective materials, and promoting permeable surfaces. This study contributes not only to a better understanding of environmental challenges related to UHI but also supports the implementation of sustainable and resilient urban revitalization solutions.

Remerciements

Avant toute chose, je tiens à exprimer ma profonde gratitude à DIEU, qui m'a accordé la force, la patience et la persévérance nécessaires pour mener à bien ce travail. L'aboutissement de ce mémoire est le fruit d'un long cheminement, ponctué d'efforts soutenus, de remises en question et d'apprentissages enrichissants. Toutefois, ce parcours n'aurait pas été possible sans le soutien et l'implication précieuse de nombreuses personnes, à qui j'adresse mes sincères remerciements.

À ma famille, je témoigne une reconnaissance infinie, en particulier à mon père adoptif Mamadou Ndoumbé Sy, ma mère Adjaratou Arame Guéye, ainsi qu'à mon épouse bien-aimée Amina Cissé Sarr, dont le soutien indéfectible a été un véritable moteur tout au long de cette aventure ainsi que toute sa famille. Vos encouragements constants, vos sacrifices silencieux et votre présence rassurante ont été essentiels dans les moments de doute et d'incertitude. À mes frères, sœurs, oncles et tantes, je vous remercie pour votre confiance inébranlable et votre affection qui ont constitué une source de motivation inestimable.

Je tiens également à exprimer ma gratitude envers mes amis, véritables compagnons de route dans cette aventure exigeante. Votre présence, votre soutien moral et vos encouragements ont été d'une aide précieuse, transformant les défis en opportunités d'apprentissage et les moments difficiles en souvenirs mémorables. Vous avez su apporter cette légèreté nécessaire, rendant le parcours plus agréable et stimulant.

Ma reconnaissance va tout particulièrement à mon encadreur, Jean-Sébastien Dessureault, dont les conseils avisés, la rigueur méthodologique et la patience ont guidé mes réflexions et enrichi ma démarche scientifique. Vos encouragements et votre disponibilité constante ont contribué à faire de ce travail une réalisation aboutie. Je remercie également Alain Thierry Iliho Manzi, pour son appui précieux et ses observations pertinentes qui ont permis d'affiner et de consolider les résultats obtenus.

Enfin, je dédie ce travail à mon fils adoré Mouhamed Bachir Sy, dont l'innocence et la joie de vivre ont été une source d'inspiration constante, mes amis chers Boucunta Ndiaye, Sokhna Khady Fall, Mansour Mbodji, M. Moustapha Fall et El Hadji Rawane Guéye, pour leur soutien indéfectible et leur confiance inébranlable ainsi que mes frères OUSSEYNOU et BATHIE SY.

Une pensée émue va à la mémoire de mon défunt père Moustapha Sy et de Fatima Sy, dont les valeurs et les enseignements continuent de m'inspirer et de guider mes pas dans la vie.

Ce mémoire est l'aboutissement d'un effort collectif, porté par le soutien et l'implication de chacun. À vous tous, je vous exprime ma plus profonde gratitude et mon immense respect.

TABLES DE MATIERES

Résumé.....	iii
Abstract	iv
Remerciements.....	v
Listes des figures.....	x
Listes des tableaux	xi
Liste des abréviations.....	xii
Chapitre 1 : Introduction.....	1
1.1 Évolution des technologies et opportunités offertes	1
1.2 Intérêt de l'étude et choix du cas d'étude.....	2
1.3 Contexte et problématique.....	3
1.4 Objectifs	3
1.5 Méthodologie.....	4
1.6 Organisation du mémoire	5
1.7 Conclusion.....	6
Chapitre 2 : Revue de littérature	8
2.1 Reconnaissance d'images.....	8
2.2 Identification et localisation des ICU.....	11
2.2.1 Identification des zones d'îlots de chaleur.....	11
2.2.2 Localisation des zones d'îlots de chaleur	12
2.3 Les îlots de chaleur urbains : causes et impacts.....	13
2.4 L'IA pour la détection des ICU	15
2.4.1 Modèles de réseaux neuronaux utilisés.....	15
2.5 Conclusion.....	18
Chapitre 3 : Méthodologie et analyses des données	20
3.1 Méthodologie.....	20
3.1.1 Approche par la classification.....	20
3.1.2 Approche par la segmentation	21
3.1.3 Approche par la détection	23

3.2	Présentation des modèles d'apprentissages automatiques utilisés.....	25
3.3	Les types d'apprentissages automatiques.....	26
3.3.1	Apprentissage non supervisé.....	26
3.3.2	Apprentissage semi-supervisé.....	27
3.3.3	Apprentissage supervisé.....	27
3.3.4	Le Modèle ResNet	28
3.3.4.1	Architecture.....	29
3.3.5	Le Modèle U-Net	31
3.3.5.1	Architecture.....	32
3.3.6	Le Modèle YOLOv8	33
3.3.6.1	Architecture.....	34
3.4	Analyse de données	36
3.4.1	Préparation des données pour la classification	37
3.4.2	Préparation des données pour la segmentation	38
3.4.3	Préparation des données pour la détection.....	40
3.5	Autres Outils.....	41
3.5.1	Google Colaboratory.....	41
3.5.2	Google Drive.....	41
3.5.3	CVAT	42
3.5.4	Apeer	42
3.5.5	Python	43
3.6	Conclusion.....	44
Chapitre 4 : Résultats et discussions		46
4.1	Détection et segmentation des îlots de chaleur	47
4.1.1	Performance des modèles.....	47
4.1.2	Résultats des prédictions.....	59
4.2	Résultats de la segmentation d'image des îlots de chaleur	59
4.2.1	Résultats obtenus avec u-net.....	60
4.2.2	Résultats obtenus avec U-Net+ResNet.....	61

4.3	Détection des îlots de chaleur avec le modèle YOLOv8	63
4.3.1	Performance du modèle	64
4.3.2	Résultats de la détection.....	68
4.4	Conclusion.....	71
Chapitre 5 : Conclusion finale et perspectives.....		73
5.1	Synthèse des résultats.....	73
5.2	Contributions de l'étude	74
5.3	Limites de l'étude	76
5.4	Perspectives d'amélioration	77
5.5	Conclusion générale	79
Bibliographie.....		81

Listes des figures

FIGURE 1: SEGMENTATION SEMANTIQUE(VAHAGN TUMANYAN 2023)	22
FIGURE 2:EXEMPLE DE DETECTION D'OBJETS EN MILIEU URBAIN A L'AIDE DE LA VISION PAR ORDINATEUR(VISION PLATFORM)	24
FIGURE 3: ARCHITECTURE RESTNET50(SUVADITYA MUKHERJEE)	30
FIGURE 4: ARCHITECTURE DE U-NET (RONNEBERGER & ALL)	32
FIGURE 5: ARCHITECTURE YOLOV8 SIMPLIFIEE(ARWAABOUATTIA)	34
FIGURE 6: ÉVOLUTION DE LA FONCTION PERTE - U-NET.....	50
FIGURE 7: ÉVOLUTION DE L'EXACTITUDE - U-NET	51
FIGURE 8:MATRICE DE CONFUSION OBTENUE APRES L'ÉVALUATION DU MODELE U- NET.....	53
FIGURE 9:ÉVOLUTION DE LA FONCTION PERTE - U-NET+RESNET	55
FIGURE 10:ÉVOLUTION DE L'EXACTITUDE - U-NET+RESNET	57
FIGURE 11: EXEMPLE DE PREDICTION DES ILOTS DE CHALEUR SUR UNE IMAGE URBAINE	61
FIGURE 12: U-NET + RESNET MONTRANT UNE BONNE COUVERTURE GENERALE AVEC QUELQUES ERREURS AUX LIMITES.	62
FIGURE 13: EVOLUTION DE LA PERTE (YOLOV8).....	65
FIGURE 14: ÉVOLUTION DES METRIQUES (PRECISION, RAPPEL ET F1-SCORE - YOLOV8).....	67
FIGURE 15: PREDICTION CORRECTE D'ILOT DE CHALEUR AVEC YOLOV8	69
FIGURE 16: MAUVAISE PREDICTION D'ILOT DE CHALEUR AVEC YOLOV8.....	70

Listes des tableaux

TABLEAU 1: SYNTHÈSE DES MODÈLES DE ML ET LEURS APPLICATIONS DANS LA DETECTION DES ILOTS DE CHALEUR.....	26
TABLEAU 2: RÉPARTITION DES DONNÉES	37
TABLEAU 3: PROPORTION DE RÉPARTITION DES DONNÉES	38
TABLEAU 4: TABLEAU DE L'AUGMENTATION DES DONNÉES.....	39
TABLEAU 5: RÉSULTATS OBTENUS POUR CHAQUE MODÈLE	48

Liste des abréviations

IA : Intelligence Artificielle

CNN : *Convolutional Neural Network* (Réseau de Neurones Convolutifs)

CVAT : *Computer Vision Annotation Tool*

GPU : *Graphics Processing Unit* (Unité de Traitement Graphique)

ICU : Îlots de Chaleur Urbains

IoU : *Intersection over Union* (Intersection sur Union)

ML : *Machine Learning* (Apprentissage Automatique)

RGB : *Red, Green, Blue* (R: Rouge, V: Vert, B: Bleu)

U-Net : *U-shaped Network* (réseau en forme de "U")

ResNet : *Residual Network* (Réseau Résiduel)

YOLO : *You Only Look Once* (Modèle de détection d'objets en temps réel)

IoT : *Internet of Things* (Internet des Objets)

RS : *Remote Sensing* (Télédétection)

ReLU : *Rectified Linear Unit* (Unité Linéaire Rectifiée)

FPN : *Feature Pyramid Network* (Réseau pyramidal de caractéristiques)

PAN : *Path Aggregation Network* (Réseau d'agrégation des chemins)

Chapitre 1 : Introduction

Introduction

Les phénomènes climatiques extrêmes et les problématiques environnementales ont pris une ampleur considérable au cours des dernières décennies, suscitant une attention croissante de la part des chercheurs, des décideurs politiques et du grand public. Parmi ces enjeux, les ICU (Îlots de Chaleur Urbains) occupent une place centrale, en raison de leurs impacts négatifs sur le bien-être des populations et sur l'environnement. Ce phénomène se caractérise par des températures nettement plus élevées dans les zones urbanisées que dans les zones rurales avoisinantes. Cette différence thermique s'explique principalement par plusieurs facteurs : la densité des constructions, la faible présence de végétation, la prédominance de matériaux imperméables et l'intensité des activités humaines.

L'importance de cette problématique s'accroît dans un contexte marqué par une urbanisation rapide et continue ainsi que par les effets du changement climatique [1]. Les ICU aggravent les vagues de chaleur, augmentent la consommation énergétique liée à la climatisation, et accentuent les risques pour la santé publique [25], notamment pour les populations vulnérables comme les personnes âgées et les enfants [2]. De plus, ils contribuent à la détérioration de la qualité de l'air et à la formation de polluants secondaires, ce qui renforce les enjeux de santé environnementale [24]. Par conséquent, la gestion de ce phénomène devient cruciale pour garantir la résilience des villes face aux défis climatiques.

Face à cette situation, la détection précise des îlots de chaleur revêt une importance capitale. Elle constitue une étape essentielle pour la mise en place de stratégies efficaces d'atténuation et d'adaptation, telles que l'augmentation des espaces verts [31], l'utilisation de matériaux de construction réfléchissants et la promotion d'aménagements urbains durables. Traditionnellement, la détection des ICU s'appuyait sur des mesures de terrain ou des capteurs fixes, mais ces méthodes présentent des limites importantes : elles sont coûteuses, chronophages et offrent une couverture spatiale limitée.

1.1 Évolution des technologies et opportunités offertes

Grâce aux progrès des technologies numériques, en particulier dans les domaines de la télédétection RS (Remote Sensing) et de l'intelligence artificielle (IA), il est désormais possible de pallier les limites des approches conventionnelles. L'exploitation des images satellites permet d'obtenir une couverture spatiale étendue et des données de haute résolution, facilitant

ainsi une analyse approfondie des environnements urbains. Ces images intègrent des informations multispectrales et thermiques précieuses pour identifier les zones susceptibles d'être affectées par ICU.

L'IA, et plus précisément l'apprentissage profond, a révolutionné les techniques d'analyse des images. Les Réseau de Neurones Convolutifs, qui figurent parmi les architectures fondamentales de l'apprentissage profond, se révèlent particulièrement efficaces pour le traitement des images. Ils permettent d'extraire automatiquement des caractéristiques complexes à partir des données visuelles et de détecter des motifs subtils associés aux ICU, tels que la densité du bâti ou la rareté des espaces verts.

1.2 Intérêt de l'étude et choix du cas d'étude

Ce mémoire se concentre sur l'application des technologies d'apprentissage profond dans la détection des îlots de chaleur urbains à partir d'images satellitaires. L'intérêt de cette recherche repose sur l'opportunité qu'offrent les techniques avancées d'analyse d'images pour répondre à un enjeu environnemental et urbain de grande envergure. L'intelligence artificielle combinée aux données géospatiales permet d'automatiser l'identification des zones à risque et de proposer des solutions adaptées à divers contextes urbains.

L'adoption de cette thématique est motivée par la nécessité croissante d'une meilleure compréhension et gestion des îlots de chaleur urbains, en particulier dans les zones sujettes à une urbanisation rapide et confrontées à des épisodes de chaleur de plus en plus fréquents et intenses. Ce phénomène, amplifié par les changements climatiques, a des répercussions considérables sur la santé publique, la consommation énergétique et la qualité de vie des populations. À Londres, il a été montré que les ICU augmentent significativement la demande énergétique estivale [41].

Cette étude ambitionne également de proposer une méthodologie reproductible, applicable à différents environnements urbains, quelle que soit leur localisation ou leur échelle. L'objectif est de concevoir une approche souple et adaptable, garantissant des résultats précis et pertinents dans divers contextes géographiques.

Ainsi, cette recherche s'inscrit dans une double perspective : d'une part, concevoir une méthode automatisée et robuste pour la détection des îlots de chaleur urbains à partir d'images satellitaires et, d'autre part, suggérer des solutions concrètes pour en atténuer les effets. Ces solutions pourraient inclure l'accroissement de la végétation, l'utilisation de matériaux

réfléchissants et la réorganisation des espaces urbains afin d'optimiser la gestion thermique des villes.

1.3 Contexte et problématique

Les ICU désignent un phénomène dans lequel les zones urbaines enregistrent des températures plus élevées que les zones rurales environnantes. Ce phénomène est principalement attribué aux activités humaines, à l'utilisation de matériaux retenant la chaleur et à l'insuffisance de végétation. Les ICU ont des conséquences significatives sur la santé publique, en augmentant les risques associés aux vagues de chaleur, sur la consommation énergétique [42], ainsi que sur l'environnement urbain en général ¹.

Malgré leur importance, l'identification et l'analyse des ICU demeurent complexes. Les approches traditionnelles, telles que les relevés de terrain, bien que fiables, sont coûteuses et limitées en termes de couverture spatiale. Les données géo-spatiales, qui fournissent des informations précieuses sur l'environnement urbain, nécessitent des outils avancés pour être exploitées efficacement. De plus, les modèles d'analyse actuels rencontrent des difficultés à s'adapter aux spécificités locales des différentes zones urbaines, rendant ainsi difficile la mise en place de stratégies adaptées à chaque contexte. Certaines recherches aussi montrent que la combinaison entre ICU et vagues de chaleur produit des effets synergiques particulièrement néfastes pour les zones urbaines [11].

Face à ces défis, la problématique à laquelle cette étude s'attache est la suivante : Comment tirer parti de l'IA pour détecter avec précision les îlots de chaleur urbains, tout en proposant des recommandations pertinentes pour en atténuer les effets ?

Ce mémoire a pour objectif de développer une approche technologique innovante, capable de fournir des solutions efficaces et adaptées aux réalités locales, afin de contribuer à une planification urbaine durable et résiliente.

1.4 Objectifs

L'objectif principal de ce mémoire est de développer une méthode basée sur l'IA pour détecter les ICU et proposer des solutions adaptées afin de réduire leurs impacts. Pour y parvenir, il est

¹ Wikipédia, « Apprentissage automatique », Wikipédia. 26 décembre 2024. Disponible sur: https://fr.wikipedia.org/w/index.php?title=Apprentissage_sautomatique&oldid=221466068

essentiel d'analyser en profondeur les causes et les conséquences des ICU dans divers contextes urbains, en tenant compte des facteurs environnementaux et anthropiques qui les aggravent.

Cette recherche repose sur la collecte et le traitement de données géo-spatiales, telles que les images satellites et thermiques, qui serviront à alimenter un modèle d'IA². Ce modèle, fondé sur des réseaux neuronaux Convolutifs, est conçu pour détecter et cartographier avec précision les zones les plus exposées aux ICU.

Une fois le modèle validé et les zones à risque identifiées, des recommandations pratiques seront formulées. Ces recommandations viseront à atténuer les effets des ICU en mettant en œuvre des solutions telles que l'augmentation des espaces verts, l'utilisation de matériaux réfléchissants et la réorganisation des infrastructures urbaines en fonction des spécificités locales.

Ainsi, ce mémoire combine une approche technologique avancée et une application concrète afin de contribuer à une planification urbaine durable et résiliente face aux défis posés par le réchauffement climatique.

1.5 Méthodologie

La méthodologie adoptée dans cette étude repose sur plusieurs étapes clés visant à assurer une détection précise et efficace des ICU :

Collecte des données : Les images satellitaires utilisées proviendront de bases de données publiques et seront complétées par des données thermiques et multi spectrales [12]. Ces images subiront un prétraitement visant à améliorer leur qualité et à assurer une homogénéité optimale pour une analyse approfondie.

Annotation des données : Des outils spécialisés, tels que CVAT (Computer Vision Annotation Tool) et Apeer³, seront employés pour générer des masques d'annotation précis. Cette étape est

² <https://www.kaggle.com/code/yaroslavnaychuk/satellite-image-segmentation>

³ « arivis Cloud - automated image analysis ». Disponible sur: <https://www.apeer.com/home>

essentielle pour garantir la pertinence des données d'entraînement et améliorer les performances des modèles de ML (Machine Learning)⁴[11].

Entraînement des modèles : Différents modèles d'apprentissage profond seront développés et entraînés afin d'exécuter les tâches de classification, de segmentation et de détection des ICU [5]. Les architectures retenues incluront des CNN (Convolutional Neural Network) spécialement adaptés à l'analyse des images géo-spatiales.

Évaluation des performances : Les modèles seront soumis à des jeux de données de validation et de test afin d'évaluer leur efficacité. Les performances seront mesurées à l'aide de métriques standardisées telles que la précision, le rappel, le score F1 (mesure de la performance combinant précision et rappel)⁵ et l'IoU (Intersection sur Union) [23]. Une comparaison entre les modèles permettra d'identifier la meilleure approche pour la détection des ICU.

Analyse des résultats et recommandations : Une analyse approfondie des résultats obtenus sera réalisée afin d'identifier les tendances et les défis liés à la détection des ICU. Sur cette base, des recommandations pratiques seront formulées pour proposer des stratégies adaptées aux réalités locales et aux bonnes pratiques en matière d'aménagement urbain durable.

1.6 Organisation du mémoire

Ce mémoire est organisé de manière à fournir une réponse structurée et cohérente aux objectifs de la recherche, tout en offrant une compréhension approfondie des ICU. Le premier chapitre établit le cadre général de l'étude en présentant les enjeux liés aux ICU dans le contexte des changements climatiques et de la planification urbaine. Ce chapitre définit les objectifs spécifiques de la recherche et décrit la méthodologie utilisée pour aborder la problématique.

⁴ J. Robert, « Python : Focus sur le langage le plus populaire », Formation Data Science | DataScientest.com. le 17 Avr 2023 Disponible sur: <https://datascientest.com/python-tout-savoir>

⁵ B. Buckley, « Qu'est-ce qu'un score F1 ? - Académie EITCA », EITCA Academy. publié le: 2023-08-09. [En ligne]. Disponible sur: <https://fr.eitca.org/intelligence-artificielle/eitc-ai-gcm1-google-cloud-apprentissage-automatique/introduction/qu%27est-ce-que-l%27apprentissage-automatique/qu%27est-ce-qu%27un-score-f1/>

Le deuxième chapitre est consacré à la revue de littérature, où les concepts clés relatifs aux ICU sont explorés. Il s'attarde sur leurs causes, leurs impacts, ainsi que sur les différentes techniques de détection disponibles. Il met également en lumière l'évolution des modèles d'IA, particulièrement les CNN, et leur application croissante dans l'analyse de données géospatiales et thermiques pour la détection des ICU.

Dans le troisième chapitre, la méthodologie et l'analyse des données sont détaillées. Cette section décrit de manière précise les différentes étapes du processus de recherche, depuis la collecte des données (telles que les images satellites et les données thermiques) jusqu'au prétraitement et à l'entraînement des modèles d'IA. Elle explique également les outils et techniques utilisés pour évaluer et valider les performances des modèles développés.

Le quatrième chapitre présente les résultats obtenus par les modèles d'IA, en mettant en évidence les zones identifiées comme étant sujettes aux ICU. Une analyse critique des résultats est réalisée en les confrontant à des études antérieures, soulignant à la fois les points forts de la méthodologie utilisée et ses limites.

Enfin, le cinquième chapitre propose une conclusion générale qui résume les contributions majeures du mémoire. Il discute de l'impact de cette recherche dans la gestion des ICU et propose des perspectives pour de futurs travaux. Ces pistes incluent l'intégration de nouvelles sources de données et l'application de la méthodologie à d'autres environnements urbains, afin d'élargir l'ampleur de l'étude et d'approfondir les connaissances sur ce phénomène complexe.

1.7 Conclusion

L'analyse présentée dans ce chapitre a permis de situer la problématique des ICU dans un contexte à la fois environnemental, social et technologique. Les ICU représentent un défi majeur pour les collectivités, notamment en raison de leurs impacts sur la santé publique, la qualité de vie et l'aménagement du territoire, en particulier dans les zones urbaines en croissance ou en transformation.

En exposant les limites des méthodes classiques de détection et de cartographie des ICU, ce chapitre a mis en lumière la pertinence d'intégrer des approches fondées sur l'IA. L'essor de l'apprentissage profond, et plus spécifiquement de techniques comme la segmentation d'images avec U-Net (U-shaped Network) et la classification avec ResNet (Residual Network),

offre des perspectives nouvelles pour automatiser, fiabiliser et enrichir l'analyse des milieux urbains à partir de données visuelles.

Dans cette optique, ce mémoire propose une méthodologie reposant sur l'exploitation d'images satellitaires et l'implémentation de modèles d'IA afin de détecter les zones à risque de chaleur urbaine. Pour bien situer cette contribution dans le champ scientifique, le chapitre suivant sera consacré à un état de l'art. Il dressera un portrait des connaissances actuelles sur les îlots de chaleur, les approches de détection existantes, ainsi que les avancées en vision par ordinateur appliquées à l'environnement urbain.

Chapitre 2 : Revue de littérature

Introduction

Les évolutions récentes dans les domaines de la RS et de l'IA ont permis de grandes avancées dans l'analyse des phénomènes environnementaux complexes, tels que les ICU. Ce phénomène, causé par une combinaison de facteurs anthropiques et géographiques, constitue un défi majeur pour les chercheurs et les urbanistes, en raison de ses effets négatifs sur la santé publique, la consommation énergétique et la qualité de vie des habitants⁶.

Avec la disponibilité croissante de données d'imagerie satellite et l'amélioration des algorithmes d'apprentissage profond, de nouvelles méthodes d'identification, de localisation et de cartographie des ICU ont vu le jour [9]. Ces approches modernes offrent une précision accrue et une meilleure couverture spatiale par rapport aux méthodes traditionnelles, telles que les relevés in situ et les capteurs fixes, souvent coûteuses et limitées en portée. Une vue d'ensemble des travaux récents souligne l'évolution rapide des approches scientifiques pour caractériser les ICU [6].

Cette revue de littérature s'intéresse aux principaux concepts et techniques utilisés pour la détection des ICU. Elle est structurée autour des axes suivants : la reconnaissance d'images, l'identification et la localisation des ICU, ainsi que les modèles d'apprentissage profond appliqués à ce type de problématique. De plus, cette revue mettra en lumière les limites des approches existantes et les perspectives d'amélioration à travers l'utilisation des nouvelles technologies et méthodes innovantes.

2.1 Reconnaissance d'images

La reconnaissance d'images constitue un domaine clé de la vision par ordinateur [12], visant à détecter et identifier des objets ou des motifs particuliers dans des images numériques en leur attribuant des étiquettes spécifiques. Ce domaine a connu un essor considérable grâce à l'introduction des CNN, qui sont une classe de modèles d'apprentissage profond particulièrement efficaces [40] pour le traitement et l'analyse des données visuelles, et à

⁶ Sarah, « Îlots de chaleur urbains (ICU) : Comprendre ce phénomène et les solutions pour le contrer », Siradel. 19 Septembre 2024 Disponible sur: <https://www.siradel.com/fr/comprendre-ilot-de-chaleur-urbain-et-les-solutions-de-lutte/>

l'accessibilité de vastes ensembles de données, tels qu'ImageNet. Ces avancées ont révolutionné l'analyse d'images et ouvert la porte à de nombreuses applications dans des secteurs variés tels que l'urbanisme, la santé et l'environnement. La télédétection thermique a joué un rôle fondamental dans l'analyse climatique urbaine, notamment grâce à ses capacités d'observation spatiale à large échelle [15]. Voogt et Oke ont démontré l'efficacité de la télédétection thermique pour évaluer la variabilité spatiale des ICU à l'échelle intra-urbaine [17].

À l'origine, la reconnaissance d'images se limitait à des tâches simples, comme la lecture automatique de caractères (OCR). Un tournant majeur a été amorcé avec l'introduction des réseaux de neurones convolutifs (CNN), conçus par Yann LeCun dans les années 1990, notamment avec le modèle LeNet pour la reconnaissance de chiffres manuscrits [52]. Cette architecture, inspirée du cortex visuel humain, a permis d'extraire automatiquement des caractéristiques visuelles hiérarchiques à partir des images, supprimant ainsi le besoin d'ingénierie manuelle des caractéristiques.

Depuis, les CNN sont devenus la pierre angulaire de la reconnaissance d'images moderne. Des architectures plus complexes et profondes, telles qu'AlexNet, VGG, ResNet ou EfficientNet, ont considérablement amélioré les performances dans des tâches variées, allant de la classification à la segmentation sémantique et à la détection d'objets. Ces avancées sont largement dues à l'augmentation de la capacité des modèles, à la disponibilité de jeux de données massifs, et à la puissance croissante du calcul parallèle via les GPU (Graphics Processing Unit).

Les progrès en reconnaissance d'images reposent donc essentiellement sur l'évolution continue des modèles de réseaux neuronaux profonds, qui ont permis d'atteindre des niveaux de précision inégalés et d'ouvrir la voie à des applications dans des domaines aussi variés que la médecine, la surveillance, ou encore la conduite autonome.

Introduit en 2012 par Krizhevsky et ses collaborateurs, AlexNet a démontré la capacité des CNN à traiter d'importants volumes de données d'images. Il repose sur l'utilisation de filtres Convolutifs de petite taille et exploite la puissance des GPU pour accélérer l'entraînement. Ce modèle a marqué une avancée majeure en réduisant de manière significative le taux d'erreur sur la base de données ImageNet [53].

AlexNet a introduit plusieurs innovations, notamment l'utilisation de la fonction d'activation ReLU (Rectified Linear Unit) pour accélérer la convergence et la technique de régularisation dropout afin de limiter le surapprentissage du modèle.

Développé par Simonyan et Zisserman en 2014, le modèle VGG a amélioré les performances en approfondissant la structure du réseau tout en conservant une architecture cohérente basée sur des couches convolutives de taille 3×3 . Grâce à sa capacité à extraire des caractéristiques de manière hiérarchique, ce modèle est devenu une référence en matière de classification d'images.

L'atout majeur de VGG réside dans sa simplicité architecturale, qui en facilite l'adaptation à divers contextes d'utilisation. Toutefois, cette approche entraîne une augmentation significative du nombre de paramètres, ce qui exige davantage de ressources en termes de mémoire et de calcul [54].

Proposé par He et al. en 2015, ResNet introduit la notion de blocs résiduels, une innovation permettant l'entraînement de réseaux neuronaux profonds comptant jusqu'à 152 couches [19]. Cette architecture résout les problèmes liés à la disparition du gradient, offrant ainsi des performances stables même pour des architectures très profondes.

Les blocs résiduels de ResNet permettent de contourner certaines couches du réseau, facilitant ainsi l'apprentissage de relations complexes sans perte de performance. Grâce à ces innovations, ResNet⁷ est aujourd'hui largement utilisé dans diverses applications de vision par ordinateur.

Développé par Tan et Le en 2019, EfficientNet propose une approche d'optimisation globale en équilibrant simultanément la profondeur, la largeur et la résolution des réseaux neuronaux [14]. Cette stratégie de mise à l'échelle permet de maximiser les performances tout en minimisant les ressources nécessaires, ce qui en fait un choix privilégié pour les applications nécessitant une analyse à grande échelle, comme l'imagerie satellitaire.

⁷ M. Abdullah, « Comparison and Architecture of pre-trained model (VGG-16, VGG-19, ResNet, GoogleNet, AlexNet... », Medium. [En ligne]. Disponible sur: <https://medium.com/@muhabd51/comparison-and-architecture-of-pre-trained-model-vgg-16-vgg-19-resnet-googlenet-alexnet-0fb459fff368>

EfficientNet repose sur un cadre d'optimisation qui ajuste automatiquement les différentes dimensions du réseau afin d'obtenir une efficacité maximale en termes de calcul et de consommation de mémoire.

Les architectures mentionnées ci-dessus constituent aujourd'hui des références incontournables pour la reconnaissance d'images et sont largement utilisées dans divers domaines, y compris pour la détection des ICU. Grâce à l'évolution constante des réseaux neuronaux et à la disponibilité croissante de données de haute qualité, les techniques de reconnaissance d'images continuent de s'améliorer et de proposer des solutions adaptées aux défis environnementaux contemporains [16].

2.2 Identification et localisation des ICU

Les techniques d'identification d'objets ont été parmi les premières à être appliquées en IA, et elles ont évolué avec l'avancement des techniques d'apprentissage profond et de vision par ordinateur, permettant des applications modernes dans des domaines aussi variés que la surveillance, la conduite autonome, la reconnaissance faciale, et la détection d'anomalies dans des environnements complexes.

La détection des îlots de chaleur urbains repose sur deux étapes fondamentales : l'identification des zones à température élevée et leur localisation précise sur une carte. Ces étapes sont essentielles pour permettre une analyse détaillée et une planification urbaine adaptée.

2.2.1 Identification des zones d'îlots de chaleur

L'identification des îlots de chaleur urbains (ICU) repose sur la capacité à détecter des zones présentant une anomalie thermique par rapport à leur environnement immédiat. Pour ce faire, deux grandes approches de traitement d'images sont privilégiées dans la littérature scientifique : la détection d'objets par boîtes englobantes et la segmentation sémantique.

D'une part, les méthodes de détection d'objets utilisent des boîtes englobantes pour localiser les régions d'intérêt dans une image. Le modèle YOLO (You Only Look Once), proposé par Redmon et al. [59], divise une image en une grille et effectue des prédictions de présence et de localisation d'objets en un seul passage dans le réseau. Ce fonctionnement à passage unique rend YOLO particulièrement rapide et adapté au traitement en temps réel de grandes quantités de données. Une autre approche performante, RetinaNet, introduite par Lin et al. [60], améliore la précision sur les classes rares en intégrant une fonction de perte appelée focal loss. Enfin,

Faster R-CNN, développé par Ren et al. [61], repose sur un système en deux étapes combinant des propositions régionales avec un réseau de classification, offrant une meilleure précision au prix d'une charge computationnelle plus importante.

D'autre part, la segmentation sémantique permet d'assigner une classe à chaque pixel de l'image, offrant ainsi une délimitation fine des zones d'intérêt. Le modèle U-Net, introduit par Ronneberger et al. [62], est particulièrement adapté à la segmentation d'images satellitaires grâce à son architecture symétrique en encodeur-décodeur qui permet la reconstruction détaillée des contours. De son côté, SegNet, conçu par Badrinarayanan et al. [63], utilise également une architecture d'encodeur-décodeur, avec une stratégie de mémorisation des indices de max-pooling pour améliorer la précision spatiale lors de la phase de décodage. Ces modèles permettent une meilleure compréhension spatiale des ICU et facilitent la génération de cartes thermiques exploitables par les urbanistes.

Cependant la relation entre les changements d'usage du sol et l'intensité des ICU est bien documentée dans la littérature [13].

2.2.2 Localisation des zones d'îlots de chaleur

Des modèles numériques permettent de simuler le comportement thermique de la ville en tenant compte des paramètres environnementaux et urbains [3]. La détection et la cartographie des îlots de chaleur urbains (ICU) reposent largement sur des approches de segmentation sémantique appliquées à des images satellitaires ou aériennes. Parmi les architectures les plus répandues dans ce domaine figurent U-Net et DeepLab. Le modèle SegNet propose également une architecture efficace d'encodeur-décodeur pour segmenter précisément les structures urbaines [32].

Le modèle U-Net, initialement proposé par Ronneberger et al. [55] pour la segmentation biomédicale, s'est rapidement imposé comme un standard dans le traitement d'images environnementales. Sa structure symétrique en encodeur-décodeur, enrichie de connexions de saut, permet une reconstruction précise des contours des objets, ce qui le rend particulièrement adapté à la détection fine dans des environnements urbains complexes. De son côté, DeepLab, introduit par Chen et al. [56], repose sur l'utilisation de convolutions dilatées (trous convolutions) permettant de capturer un contexte spatial plus étendu sans perte de résolution voir aussi [7], [20] pour une présentation détaillée de DeepLab et ses variantes. Cette capacité

est essentielle pour la segmentation de structures urbaines diffuses comme les zones asphaltées, les toits ou les sols imperméables.

Afin d'améliorer la performance de ces modèles, plusieurs stratégies sont documentées dans la littérature. L'augmentation des données, telle que présentée par Perez et Wang [57], permet d'accroître la diversité du jeu d'entraînement par l'application de transformations géométriques et photométriques (rotation, réflexion, modification de la luminosité, etc.). Cette diversité renforce la robustesse des modèles face à la variabilité des conditions réelles d'acquisition. L'augmentation des données permet d'améliorer la robustesse des modèles de détection dans des environnements urbains complexes [33].

Par ailleurs, la régularisation joue un rôle crucial pour limiter le surapprentissage, notamment dans le cas de petits jeux de données annotés. Le dropout [58] consiste à désactiver aléatoirement certaines unités neuronales pendant l'entraînement, ce qui permet de réduire la co-adaptation entre neurones et aussi permet de limiter le surapprentissage lors de l'entraînement des modèles de segmentation [34]. Combiné à la régularisation L2, cette méthode contribue à améliorer la généralisation du modèle sur des données non vues. La régularisation par décroissance du poids (weight decay) aussi constitue également une stratégie efficace pour améliorer la généralisation des modèles [35] mais une version améliorée de cette technique, le decoupled weight decay, permet une régularisation plus efficace dans les optimisateurs modernes comme Adam [36].

Enfin, l'utilisation de validations croisées constitue une bonne pratique méthodologique pour évaluer la robustesse des modèles. En divisant le jeu de données en plusieurs sous-ensembles d'entraînement et de validation, cette méthode permet de vérifier que les performances obtenues ne sont pas dues à un biais d'échantillonnage. Les Fully Convolutional Networks ont marqué un tournant dans la segmentation d'images, en particulier dans le contexte de la localisation des ICU [29].

2.3 Les îlots de chaleur urbains : causes et impacts

Les îlots de chaleur urbains (ICU) résultent d'une synergie entre des facteurs environnementaux et anthropiques qui contribuent à une augmentation significative des températures dans les zones densément urbanisées [64]. L'un des facteurs principaux est l'omniprésence de matériaux de construction imperméables tels que le béton, l'asphalte et les toitures sombres. Ces matériaux, caractérisés par une forte capacité thermique et un albédo

faible, absorbent une grande quantité de rayonnement solaire durant la journée et restituent progressivement cette chaleur la nuit, limitant ainsi le refroidissement naturel des environnements urbains [65].

Parallèlement, la réduction de la couverture végétale accentue le phénomène. L'absence d'espaces verts et d'arbres empêche le déroulement optimal de l'évapotranspiration, un mécanisme naturel essentiel au refroidissement de l'air ambiant, par lequel l'humidité du sol est transférée vers l'atmosphère. Cette fonction régulatrice, bien documentée dans les travaux de Jenerette et al., est particulièrement cruciale dans les environnements urbains fortement minéralisés [66].

Les activités humaines intensives contribuent également à l'amplification du phénomène. Les émissions de chaleur générées par la circulation automobile, les processus industriels ou encore l'utilisation massive des systèmes de climatisation, forment ce qu'on appelle la chaleur anthropique. Celle-ci crée un effet rétroactif : l'élévation des températures entraîne une augmentation de la demande en climatisation, qui à son tour intensifie les rejets de chaleur [67]. Ce phénomène est également renforcé par l'effet combiné de l'urbanisation rapide et du changement climatique, comme le confirment plusieurs études de synthèse récentes [5]. Une étude japonaise souligne que l'élévation des températures urbaines peut être un indicateur pertinent du réchauffement climatique à l'échelle locale [22].

Les conséquences des ICU se manifestent à divers niveaux. D'un point de vue sanitaire, les températures extrêmes exacerbent les risques de maladies cardiovasculaires et respiratoires, en particulier chez les populations les plus vulnérables (personnes âgées, enfants, personnes souffrant de pathologies chroniques). Ces impacts ont été largement documentés, notamment en ce qui concerne l'augmentation de la mortalité lors d'épisodes de chaleur extrême [68].

Sur le plan énergétique, les ICU entraînent une augmentation marquée de la consommation d'électricité, principalement liée à l'usage accru de la climatisation. Cette surcharge du réseau peut entraîner des pics de consommation, voire des pannes électriques, tout en augmentant les coûts énergétiques et les émissions de gaz à effet de serre associées [69]. Les ICU influencent directement aussi la performance énergétique des bâtiments [10], ce qui accentue la consommation d'énergie urbaine [8].

Enfin, les effets environnementaux ne sont pas négligeables. Les ICU perturbent les microclimats urbains et affectent la biodiversité locale. Les variations thermiques modifient les

cycles de reproduction, les comportements et les habitats des espèces végétales et animales urbaines, réduisant leur capacité à s'adapter aux changements environnementaux [70].

Ces constats appellent à la mise en place de stratégies d'atténuation fondées sur un urbanisme durable, intégrant des solutions telles que l'augmentation de la végétation, l'utilisation de matériaux à haut albédo et une meilleure gestion énergétique des infrastructures urbaines.

2.4 L'IA pour la détection des ICU

L'application de l'IA, et plus particulièrement des réseaux de neurones profonds, a permis d'améliorer de manière significative la détection et la cartographie des ICU. En automatisant le traitement d'images satellitaires ou thermiques, ces modèles réduisent considérablement le temps nécessaire à l'analyse et facilitent la production de résultats exploitables à grande échelle [71]. Le recours aux techniques de Big Data et de ML permet d'exploiter de vastes ensembles de données urbaines pour modéliser les ICU avec une précision accrue [27].

2.4.1 Modèles de réseaux neuronaux utilisés

Les Convolutional Neural Networks (CNN) sont particulièrement bien adaptés à l'analyse des données issues de la télédétection. Leur capacité à extraire automatiquement des caractéristiques hiérarchiques à partir d'images à haute résolution permet une classification précise des zones selon leurs propriétés thermiques [72]. Les CNN, grâce à leurs couches convolutives successives, détectent des motifs à différentes échelles spatiales et peuvent s'adapter à divers types de données (images thermiques, multispectrales ou infrarouges), en ajustant le nombre de canaux d'entrée [73].

Toutefois, l'entraînement de réseaux neuronaux profonds peut se heurter à des difficultés [21], notamment liées à l'initialisation des poids, comme l'ont démontré Glorot et Bengio [38]. Afin d'améliorer la stabilité de l'apprentissage, Ioffe et Szegedy ont introduit la technique de batch normalization, qui réduit les variations internes au sein des couches profondes [39]. Le modèle Network in Network remplace certaines couches convolutives classiques par des micro-réseaux, augmentant ainsi la capacité de discrimination locale [43]. L'ouvrage de Bengio, Goodfellow et Courville constitue une référence essentielle pour comprendre les fondements des réseaux neuronaux profonds [28]. Des travaux fondamentaux tels que celui de LeCun, Bengio et Hinton ont également jeté les bases du deep learning moderne [81]. Le livre «Deep

Learning with Keras » de Gulli et Pal offre également une introduction pratique aux réseaux neuronaux profonds [95].

Dans le but d'améliorer la généralisation des modèles, des architectures hybrides ont été développées, combinant par exemple U-Net [18] et ResNet. Les réseaux résiduels profonds ont été améliorés par l'introduction des connexions d'identité, facilitant la propagation du gradient dans les architectures très profondes [44]. Le réseau DenseNet, introduit par Huang et al., connecte chaque couche à toutes les couches suivantes, favorisant la réutilisation des caractéristiques [45]. Le modèle Xception, fondé sur des convolutions séparables en profondeur, améliore l'efficacité sans compromettre la performance [46]. EfficientNet, conçu par Tan et Le, optimise simultanément la profondeur, la largeur et la résolution du réseau [47]. Ce type d'architecture hybride a montré une efficacité notable pour la détection des ICU [74].

Les réseaux à double flux (Two-Stream CNN), proposés par Simonyan et Zisserman, exploitent des flux parallèles pour extraire simultanément les caractéristiques spatiales et temporelles [48]. Des variantes telles que U-Net++ et DeepLabV3+ introduisent des connexions denses et des convolutions dilatées pour mieux capter les détails thermiques fins [75]. CCNet exploite l'attention croisée pour une segmentation plus précise des relations contextuelles [60], tandis que Hybrid Task Cascade combine détection et segmentation pour améliorer les performances globales [61]. PSPNet intègre un mécanisme pyramidal pour mieux contextualiser les scènes urbaines complexes [62].

Les réseaux convolutifs multi-échelles se montrent particulièrement efficaces pour classifier des scènes satellitaires complexes [49]. Zhang et al. ont démontré que le deep learning permet de classifier finement l'occupation du sol à partir d'images haute résolution [50]. L'utilisation de données multispectrales et d'images satellites haute résolution, comme chez Liu et al., renforce la précision de classification des usages urbains [51]. Ma et al. proposent une méthode de fusion de données multisources afin de stabiliser les prédictions [52]. De même, l'adaptation de U-Net aux données multispectrales permet une meilleure différenciation des types de couverture terrestre [53]. Zhang et al. ont également utilisé le deep residual U-Net pour extraire avec précision des formes complexes comme les routes [54]. Li et al. montrent l'intérêt de l'intégration multi-échelle dans la classification des surfaces [55].

L'atrous convolution, revisitée par Chen et al., permet d'effectuer une segmentation sémantique détaillée sans perte de résolution [63]. OcNet exploite le contexte global de la scène pour

renforcer la cohérence de la segmentation [64]. Les réseaux résiduels dilatés élargissent le champ réceptif sans altérer la résolution spatiale [65]. Wu et al. soulignent l'importance d'un équilibre entre profondeur et largeur pour optimiser les architectures ResNet [66]. L'introduction des Fully Convolutional Networks (FCN) a permis la segmentation de bout en bout sans couches entièrement connectées [67]. Enfin, le dataset ADE20K a permis une évaluation standardisée de nombreux modèles de segmentation [68].

Les images multispectrales, particulièrement celles intégrant l'infrarouge thermique, jouent un rôle fondamental dans la caractérisation des ICU [79]. Les modèles utilisant des données multi-sources obtiennent des performances élevées [76]. Le dataset SEN12MS constitue une ressource précieuse pour l'entraînement de modèles sur des données géoréférencées [82]. Voogt et Oke ont également montré l'efficacité de la télédétection thermique pour analyser le climat urbain [84]. Des études comme celle de Liu et al. démontrent l'efficacité de l'apprentissage profond pour détecter les ICU à partir de données multiples [83]. Zheng et al. ont montré que la fusion de données multi-sources permet de surveiller l'environnement thermique urbain avec une précision accrue [85]. L'outil CVAT facilite l'annotation de ces données complexes [89].

Grad-CAM est souvent utilisé pour visualiser les régions activées par le modèle lors de la prise de décision [70]. Zeiler et Fergus ont également proposé une méthode de visualisation des activations internes [71]. Mahendran et Vedaldi ont introduit une méthode d'inversion des activations pour visualiser les représentations internes [100]. La plateforme Apeer, devenue Arivis, offre des outils puissants pour interpréter les résultats sur des données d'imagerie scientifique [90].

Pour améliorer la robustesse, des techniques comme l'*ensemble learning* sont utilisées pour agréger les prédictions de plusieurs modèles [77]. La *focal loss*, introduite par Lin et al., traite efficacement le déséquilibre de classes [58]. Les modèles auto-supervisés permettent d'utiliser des jeux de données non annotés, facilitant leur application à des contextes de données rares [80].

Des outils comme PyTorch [87] et TensorFlow [94] facilitent la mise en œuvre et l'entraînement des modèles. Le *framework* PyTorch, introduit par Paszke et al., facilite le développement de modèles profonds flexibles et performants grâce à son mode impératif et sa gestion optimisée des calculs [91]. Des environnements *cloud* comme Google Cloud Platform

(via les travaux de Bisong [88]) permettent une scalabilité et un déploiement rapide. Ces méthodes s'appuient également sur des bibliothèques de calcul scientifique comme NumPy [96], Pandas [97], Matplotlib [98] et Seaborn [99], essentielles pour le traitement et la visualisation des données.

Les synthèses récentes sur la segmentation sémantique montrent l'évolution rapide des architectures [101]. Enfin, YOLOv8, par Swathi et Challa, représente une avancée majeure dans la détection d'objets en temps réel grâce à sa rapidité et précision [102], [103]. YOLOv8, comme présenté par Wang et al., introduit des innovations structurantes dans la détection rapide et efficace en environnement urbain dense [92].

2.5 Conclusion

La revue de littérature menée dans les sections précédentes a permis de mettre en évidence les avancées majeures réalisées dans le domaine de la détection des ICU grâce aux techniques d'IA, et en particulier aux réseaux de neurones profonds. Ces approches ont profondément modifié les méthodologies traditionnelles d'analyse des images satellitaires et thermiques en introduisant des capacités de traitement automatique [78], d'extraction de caractéristiques complexes et d'analyse à grande échelle [81].

Parmi ces contributions, les CNN se sont imposés comme une architecture de référence, notamment pour leur capacité à apprendre des représentations hiérarchiques d'images complexes, qu'elles soient issues du visible, de l'infrarouge ou du multispectral [82]. Ces modèles ont été exploités avec succès pour la détection fine des ICU, en particulier dans des environnements urbains denses où les variations thermiques sont fortement hétérogènes [83].

Cependant, les approches conventionnelles, telles que les méthodes statistiques ou les techniques fondées sur des règles expertes, montrent leurs limites dès lors qu'il s'agit de traiter des données à large échelle, hétérogènes et temporellement variables [84]. Ces approches peinent également à s'adapter à la diversité des contextes géographiques et climatiques, ce qui compromet leur capacité de généralisation.

L'IA, à l'inverse, s'avère particulièrement efficace pour traiter de grandes quantités de données multisources (images satellitaires, données topographiques, météorologiques, etc.) et pour automatiser les processus d'analyse. Cette capacité d'intégration et d'adaptation ouvre la voie

à des systèmes intelligents capables de fournir des cartes thermiques précises et évolutives, favorisant ainsi une meilleure compréhension des dynamiques urbaines [85].

Néanmoins, malgré ces avancées, plusieurs défis subsistent. D'une part, la généralisation des modèles reste une problématique clé, notamment dans les régions où peu de données annotées sont disponibles. D'autre part, l'intégration de nouvelles sources de données, comme les capteurs IoT (Internet of Things) ou les données socio-économiques, représente une opportunité pour enrichir les modèles, mais nécessite de développer des architectures capables de gérer cette complexité croissante [86].

Ces constats soulignent la nécessité d'une approche méthodologique rigoureuse, combinant des techniques d'optimisation de modèles, de validation croisée et d'analyse comparative, pour évaluer l'efficacité réelle des outils d'IA dans le contexte spécifique des ICU. Le chapitre suivant est ainsi consacré à la présentation de la méthodologie mise en œuvre dans cette étude, depuis la collecte et le prétraitement des données jusqu'au développement du modèle de détection et à l'évaluation de ses performances.

Chapitre 3 : Méthodologie et analyses des données

Dans ce chapitre, nous présentons la méthodologie développée pour aborder la problématique de détection et de localisation des îlots de chaleur urbains dans le cadre de ce mémoire. Nous débutons par une description des algorithmes d'apprentissage profond utilisés pour la segmentation et la classification des zones thermiques dans les images urbaines, en mettant un accent particulier sur les modèles tels que U-Net et ResNet. Ensuite, nous détaillons les approches adoptées pour le traitement et la préparation des données, incluant les étapes de prétraitement, d'augmentation des données et de génération des masques thermiques nécessaires à l'entraînement des modèles.

La solution proposée vise à automatiser l'identification des zones à risque à l'aide de techniques de vision par ordinateur⁸ et d'IA. Cette démarche permet non seulement d'accélérer l'analyse des données thermiques et spatiales, mais aussi d'améliorer la précision des cartes de risque, offrant ainsi un outil précieux pour l'aménagement urbain et la planification des interventions climatiques.

3.1 Méthodologie

Pour résoudre la problématique de la détection des îlots de chaleur, nous explorons trois approches complémentaires : la classification, la segmentation et l'identification. La **classification** utilise des modèles comme ResNet pour distinguer les zones affectées des environnements normaux ; la **segmentation** applique des réseaux tels que U-Net afin de diviser précisément les images et localiser les zones de chaleur ; et l'**identification** analyse les segments obtenus pour valider les régions à risque et générer des cartes exploitables en vue d'interventions ciblées.

3.1.1 Approche par la classification

La classification est une tâche de ML où un modèle attribue des étiquettes ou des catégories aux données en fonction de caractéristiques spécifiques⁹. Dans le cadre de notre mémoire sur

⁸ Abirami Vina « Tout ce que tu dois savoir sur la vision par ordinateur en 2025 ». Disponible sur: <https://www.ultralytics.com/fr/blog/everything-you-need-to-know-about-computer-vision-in-2025>

⁹ I. Costa, « Types et applications apprentissage automatique | KAIZENTM ». Consulté le: 29 décembre 2024. Disponible sur: <https://kaizen.com/fr/publications/types-applications-apprentissage-automatique/>

la détection des îlots de chaleur, l'approche par classification consiste à prédire les zones présentant des températures élevées, en les classant comme zones de chaleur ou non-chaleur.

Nous utilisons ici la classification binaire, où le modèle distingue deux classes : les zones avec des îlots de chaleur et les autres zones. L'objectif est de déterminer avec précision les zones urbaines où les risques d'îlots de chaleur sont les plus élevés, afin de cibler des actions de gestion et de prévention.

3.1.2 Approche par la segmentation

L'approche par la segmentation consiste à diviser une image en plusieurs régions homogènes ou segments, qui permettent de mieux comprendre et analyser son contenu. Cette méthode est cruciale dans notre étude, car elle permet d'isoler les zones à forte chaleur dans les images thermiques ou les images prises par satellite afin de détecter et localiser les îlots de chaleur¹⁰. La segmentation est utilisée pour distinguer les zones présentant des températures anormalement élevées (îlots de chaleur) des autres parties de l'image, qui peuvent correspondre à des zones moins chaudes, telles que les espaces verts, les zones d'eau ou les bâtiments à faible émission de chaleur. La segmentation sémantique est une technique qui, plutôt que de simplement séparer l'image en sections, attribue des interprétations sémantiques distinctes à chaque région de l'image comme indique dans la Figure 1 ci-dessous.

¹⁰R. Dwivedi, « My experiment with UNet - building an image segmentation model », Analytics India Magazine. 24 juillet 2020. Disponible sur: <https://analyticsindiamag.com/deep-tech/my-experiment-with-unet-building-an-image-segmentation-model/>

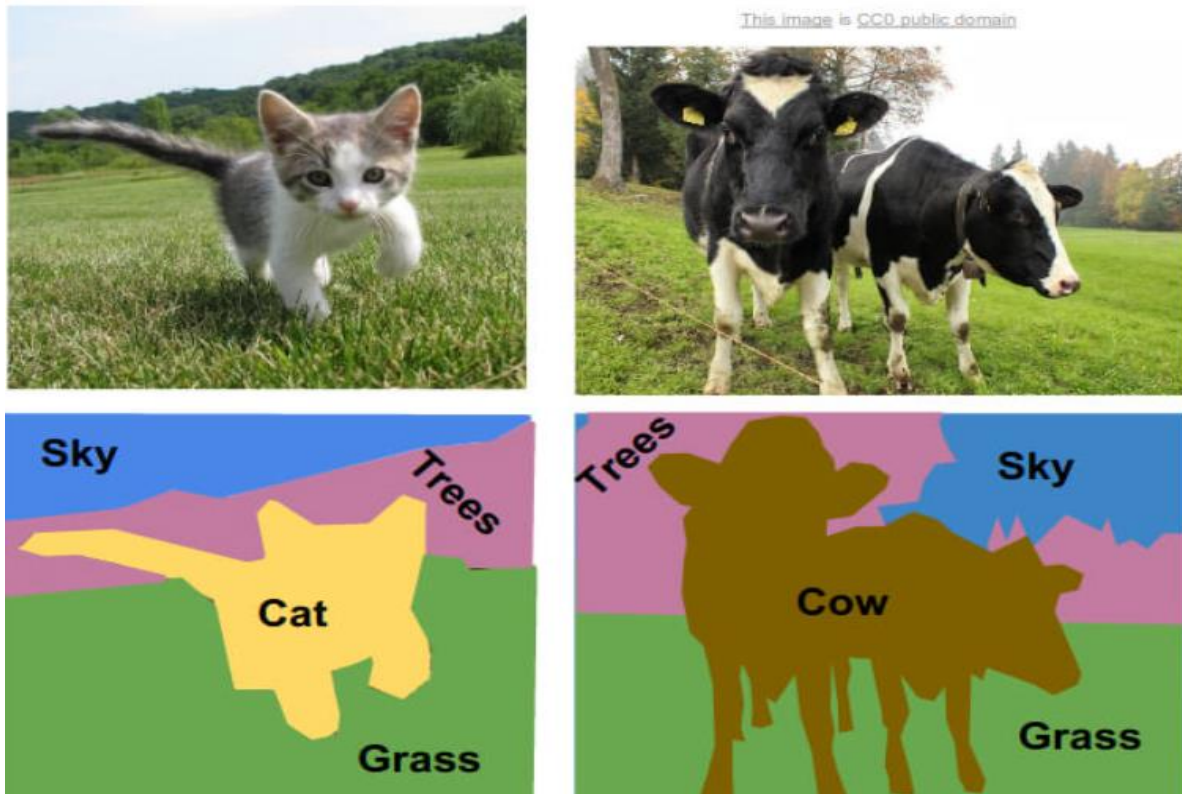


Figure 1: Segmentation sémantique [\[101\]](#)

Il existe plusieurs méthodes de segmentation, mais dans notre étude, nous utilisons principalement la segmentation sémantique, une technique avancée qui permet de classer chaque pixel de l'image en fonction de sa classe d'appartenance. Cela signifie qu'au lieu de simplement diviser l'image en zones homogènes, nous attribuons une signification à chaque pixel, en lui associant une étiquette correspondant à un type particulier de zone, comme un îlot de chaleur, une zone de végétation, une zone d'eau, etc. Cette approche est particulièrement adaptée aux images thermiques, où la différence de température entre les différentes zones est marquée et permet une segmentation précise.

La segmentation par seuil pourrait également être utilisée dans notre cas, en se basant sur une valeur seuil de température pour distinguer les zones chaudes des autres. Cependant, cette méthode peut ne pas être aussi précise que la segmentation sémantique, car elle ne tient pas compte de la variation locale des températures. Elle est donc souvent combinée avec d'autres techniques pour améliorer la qualité du découpage.

Dans la segmentation sémantique, chaque pixel est attribué à une classe, ce qui permet de générer des masques d'images où chaque zone de l'image a une étiquette unique. Par exemple, un pixel dans une zone urbaine à haute température sera étiqueté comme "îlot de chaleur",

tandis qu'un pixel dans une zone verte sera étiqueté comme "végétation". Cela permet de mieux comprendre la distribution des îlots de chaleur dans une zone géographique et d'identifier précisément les endroits où des interventions peuvent être nécessaires pour réduire la chaleur urbaine.

Pour cette tâche, nous utilisons des modèles de CNN et des architectures comme U-Net [6], qui sont bien adaptés pour la segmentation sémantique, notamment dans les tâches de segmentation d'images thermiques ou satellites. U-Net se distingue par sa capacité à effectuer une segmentation précise, même avec un nombre limité d'images d'entraînement, grâce à son architecture symétrique avec des couches de contraction et d'expansion. Cela permet de maintenir un contexte spatial tout en affinant la segmentation au niveau des pixels.

L'un des avantages de la segmentation sémantique est qu'elle permet une analyse fine des caractéristiques de chaque zone, rendant ainsi possible la détection et l'évaluation des îlots de chaleur de manière plus détaillée que d'autres méthodes de segmentation plus simples. Cette précision est essentielle pour la gestion des îlots de chaleur, car elle aide à identifier les zones spécifiques à surveiller ou à traiter pour réduire la température locale et améliorer la qualité de vie urbaine.

3.1.3 Approche par la détection

Notre troisième approche se concentre sur la détection d'objets à l'aide du modèle YOLO. Cette tâche consiste à identifier et localiser des éléments spécifiques au sein d'une image. La détection d'objets, qui constitue une tâche essentielle en vision par ordinateur, a profondément transformé la manière dont les systèmes d'IA analysent les informations visuelles. Contrairement à la **classification d'images**, où l'objectif est de classer une image entière en une seule catégorie, la détection d'objets permet non seulement d'identifier des objets spécifiques au sein d'une image ou d'une séquence vidéo, mais aussi de les localiser précisément. Ce processus implique la reconnaissance des objets, leur localisation dans l'image, ainsi que la détermination de leur position à l'aide d'une **boîte de délimitation**. En ce sens, la détection d'objets sert de pont entre la classification d'images et des tâches plus avancées, telles que la **segmentation d'images**, où l'objectif est d'attribuer un label à chaque pixel de l'image, identifiant ainsi l'objet auquel il appartient. Dans la Figure 2 nous avons plusieurs objets, tels que des voitures, des vélos, une personne, un camion et un feu de circulation, sont détectés et étiquetés avec des boîtes de délimitation de différentes couleurs

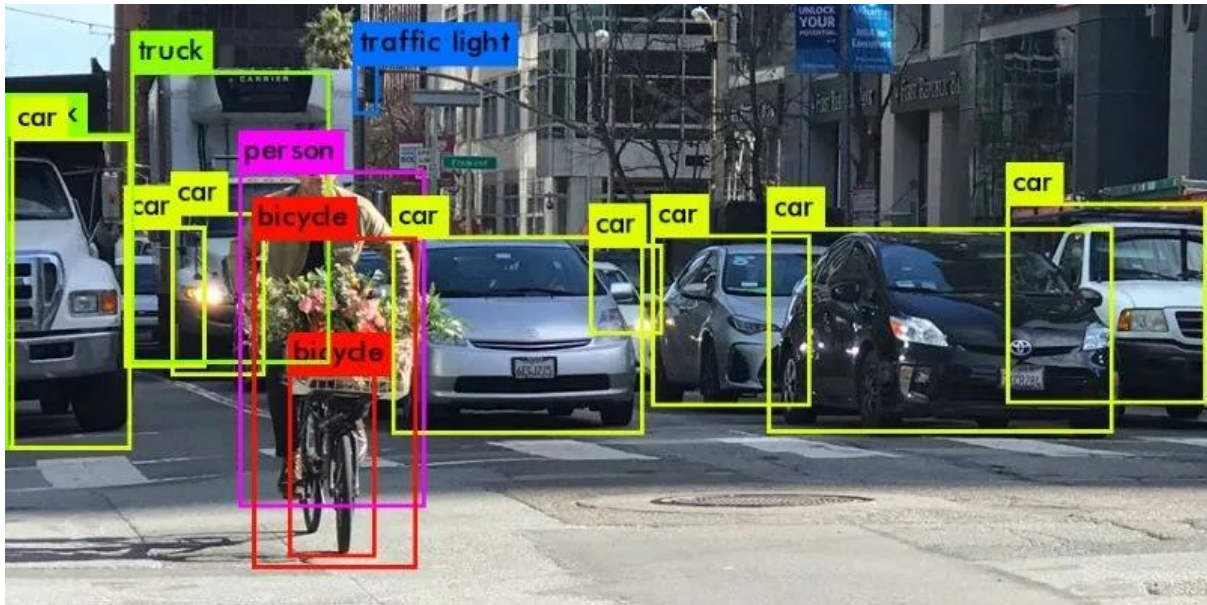


Figure 2: Exemple de détection d'objets en milieu urbain à l'aide de la vision par ordinateur [\[102\]](#)

Ce processus utilise des modèles d'apprentissage profond comme YOLO ou Faster R-CNN pour identifier et localiser les objets dans une scène urbaine. Chaque objet est classé avec une étiquette correspondant à sa catégorie. Ces technologies sont appliquées dans des domaines comme la surveillance et la conduite autonome.

Dans le contexte de notre mémoire, les objectifs de cette approche sont les suivants :

Détection des zones en îlot de chaleur : Nous utilisons le modèle YOLO pour localiser et identifier les zones en îlot de chaleur dans les images urbaines. Ce modèle est particulièrement adapté à ce type de détection, car il permet de repérer rapidement des objets ou des zones d'intérêt dans une seule passe grâce à son architecture d'un seul réseau convolutif.

Localisation précise des zones de chaleur : Une fois que les zones chaudes sont détectées, le modèle YOLO permet de déterminer leur localisation exacte sur l'image, facilitant ainsi l'analyse des zones les plus à risque. Cela permet de mieux comprendre l'étendue des îlots de chaleur et d'intervenir plus efficacement.

L'objectif final de cette approche est de pouvoir non seulement localiser les zones chaudes mais aussi d'évaluer l'ampleur du phénomène pour une analyse plus approfondie. Grâce à la rapidité de YOLO, cette approche est particulièrement efficace pour les grandes images où plusieurs zones d'intérêt doivent être identifiées et localisées simultanément.

3.2 Présentation des modèles d'apprentissages automatiques utilisés

Le ML constitue depuis plusieurs années la principale famille des méthodes de l'IA [4]. Il est conçu pour fournir aux systèmes informatiques la capacité d'apprendre à partir de données, de reconnaître des modèles et d'effectuer des tâches sans avoir à être explicitement programmé pour chaque situation spécifique. Dans cette recherche, nous nous concentrons sur les techniques d'apprentissage profond, un sous-domaine du ML particulièrement efficace pour traiter des données volumineuses et complexes telles que les images satellite.

Les modèles de ML adoptés dans cette étude ont été soigneusement choisis pour répondre aux besoins spécifiques des tâches liées à la détection des îlots de chaleur. Le Tableau 1 présente une correspondance entre trois tâches spécifiques d'analyse d'images satellites et les modèles de deep learning les plus adaptés à chacune.

Le Tableau 1 est structuré en trois colonnes : la première colonne décrit la tâche réalisée, la deuxième indique le modèle utilisé et la troisième justifie son choix en expliquant son rôle.

La première ligne concerne la classification des zones urbaines, qui est effectuée à l'aide du modèle ResNet. Ce réseau neuronal profond est utilisé pour classifier différents types de zones urbaines, qu'elles soient résidentielles, industrielles ou naturelles, en exploitant des images satellites haute résolution. Son rôle est donc d'assigner une catégorie précise à chaque région analysée en se basant sur les caractéristiques visuelles des images.

La deuxième ligne du tableau est dédiée à la segmentation des îlots de chaleur, une tâche assurée par U-Net. Ce modèle est conçu pour segmenter précisément les zones touchées par les îlots de chaleur en générant des masques à partir des données d'entraînement. Contrairement à la classification, qui attribue une catégorie globale à une image ou une portion d'image, la segmentation permet une détection pixel par pixel des zones concernées, offrant ainsi une vision plus détaillée et localisée du phénomène.

Enfin, la dernière ligne traite de la détection des zones à risque, réalisée avec YOLOv8. Ce réseau est spécialisé dans la détection d'objets en temps réel et permet d'identifier et de localiser avec précision les régions critiques affectées par les îlots de chaleur. En travaillant sur de larges images satellites, il est capable de fournir des indications précises sur l'emplacement des zones à surveiller.

Tableau 1: Synthèse des modèles de ML et leurs applications dans la détection des îlots de chaleur

Tâche	Model	Rôle et justification
Classification des zones urbaines	ResNet	Un réseau neuronal profond utilisé pour classifier les types de zones urbaines (résidentielles, industrielles, naturelles, etc.) à partir d'images satellite haute résolution.
Segmentation des îlots de chaleur	U-Net	Un modèle conçu pour segmenter précisément les zones affectées par les îlots de chaleur, en utilisant des masques générés à partir des données d'entraînement.
Détection des zones à risque	YOLOv8	Un réseau performant en temps réel pour détecter et localiser les régions critiques d'îlots de chaleur dans de larges images satellite.

Ce tableau met donc en évidence l'utilisation de trois approches complémentaires pour l'analyse des îlots de chaleur et des environnements urbains : la classification avec ResNet, la segmentation avec U-Net et la détection avec YOLOv8.

3.3 Les types d'apprentissages automatiques

3.3.1 Apprentissage non supervisé

L'apprentissage non supervisé est une approche de ML où le modèle est entraîné à partir de données non étiquetées, c'est-à-dire sans que chaque entrée soit associée à une sortie attendue. L'objectif principal de cette méthode est de découvrir des structures ou des relations cachées dans les données, sans intervention humaine pour fournir des étiquettes. En appliquant des techniques d'apprentissage non supervisé, le modèle tente d'identifier des motifs, des groupes ou des anomalies dans les données de manière autonome. Ce type d'apprentissage est

particulièrement utile lorsqu'il est difficile d'obtenir des données étiquetées ou lorsque l'on cherche à explorer un ensemble de données à grande échelle pour en extraire des informations pertinentes sans avoir de classes prédéfinies. Les algorithmes utilisés en apprentissage non supervisé incluent les k-means, les cartes auto-organisatrices (SOM), ou encore les réseaux de neurones génératifs adverses (GAN). Dans le contexte de l'analyse des ICU, l'apprentissage non supervisé pourrait être utilisé pour identifier des clusters ou des regroupements de zones urbaines présentant des caractéristiques thermiques similaires, sans avoir besoin de les préclasser au préalable. Cela permettrait d'explorer de nouvelles régions d'intérêt et d'obtenir des résultats potentiellement inattendus qui pourraient ne pas être apparus dans des analyses supervisées traditionnelles.

3.3.2 Apprentissage semi-supervisé

L'apprentissage semi-supervisé combine les avantages de l'apprentissage supervisé et non supervisé. Dans cette approche, une petite quantité de données étiquetées est utilisée en complément d'une grande quantité de données non étiquetées. Ce type d'apprentissage est souvent appliqué lorsque l'étiquetage des données est coûteux ou chronophage, mais qu'il est encore possible d'obtenir suffisamment d'exemples étiquetés pour guider le modèle. L'algorithme exploite les données non étiquetées pour mieux comprendre les structures sous-jacentes du problème, tout en s'appuyant sur les étiquettes disponibles pour affiner ses prédictions. Cela permet d'améliorer les performances du modèle par rapport à un apprentissage supervisé utilisant uniquement des données étiquetées, tout en réduisant le besoin de données annotées. Dans le cadre de ce mémoire, l'apprentissage semi-supervisé pourrait être appliqué pour la détection et la classification des îlots de chaleur urbains, où des images satellites ou thermiques non étiquetées peuvent être utilisées pour compléter un ensemble de données étiquetées plus réduit. Par exemple, une petite quantité d'images déjà annotées avec des zones thermiques élevées pourrait être utilisée pour guider l'apprentissage du modèle, tandis que les images non étiquetées pourraient être exploitées pour mieux comprendre et affiner la segmentation des ICU. Cette méthode permettrait ainsi d'augmenter la précision des modèles tout en limitant le besoin en données étiquetées.

3.3.3 Apprentissage supervisé

L'apprentissage supervisé est une approche de ML où un modèle est entraîné à partir d'un ensemble de données préalablement étiqueté. Chaque donnée d'entrée est associée à une sortie attendue, ce qui permet au modèle d'apprendre les relations sous-jacentes entre les

caractéristiques des données et les classes correspondantes. L'objectif principal de cette méthode est d'apprendre à généraliser à partir des exemples fournis afin de pouvoir prédire avec précision les résultats sur de nouvelles données non vues lors de l'entraînement. Ce processus repose sur plusieurs étapes essentielles, notamment la collecte et l'annotation d'un ensemble de données structuré, où chaque image ou unité d'information est associée à une catégorie spécifique. Par la suite, le modèle est entraîné en ajustant ses paramètres internes pour minimiser l'erreur entre ses prédictions et les étiquettes réelles, grâce à une fonction de coût et un algorithme d'optimisation tel que la descente de gradient.

Dans le cadre de ce mémoire, l'apprentissage supervisé a été utilisé pour la détection et la classification des ICU. D'une part, il a permis d'identifier les zones urbaines affectées par les ICU en exploitant des images satellites et thermiques. Le modèle a été entraîné à reconnaître les zones où la température est anormalement élevée, en classant les pixels des images selon différentes catégories thermiques, ce qui a permis de générer des cartes de chaleur détaillées. D'autre part, cette approche a été appliquée à la classification des types de zones urbaines. L'architecture ResNet a été utilisée pour distinguer les différentes zones urbaines en fonction de leur occupation du sol, telles que les zones résidentielles, industrielles, commerciales et naturelles. Grâce à sa profondeur et à sa capacité à extraire des caractéristiques complexes, ResNet a permis d'identifier avec une grande précision les variations spatiales et thermiques des ICU.

L'application de l'apprentissage supervisé dans cette étude a joué un rôle clé dans la structuration et l'analyse des données. En combinant classification et segmentation, cette approche a permis non seulement de cartographier les zones touchées par les ICU, mais aussi d'obtenir une meilleure compréhension des facteurs environnementaux qui contribuent à leur formation. Ces résultats offrent ainsi des pistes d'optimisation pour les futures stratégies de gestion urbaine visant à atténuer l'impact des îlots de chaleur.

3.3.4 Le Modèle ResNet

Le ResNet (Residual Network) est une architecture de CNN introduite en 2015 par He, Zhang et al. dans le but de résoudre certains des défis liés à l'entraînement des réseaux neuronaux profonds, notamment la disparition du gradient. Cette architecture a révolutionné la vision par ordinateur en permettant l'entraînement de réseaux extrêmement profonds tout en maintenant une efficacité et une précision élevées.

Le ResNet s'est imposé comme une référence dans plusieurs tâches de vision par ordinateur, telles que la classification d'images, la détection d'objets et la segmentation sémantique¹¹. Il se distingue par l'introduction des blocs résiduels, une innovation majeure qui permet d'ajouter des connexions de saut direct entre les couches du réseau. Ces connexions facilitent le flux du gradient pendant l'entraînement, évitant ainsi l'explosion ou la disparition du gradient, un problème courant dans les réseaux très profonds. Cette approche a permis d'entraîner des réseaux atteignant jusqu'à 152 couches, surpassant les architectures précédentes en termes de performances.

En 2015, ResNet a marqué un tournant dans le domaine de la vision par ordinateur en remportant le ImageNet Large Scale Visual Recognition Challenge (ILSVRC) (Fan, Ma et al., 2021), un concours prestigieux de ML. Il a surpassé des modèles emblématiques tels qu'AlexNet, VGG19 et GoogleNet, en obtenant un taux d'erreur significativement plus faible, démontrant ainsi l'efficacité des blocs résiduels dans l'amélioration des performances des réseaux profonds.

Grâce à sa robustesse et à sa capacité à généraliser efficacement sur divers ensembles de données, ResNet est devenu un modèle incontournable dans la recherche et les applications industrielles, notamment en imagerie médicale, en reconnaissance faciale et en RS.

3.3.4.1 Architecture

La Figure 3 présente l'architecture ResNet50 qui est une déclinaison avancée des réseaux résiduels (ResNet), caractérisée par sa profondeur de 50 couches.

¹¹ Bastien MAURICE « Différences entre classification d'image, localisation/détection & segmentation d'objets - Deeply Learning », 2 décembre 2020 Disponible sur: https://deeplylearning.fr/cours-theoriques-deep-learning/differences-entre-classification-image-localisation-detection-segmentation-dobjets/#google_vignette

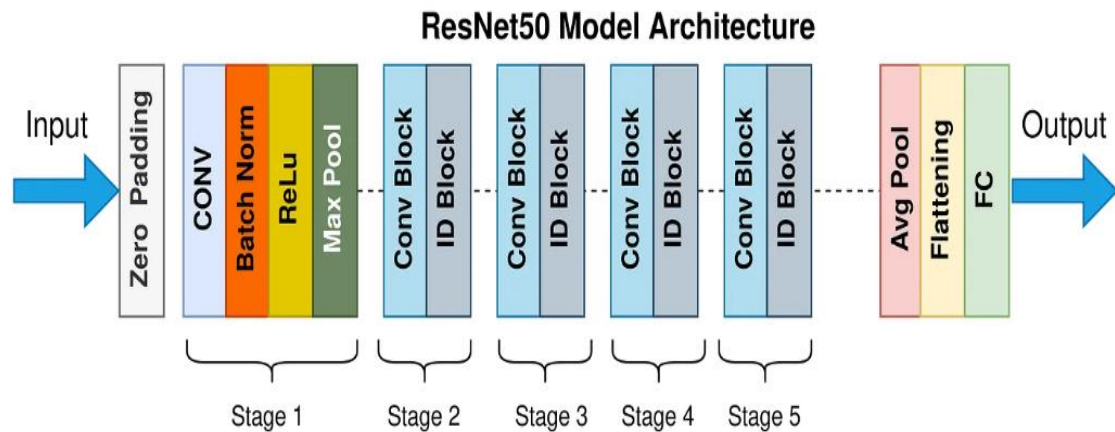


Figure 3: Architecture Restnet50[\[100\]](#)

Conçue pour surmonter les limitations des réseaux neuronaux profonds, elle repose sur l'innovation des blocs résiduels, qui introduisent des connexions de saut (skip connections). Ces connexions permettent de contourner certaines couches du réseau, facilitant ainsi la propagation des gradients et prévenant les problèmes de disparition ou d'explosion des gradients rencontrés lors de l'apprentissage.

L'image d'entrée passe d'abord par une convolution initiale de grande taille (7x7), suivie d'une opération de pooling. Cette étape permet de réduire la taille de l'image tout en préservant les informations essentielles nécessaires à l'extraction des caractéristiques. Ensuite, l'architecture de ResNet50 est structurée en quatre grandes étapes, chacune composée de plusieurs blocs résiduels. Chaque bloc résiduel incorpore plusieurs couches de convolution, utilisant des filtres de petite taille (3x3 et 1x1), accompagnées de techniques de normalisation par lots (Batch Normalization) pour stabiliser l'apprentissage. De plus, une fonction d'activation ReLU est appliquée après chaque convolution afin d'introduire la non-linéarité nécessaire à la modélisation des relations complexes dans les données.

L'une des innovations clés de ResNet50 réside dans ses connexions résiduelles, qui facilitent la propagation efficace des gradients à travers les couches profondes du réseau. Contrairement aux architectures classiques où les informations peuvent se perdre en raison des transformations successives, les connexions de saut permettent au réseau d'apprendre des résidus plutôt que des transformations directes, ce qui accélère l'apprentissage et améliore la stabilité de l'entraînement.

Après avoir traversé l'ensemble des blocs résiduels, les caractéristiques extraites sont aplaties et envoyées vers une couche entièrement connectée. Enfin, une fonction d'activation softmax est appliquée pour générer des probabilités associées aux différentes classes, permettant ainsi la prédiction finale.

L'architecture ResNet50 présente plusieurs avantages notables. Grâce à sa grande profondeur, elle est capable de capturer des caractéristiques complexes dans les images, ce qui en fait un choix performant pour des tâches de classification et de segmentation. De plus, sa stabilité d'apprentissage due aux connexions résiduelles permet d'entraîner efficacement des réseaux très profonds, sans souffrir des problèmes habituels de disparition du gradient. Enfin, sa polyvalence en fait un backbone privilégié pour diverses architectures avancées, telles que U-Net ou YOLO, où il est utilisé pour des applications allant de la classification d'images à la détection et segmentation d'objets.

3.3.5 Le Modèle U-Net

Le modèle U-Net, proposé en 2015 par Ronneberger et al., constitue une avancée majeure dans les architectures de réseaux neuronaux convolutionnels. Spécifiquement conçu pour les tâches de segmentation d'images, il a été initialement appliqué à des données biomédicales, mais s'est depuis imposé dans des domaines variés, notamment pour la détection des ilots de chaleur grâce à sa capacité à segmenter des zones complexes avec une grande précision. Ce modèle est particulièrement adapté aux données biomédicales, mais il a depuis été largement utilisé pour d'autres applications, telles que la segmentation des ilots de chaleur.

Sur la Figure 4 on a la présentation de l'architecture U-Net qui repose sur une structure en U, composée de deux parties principales : une partie contractante (encodeur) et une partie expansive (décodeur).

3.3.5.1 Architecture

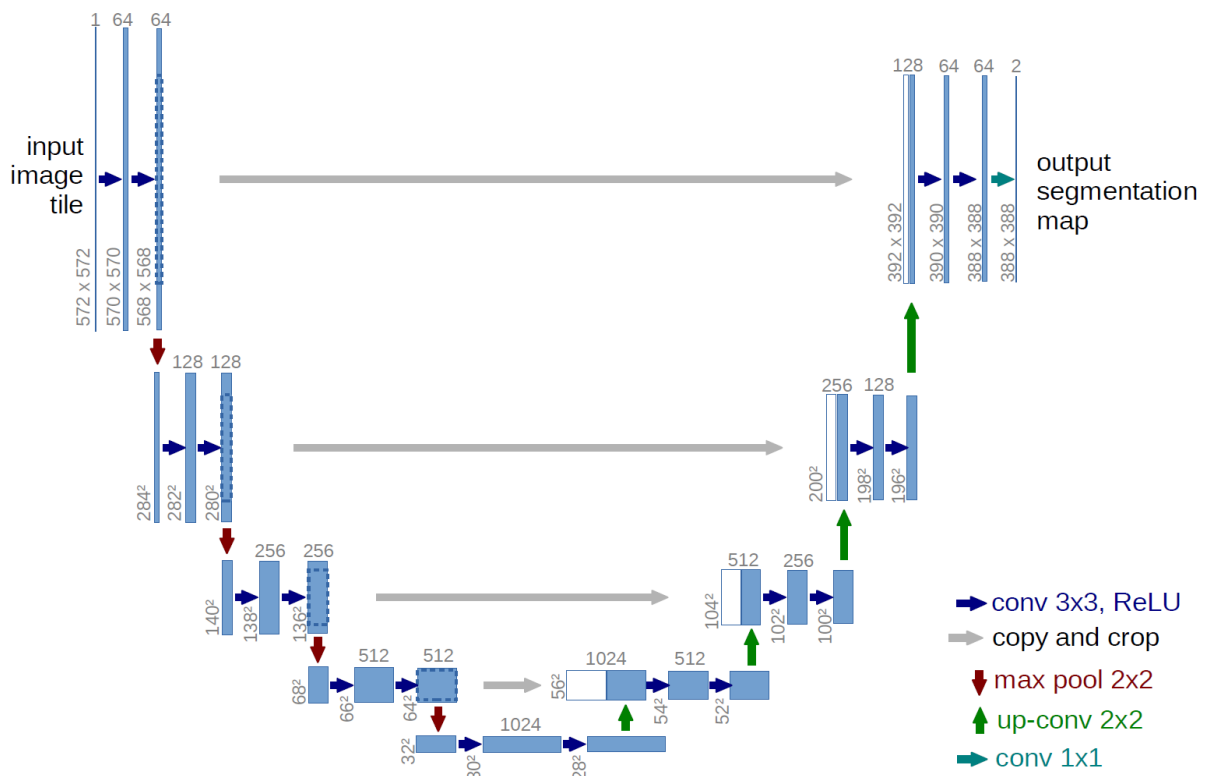


Figure 4: architecture de U-net [6]

L'architecture présentée à la Figure 4 permet d'extraire efficacement des caractéristiques complexes tout en préservant la précision spatiale, ce qui en fait un modèle particulièrement performant pour la segmentation d'images.

La partie contractante, semblable à un réseau convolutif classique, est responsable de l'extraction progressive des caractéristiques de l'image. Elle est constituée de plusieurs blocs convolutifs comprenant des couches de convolution 3×3, suivies d'une fonction d'activation ReLU, et des opérations de max pooling 2×2 qui réduisent la dimension spatiale de l'image tout en augmentant le nombre de filtres. À mesure que l'image traverse ces couches, elle devient plus compacte et ses caractéristiques plus abstraites. Cependant, cette compression entraîne une perte d'informations spatiales essentielles à la segmentation précise.

Au cœur de l'architecture se trouve le goulot d'étranglement, une étape où la résolution spatiale est réduite au minimum tandis que le nombre de canaux de caractéristiques est maximisé. Ce goulot d'étranglement joue un rôle essentiel dans la capture des motifs les plus pertinents de l'image tout en éliminant les détails non essentiels. Il constitue la transition entre la compression des données réalisée par l'encodeur et la reconstruction progressive effectuée par le décodeur.

La partie expansive a pour objectif de restaurer la résolution originale de l'image en appliquant des couches de convolution transposée (*up-convolution*). Ces couches augmentent progressivement la taille des images tout en affinant les informations de segmentation. L'une des innovations clés de U-Net réside dans ses connexions de saut (*skip connections*), qui relient directement les couches correspondantes de l'encodeur et du décodeur. Ces connexions permettent de réinjecter les informations spatiales perdues lors de la compression, garantissant ainsi une segmentation plus précise et détaillée. Une fois la reconstruction terminée, une couche finale de convolution 1×1 est appliquée pour produire une carte de segmentation, où chaque pixel est assigné à une classe spécifique, comme une zone d'îlot de chaleur ou une zone normale.

L'architecture U-Net présente plusieurs avantages qui en font une référence dans le domaine de la segmentation d'images. Sa capacité à préserver la précision spatiale grâce aux connexions de saut lui permet d'identifier des contours fins et de segmenter avec une grande fidélité des objets complexes. Son efficacité computationnelle en fait un choix pertinent pour l'analyse de grandes images, comme celles issues de la RS ou de la cartographie urbaine. De plus, son adaptabilité lui permet d'être appliqué à divers types de données, allant des images médicales aux images satellites utilisées pour l'étude des îlots de chaleur urbains.

Grâce à cette architecture optimisée, U-Net est largement utilisé pour détecter et cartographier les îlots de chaleur en analysant les variations thermiques des surfaces urbaines et en distinguant les zones les plus exposées à la chaleur. Il constitue ainsi un outil puissant pour l'analyse environnementale et l'aide à la planification urbaine.

3.3.6 Le Modèle YOLOv8

YOLOv8 est une architecture avancée spécifiquement optimisée pour la détection des zones à risque, comme les îlots de chaleur, dans des images satellite de haute résolution ¹². Elle excelle dans l'analyse en temps réel grâce à sa capacité à traiter les données d'une manière rapide et efficace, adaptée aux besoins de notre étude. Conçue pour localiser rapidement et précisément des objets dans une image, cette architecture est particulièrement utile pour identifier des zones

¹² Arwaabouattia « Car Object Detection Using Ultralytics YOLOv8 ». Publié le 29 décembre 2024. [En ligne]. Disponible sur: <https://app.readytensor.ai/publications/car-object-detection-using-ultralytics-yolov8-hiZGDyIXmIIX>

à risque et des anomalies. Dans la Figure 5 nous avons la présentation de l'architecture de yolov8 simplifiée.

3.3.6.1 Architecture

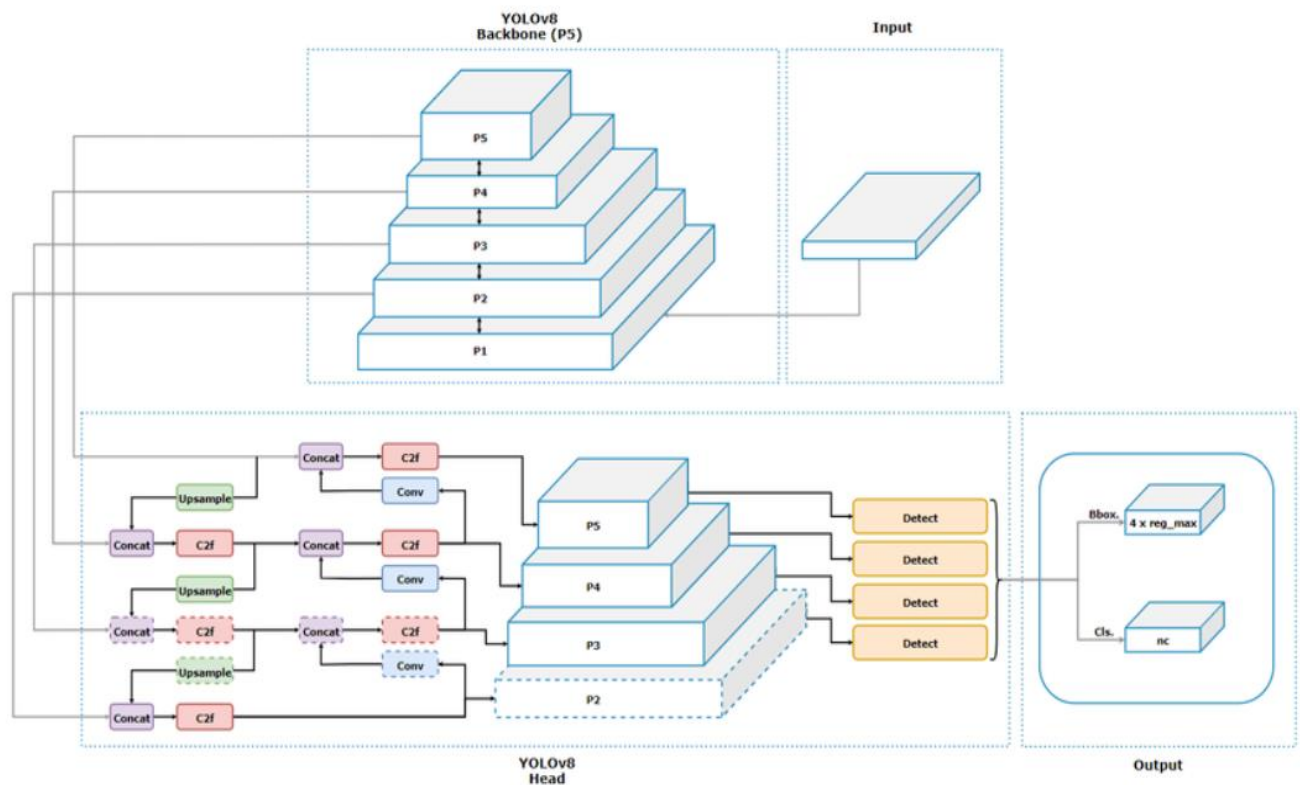


Figure 5: architecture YOLOV8 simplifiée [103]

YOLOv8 conserve la structure fondamentale des modèles YOLO tout en intégrant des améliorations significatives pour optimiser la détection d'objets. Son architecture repose sur plusieurs modules interconnectés, chacun jouant un rôle clé dans l'extraction des caractéristiques et la prédiction des objets détectés.

L'image d'entrée est d'abord normalisée et redimensionnée afin de s'adapter aux dimensions requises par le modèle. Cette étape assure l'uniformisation des données et optimise leur passage à travers les couches convolutives qui suivent. Le modèle est capable de traiter aussi bien des images RGB (Red, Green, Blue) que des images multispectrales, selon les besoins spécifiques de l'application.

Le *backbone*, basé sur une version améliorée du CSPNet (*Cross Stage Partial Network*), joue un rôle central dans l'extraction des caractéristiques. Il divise les caractéristiques extraites en deux flux parallèles, ce qui permet une meilleure fusion des informations et une réduction du

nombre de calculs nécessaires. Grâce à cette architecture optimisée, le modèle améliore l'apprentissage des représentations visuelles, facilitant ainsi la détection des objets dans des environnements complexes.

Le neck, qui constitue l'interface entre le *backbone* et la tête de détection, est composé de deux modules essentiels : le FPN (*Feature Pyramid Network*) et le PAN (*Path Aggregation Network*). Le FPN permet d'extraire des caractéristiques à différentes échelles, rendant possible la détection d'objets de tailles variées au sein d'une même image. Le PAN, quant à lui, améliore la propagation des caractéristiques en optimisant les connexions entre les couches supérieures et inférieures du réseau. Cette combinaison assure une gestion efficace des informations issues des différentes couches du réseau, renforçant ainsi la capacité du modèle à détecter aussi bien les petits objets que ceux présentant des variations de forme ou de texture.

Le module de détection (*head*) constitue l'élément clé de YOLOv8 pour prédire les objets détectés. Il génère les boîtes englobantes (*bounding boxes*), attribue un score de confiance à chaque détection et assigne une classe prédictive aux objets identifiés. Une des améliorations majeures de YOLOv8 réside dans l'affinement des prédictions grâce à un mécanisme de correction permettant de minimiser les erreurs et d'accroître la précision globale du modèle. Ce module assure ainsi une détection plus fiable en réduisant le nombre de fausses détections et en améliorant la localisation des objets.

La sortie finale du modèle comprend plusieurs éléments essentiels pour l'interprétation des résultats. Les boîtes englobantes permettent de localiser précisément les objets détectés, tandis que les scores de confiance indiquent la probabilité que l'objet appartienne à une classe spécifique. Enfin, les classes prédictives associent chaque détection à une catégorie identifiable, permettant ainsi une interprétation claire des résultats.

Parmi les innovations notables de YOLOv8, on retrouve une optimisation avancée des connexions résiduelles, qui améliore la gestion des informations tout au long du réseau et renforce la capacité du modèle à détecter des objets de petite taille ou situés dans des zones densément peuplées. Le module de tête de détection affiné permet d'améliorer la localisation précise des objets en minimisant les erreurs liées aux prédictions multiples. Enfin, YOLOv8 se distingue par son efficacité computationnelle accrue, qui lui permet de conserver la rapidité

caractéristique des modèles YOLO¹³ tout en atteignant une précision optimisée. Cette amélioration le rend particulièrement adapté aux applications en temps réel, nécessitant une exécution rapide et fiable, comme la surveillance urbaine, la reconnaissance d'objets dans des images satellites et la détection des îlots de chaleur.

En conclusion, l'utilisation combinée de ces modèles, tels que ResNet pour la classification, U-Net pour la segmentation, et YOLOv8 pour la détection, a permis d'atteindre une précision élevée dans l'identification et la localisation des îlots de chaleur. Cette synergie a fourni des résultats adaptés à l'analyse spatiale à grande échelle, renforçant l'efficacité de notre démarche méthodologique. Le choix de ces modèles repose sur leur capacité à traiter de grandes quantités de données tout en offrant des résultats interprétables et fiables. Le chapitre suivant présentera les résultats obtenus et leur analyse.

3.4 Analyse de données

Dans le cadre de cette étude sur la détection des îlots de chaleur, une collection de données d'images satellites a été obtenue à partir de sources publiques disponibles sur Kaggle, une plateforme populaire dédiée au partage de jeux de données et à l'expérimentation de modèles de ML. Kaggle propose une large variété de jeux de données accessibles librement, ainsi que des notebooks interactifs permettant d'exécuter du code directement dans un environnement cloud.

Pour accéder aux données, il a été nécessaire de créer un compte Kaggle. Une fois inscrit, il est possible de parcourir les différentes catégories de jeux de données et de télécharger ceux qui sont pertinents pour une étude donnée. Dans notre cas, nous avons sélectionné un jeu de données spécifique contenant 38 images haute résolution capturées par des satellites, associées à des masques annotés représentant les zones d'intérêt, comme les surfaces urbaines, les espaces verts et les plans d'eau.

¹³ Ultralytics, « YOLO Mesures de performance ». Consulté le: 9 janvier 2025. [En ligne]. Disponible sur: <https://docs.ultralytics.com/fr/guides/yolo-performance-metrics>

Après avoir téléchargé les données, elles ont été chargées dans notre environnement de travail en utilisant des bibliothèques Python telles que Pandas¹⁴, NumPy et OpenCV pour le traitement des images. Afin de faciliter la manipulation des fichiers et leur structuration, les images ont été organisées dans des répertoires distincts : un répertoire pour les images brutes et un autre pour les masques correspondants.

Afin d'améliorer la diversité et la représentativité des données, une étape d'augmentation a été appliquée. Cette augmentation a permis de générer un total de 35 292 images à partir des 38 images d'origine, en appliquant diverses transformations, telles que les rotations, les flips et les variations de luminosité. Ces techniques d'augmentation ont été essentielles pour enrichir le jeu de données, augmenter sa variabilité et améliorer la robustesse des modèles d'apprentissage.

Le Tableau 2 présente la répartition des données initialement et après augmentation.

Tableau 2: répartition des données

Type d'image	Nombre initial	Nombre après augmentation
Images satellites	38	35 292
Masques d'annotation	38	35292

Le traitement de nos données a été divisé en trois étapes principales pour répondre aux objectifs spécifiques de chaque modèle utilisé dans notre méthodologie.

3.4.1 Préparation des données pour la classification

Dans la première étape, l'objectif principal est d'identifier les différentes zones urbaines et de les classer selon leur potentiel de formation d'îlots de chaleur [52].

¹⁴ W. McKinney, pandas: A foundational library for data analysis in Python, 2008. Disponible sur : [https://en.wikipedia.org/wiki/Pandas_\(software\)](https://en.wikipedia.org/wiki/Pandas_(software))

Les images satellites ont été chargées dans leur format brut depuis les répertoires sources, puis étiquetées selon la catégorie de surface qu'elles représentent. Les catégories incluent les zones urbaines denses, les espaces verts, les plans d'eau et les zones mixtes.

Après l'étiquetage, les images ont été redimensionnées à une taille standard de 256x256 pixels afin de garantir une entrée uniforme au modèle ResNet. Cette étape est cruciale pour assurer la cohérence des dimensions des images lors de l'entraînement.

Une normalisation des pixels a ensuite été effectuée, avec des valeurs ramenées dans l'intervalle [0, 1]. Cette opération permet d'améliorer la convergence du modèle lors de la phase d'apprentissage.

Enfin, les données préparées ont été divisées en deux ensembles distincts : un ensemble d'entraînement représentant 80 % des données totales, et un ensemble de validation représentant les 20 % restants. Cette division est essentielle pour évaluer les performances du modèle sur des données non vues pendant l'entraînement.

Afin de mieux comprendre la répartition des données utilisées dans cette étude, le Tableau 3 présente la proportion de répartition des données en fonction des différentes catégories. Ce tableau offre une vue d'ensemble détaillée des données disponibles pour l'entraînement et l'évaluation du modèle, permettant ainsi d'analyser la distribution des échantillons entre les différentes classes. Cette répartition est essentielle pour garantir une représentation équilibrée des données et pour évaluer la performance du modèle de manière optimale.

Tableau 3: Proportion de répartition des données

Ensemble	Proportion	Nombre d'images
Entrainement	80%	28234
Validation	20%	7058

3.4.2 Préparation des données pour la segmentation

La deuxième étape concerne la segmentation des zones à risque. L'objectif ici est de délimiter précisément les surfaces présentant des températures élevées dans les images satellites à l'aide du modèle U-Net.

Chaque image satellite a été associée à un masque binaire, où chaque pixel du masque indique la présence (1) ou l'absence (0) d'un îlot de chaleur. Cette étape d'association est fondamentale pour entraîner le modèle à reconnaître les zones pertinentes.

Étant donné la grande taille des images d'origine (512x512 pixels), celles-ci ont été découpées en blocs plus petits de 128x128 pixels. Ce découpage permet d'augmenter le nombre d'échantillons et de faciliter l'apprentissage du modèle, tout en réduisant les besoins en mémoire.

Pour améliorer la robustesse du modèle et éviter les problèmes de surapprentissage, une augmentation des données a été réalisée. Les transformations appliquées incluent des rotations aléatoires jusqu'à 90 degrés, des flips horizontaux et verticaux, des zooms, des translations ainsi que des variations de luminosité et de contraste. Ces augmentations permettent de générer de nouvelles images à partir des données d'origine, enrichissant ainsi le jeu de données et améliorant la capacité du modèle à généraliser.

Le Tableau 4 présente les techniques de transformations appliquées pour augmenter les données avec leurs descriptions.

Tableau 4: Tableau de l'augmentation des données

Transformation appliquée	Description
Rotation	Rotations aléatoires des images jusqu'à 90 degrés
Flip	Inversions horizontales et verticales
Zoom	Zooms aléatoires sur des portions de l'image
Translation	Déplacement des images sur différentes directions
Variation de luminosité	Modifications aléatoires de la luminosité et du contraste des images

Comme pour la classification, les données augmentées ont été divisées en ensembles d'entraînement et de validation selon les mêmes proportions de 80 % et 20 %

3.4.3 Préparation des données pour la détection

La troisième étape consiste à préparer les données pour la détection des îlots de chaleur à l'aide du modèle YOLOv8. Contrairement à la segmentation, cette étape vise à localiser les zones d'intérêt en dessinant des boîtes englobantes (*bounding boxes*) autour des régions concernées¹⁵.

Les masques d'annotation ont d'abord été convertis en boîtes englobantes. Chaque boîte délimite une région présentant une température élevée et est associée à une classe spécifique : 1 pour les îlots de chaleur et 0 pour les zones normales.

Les coordonnées de chaque boîte ont été extraites et formatées selon les exigences de YOLO. Ce format inclut les informations suivantes : les coordonnées du centre de la boîte, la largeur et la hauteur de celle-ci, ainsi que la classe associée.

Une fois les boîtes générées et étiquetées, les données ont été divisées en ensembles d'entraînement et de validation selon les proportions habituelles de 80 % et 20 %. Cette séparation garantit que les performances du modèle seront évaluées sur des données indépendantes de celles utilisées pour l'apprentissage.

Enfin, des métriques d'évaluation spécifiques, telles que l'IoU¹⁶ et la précision, seront utilisées pour mesurer la qualité des détections effectuées par le modèle YOLOv8. Ces métriques sont essentielles pour quantifier la capacité du modèle à localiser avec précision les îlots de chaleur.

Il est essentiel de souligner que ces trois étapes distinctes de préparation des données ont permis de constituer un jeu de données solide et adapté aux besoins spécifiques de chaque modèle utilisé dans cette étude. La classification a permis de différencier les types de zones urbaines selon leur potentiel de formation d'îlots de chaleur, la segmentation a aidé à délimiter précisément les zones présentant des températures élevées, et la détection a fourni une localisation claire des régions à risque. Ces étapes méthodologiques rigoureuses offrent une

¹⁵ Vision Platform, « YOLOv8 : Détection d'objets de pointe en vision par ordinateur (computer vision) », visionplatform. Publié 18 Mai 2024. Disponible sur: <https://www.visionplatform.ai/fr/yolov8-detection-dobjets-de-pointe-en-vision-par-ordinateur-computer-vision/>

¹⁶ G. Boesch, « What is Intersection over Union (IoU)? », viso.ai. Consulté le: 9 janvier 2025. [En ligne]. Disponible sur: <https://viso.ai/computer-vision/intersection-over-union-iou/>

base fiable pour la suite de l'étude, notamment dans l'évaluation des performances des modèles et l'interprétation des résultats obtenus, qui seront détaillées dans le chapitre suivant.

3.5 Autres Outils

Dans le cadre de cette mémoire, plusieurs outils ont été mobilisés pour l'annotation des images, le stockage des données et l'exécution des modèles d'apprentissage automatique. Les principales plateformes employées incluent Google Colaboratory, Google Drive, le CVAT, Apeer et le langage Python avec ses bibliothèques spécialisées.

3.5.1 Google Colaboratory

Développé par Google Research, Google Colaboratory (Colab) est un environnement de développement interactif basé sur le cloud, conçu pour exécuter du code Python au sein de notebooks Jupyter hébergés à distance. Il s'est rapidement imposé comme une solution privilégiée dans la recherche en IA, notamment en raison de la mise à disposition gratuite de GPU et TPU, ressources essentielles pour l'apprentissage profond [87]. Colab permet aux chercheurs de contourner les limitations matérielles locales, tout en bénéficiant d'un accès à un environnement préconfiguré avec des bibliothèques comme TensorFlow, Keras ou OpenCV.

L'environnement supporte également l'exécution collaborative en temps réel, ce qui facilite les projets d'équipe. Enfin, l'intégration native avec Google Drive permet un accès et un stockage directs des données et des résultats [88].

3.5.2 Google Drive

Google Drive est un service de stockage en nuage proposé par Google depuis 2012. Dans le cadre de cette étude, il a été utilisé comme support principal de stockage des images satellites, des masques annotés et des résultats de modélisation. Son intégration avec Colab a permis de charger, sauvegarder et partager les fichiers directement depuis les notebooks, optimisant le flux de travail collaboratif et assurant une traçabilité complète des expériences [89].

Dans ce mémoire, Google Drive a été utilisé comme principal espace de stockage, garantissant une accessibilité immédiate aux données sans contraintes liées aux capacités matérielles locales. Cette infrastructure cloud a permis de stocker et d'organiser efficacement les images satellites et leurs masques annotés dans des répertoires distincts, assurant ainsi une meilleure gestion des fichiers à chaque étape du prétraitement et de l'entraînement des modèles.

Un autre avantage clé de Google Drive réside dans sa capacité à sauvegarder automatiquement les checkpoints des modèles et les résultats finaux, évitant ainsi toute perte de données en cas d'interruption de l'exécution sur Google Colab. De plus, la synchronisation fluide avec Google Drive API a facilité le chargement et le partage des fichiers, optimisant le workflow collaboratif et la traçabilité des expériences.

En combinant stockage cloud évolutif, sécurisation des données et intégration fluide avec l'environnement de développement, Google Drive s'est imposé comme une solution incontournable pour la gestion des données et des modèles au sein de ce projet.

3.5.3 CVAT

Le CVAT est un outil open-source développé par Intel en 2018, spécifiquement conçu pour l'annotation d'images dans le cadre de tâches de vision par ordinateur [90]. Il prend en charge une grande variété de formats d'annotation : boîtes englobantes, polygones, masques, etc. Dans ce mémoire, CVAT a été utilisé pour la création manuelle de masques précis, nécessaires à l'entraînement du modèle U-Net. Son interface permet une annotation fine, ce qui est crucial pour les modèles de segmentation où la qualité des données d'entrée a un impact direct sur la performance des prédictions [91].

Dans le cadre de ce mémoire, CVAT a joué un rôle clé en permettant la création manuelle d'annotations détaillées sur des images satellites. Ces annotations ont été utilisées pour délimiter les zones urbaines, les espaces verts et les plans d'eau, générant ainsi des masques d'entraînement optimisés pour le modèle U-Net.

L'annotation précise des images représente une étape critique dans le développement des modèles de segmentation, car elle influence directement la qualité des données d'entrée et, par conséquent, la précision des résultats obtenus. Grâce à CVAT, il a été possible d'assurer une cohérence et une haute-fidélité des annotations, renforçant ainsi la robustesse du modèle de segmentation dans l'identification des différentes zones du paysage urbain analysé.

3.5.4 Apeer

Apeer, initialement développé par ZEISS, est une plateforme cloud destinée à l'analyse et à l'annotation d'images scientifiques. Elle combine des outils d'IA à des interfaces de traitement collaboratif, facilitant ainsi l'annotation semi-automatique de jeux de données complexes. Elle est particulièrement utile pour les images nécessitant une validation humaine dans des cas

ambigus [92]. Dans ce projet, Apeer a été utilisée en complément de CVAT pour accélérer l'annotation et améliorer la précision des masques, grâce à une approche hybride alliant automatisation et correction manuelle.

Dans le cadre de notre mémoire, Apeer a été utilisé en complément de CVAT afin d'annoter des images présentant des caractéristiques complexes ou ambiguës, nécessitant une annotation plus précise. Grâce à ses capacités d'annotation semi-automatique, la plateforme a permis d'accélérer considérablement le processus d'annotation, réduisant le temps nécessaire à la création du jeu de données tout en améliorant la précision des masques segmentés destinés à l'entraînement du modèle de deep learning.

L'intégration de l'IA dans Apeer a facilité l'automatisation partielle de certaines tâches répétitives, tout en laissant aux utilisateurs la possibilité d'ajuster manuellement les annotations pour garantir une qualité optimale des données labellisées. Cette approche hybride entre annotation automatique et validation humaine a contribué à une meilleure efficacité dans la constitution du dataset, optimisant ainsi les performances des modèles entraînés.

3.5.5 Python

Le langage Python s'est imposé comme un outil incontournable dans les domaines de la science des données, de l'intelligence artificielle et de l'apprentissage automatique, en raison de sa syntaxe claire, de sa lisibilité et de la richesse de son écosystème de bibliothèques. Créé en 1991 par Guido van Rossum, Python est aujourd'hui largement utilisé pour le développement de modèles d'apprentissage automatique et de l'apprentissage profond, grâce à sa capacité à intégrer facilement des bibliothèques spécialisées et à sa communauté active [93].

Dans le cadre de cette étude, plusieurs bibliothèques essentielles de l'écosystème Python ont été mobilisées pour concevoir et optimiser les modèles. TensorFlow, développé par Google Brain en 2015, est une bibliothèque open-source permettant la construction et l'entraînement de modèles de deep learning à grande échelle. Elle offre une flexibilité et une performance accrues, notamment grâce à son architecture basée sur des graphes de calculs distribués [94]. Keras, introduite la même année par François Chollet, est une interface haut niveau pour TensorFlow, facilitant la création rapide de modèles de réseaux neuronaux grâce à une API intuitive et modulaire [95].

Pour les opérations numériques avancées, NumPy a été utilisé. Cette bibliothèque, fondée par Travis Oliphant en 2006, fournit des structures de données performantes, telles que les tableaux multidimensionnels, et des fonctions mathématiques efficaces, essentielles pour le traitement et l'analyse des données scientifiques [96].

La gestion et l'analyse des ensembles de données tabulaires ont été assurées par Pandas, créée par Wes McKinney en 2008. Pandas offre des structures de données flexibles, comme les DataFrames, et une panoplie d'outils pour le nettoyage, la transformation et l'analyse des données, facilitant ainsi la manipulation de grands volumes de données [97].

Enfin, pour la visualisation des résultats et l'illustration des performances des modèles, les bibliothèques Matplotlib et Seaborn ont été employées. Matplotlib, développée par John D. Hunter en 2003, est une bibliothèque de traçage 2D puissante, permettant la création de graphiques statiques, animés et interactifs [98]. Seaborn, introduite par Michael Waskom en 2012, est construite sur Matplotlib et offre une interface de haut niveau pour la création de visualisations statistiques attractives et informatives, facilitant l'exploration et la compréhension des données [99].

L'intégration de ces bibliothèques a permis de couvrir l'ensemble des étapes du pipeline de traitement, de la préparation des données jusqu'à l'interprétation des performances des modèles, garantissant ainsi une approche méthodique et rigoureuse tout au long de l'étude.

3.6 Conclusion

Ce chapitre a permis de détailler les étapes méthodologiques mises en œuvre pour concevoir un système efficace de détection des îlots de chaleur urbains à l'aide de techniques avancées d'IA et d'analyse d'images. En combinant la collecte de données géospatiales, le prétraitement rigoureux des images et l'entraînement de modèles de réseaux neuronaux adaptés, une base solide a été établie pour garantir des résultats fiables et exploitables.

Les approches adoptées ont été validées à travers des expérimentations sur des ensembles de données annotées manuellement, et les performances ont été évaluées selon des critères précis, tels que la précision, le rappel et le score F1. Ces étapes méthodologiques ont permis de sélectionner le modèle offrant les meilleurs résultats, tout en assurant une généralisation suffisante pour son application à de nouvelles données.

Cette méthodologie, rigoureuse et scalable, constitue une avancée notable dans le domaine de l'analyse géospatiale et de la gestion des ICU. Elle ouvre des perspectives intéressantes pour la planification urbaine durable et la réduction des impacts liés au réchauffement climatique dans les villes.

Dans le chapitre suivant, nous présenterons les résultats obtenus à partir de cette méthodologie, ainsi qu'une discussion approfondie sur leur pertinence, leurs implications pratiques et les éventuelles limitations rencontrées au cours de cette étude.

Chapitre 4 : Résultats et discussions

Introduction

Ce chapitre se concentre sur la présentation et l'analyse des résultats obtenus à partir de l'application des modèles de ML détaillés dans le chapitre précédent. Ces modèles ont été développés et optimisés pour répondre aux objectifs spécifiques de cette étude, à savoir la classification des zones thermiques, la segmentation précise des ICU et la localisation géographique des zones les plus affectées par le phénomène. L'objectif principal de cette phase est d'évaluer l'efficacité des modèles utilisés et de tirer des conclusions sur la pertinence des approches méthodologiques adoptées.

Nous commencerons par exposer de manière détaillée les résultats issus de chaque tâche, en fournissant des métriques d'évaluation précises, telles que la précision, le rappel et le score F1. Ces résultats seront illustrés à l'aide de graphiques, de tableaux comparatifs et de cartes thermiques générées par nos modèles. Chaque sous-section mettra en lumière la performance individuelle des modèles appliqués et les défis rencontrés lors de leur mise en œuvre.

Par la suite, une analyse critique des résultats sera réalisée. Cette étape consistera à interpréter les résultats obtenus en identifiant les facteurs ayant contribué aux performances observées, ainsi que les éventuelles limites des modèles. Nous chercherons à comprendre pourquoi certaines approches ont mieux fonctionné que d'autres et dans quels contextes spécifiques elles pourraient être améliorées. Cette réflexion permettra d'approfondir la compréhension des phénomènes thermiques urbains et d'évaluer la robustesse des modèles face aux variations des données d'entrée.

Enfin, une discussion sera menée sur les perspectives d'amélioration des modèles et les pistes de recherche futures. Cette discussion portera notamment sur l'optimisation des hyperparamètres, l'ajout de nouvelles sources de données, comme des données Météorologiques et topographiques, ainsi que l'intégration de techniques d'apprentissage avancées, telles que l'apprentissage semi-supervisé et auto-supervisé. Nous envisagerons également la possibilité d'étendre cette approche à d'autres villes ou régions présentant des caractéristiques climatiques et urbaines différentes, afin de tester la généralisation de la méthodologie proposée.

Ce chapitre vise ainsi non seulement à évaluer la qualité des résultats obtenus, mais également à offrir une analyse approfondie permettant d'identifier les axes d'amélioration possibles et d'ouvrir la voie à des recherches complémentaires. Grâce à cette démarche, nous espérons fournir des outils et des recommandations utiles aux urbanistes et décideurs pour atténuer les effets des îlots de chaleur urbains, tout en contribuant à la création de villes plus résilientes et durables face au changement climatique.

4.1 Détection et segmentation des îlots de chaleur

La détection et la segmentation des îlots de chaleur urbains représentent des étapes cruciales pour l'analyse spatiale et la planification urbaine dans le contexte des changements climatiques. Ces deux tâches visent à repérer les zones où les températures sont anormalement élevées par rapport aux alentours et à délimiter précisément les contours de ces zones d'îlots de chaleur.

Pour ce faire, nous avons utilisé des modèles de ML spécialisés dans la vision par ordinateur, notamment U-Net, U-Net+ResNet et YOLOv8. Ces modèles permettent d'exploiter des images satellitaires et thermiques afin de produire des cartes thermiques segmentées, où chaque pixel ou région est classée selon son niveau de chaleur.

4.1.1 Performance des modèles

L'évaluation des performances des modèles repose sur l'utilisation de plusieurs métriques essentielles, appliquées à un ensemble de test afin d'analyser leur efficacité et leur précision dans la détection des îlots de chaleur. La précision est une mesure qui indique le pourcentage de prédictions correctes parmi l'ensemble des prédictions effectuées. Une précision élevée signifie que le modèle produit peu de fausses détections, ce qui est crucial pour éviter les erreurs d'identification et assurer une fiabilité maximale dans l'interprétation des résultats. Toutefois, la précision seule ne permet pas d'avoir une vue complète des performances du modèle, car elle ne prend pas en compte les cas où des îlots de chaleur présents dans l'image ne sont pas détectés.

Pour pallier cette limite, le score F1 est utilisé comme un indicateur complémentaire. Il représente la moyenne harmonique entre la précision et le rappel, permettant d'évaluer la capacité du modèle à identifier correctement les îlots de chaleur tout en minimisant les erreurs d'omission et de fausses alarmes. Un score F1 élevé signifie que le modèle parvient à équilibrer

efficacement la précision et le rappel, garantissant ainsi une détection plus fiable des zones concernées.

Par ailleurs, l'IoU constitue une métrique clé pour évaluer la qualité des détections spatiales. Elle mesure le degré de superposition entre la zone prédite par le modèle et la zone réelle de l'îlot de chaleur en calculant le rapport entre l'intersection et l'union de ces deux régions. Un IoU élevé indique une forte correspondance entre la zone identifiée par l'algorithme et la zone de référence, signifiant que le modèle localise avec précision les contours et l'étendue des îlots de chaleur. À l'inverse, un faible IoU met en évidence un problème de mauvaise délimitation des zones, ce qui peut entraîner des erreurs d'interprétation dans les analyses environnementales et urbaines.

L'utilisation combinée de ces métriques permet d'obtenir une vision détaillée des performances des modèles en mettant en lumière leurs forces et leurs limites. Cette approche facilite également l'identification des axes d'amélioration, en orientant les ajustements à apporter aux algorithmes pour affiner la qualité des prédictions et optimiser la robustesse du système de détection des îlots de chaleur.

Le Tableau 5 ci-dessous présente les résultats des métriques obtenus pour chaque modèle

Tableau 5: résultats obtenus pour chaque model

Model	Précision	Score F1(%)	IoU(%)
U-Net	85,13	85,13	46,68
U-Net+ ResNet	--	89,46	63,40
YOLOv8	92,60	--	84,39

Interprétation détaillée des métriques utilisées

L'évaluation de la performance des modèles repose sur plusieurs indicateurs quantitatifs qui permettent de mesurer la fiabilité et l'efficacité des prédictions. Parmi ces métriques, la précision joue un rôle clé en indiquant la proportion de prédictions correctes parmi celles identifiées comme positives par le modèle. Elle permet ainsi de quantifier la capacité du modèle à détecter uniquement les véritables îlots de chaleur, en minimisant les fausses alertes. Une précision élevée traduit une grande fiabilité du modèle dans la classification des zones affectées, ce qui est essentiel pour éviter la détection erronée de zones normales comme étant des îlots de chaleur. Cependant, un modèle peut présenter une haute précision tout en manquant

certaines zones importantes si son rappel est faible, d'où la nécessité d'un équilibre entre ces métriques.

Pour compléter cette évaluation, le score F1 est utilisé comme une mesure plus globale de la performance du modèle. Il correspond à la moyenne harmonique entre la précision et le rappel, offrant ainsi une évaluation équilibrée entre la capacité du modèle à détecter les îlots de chaleur réels et sa propension à éviter les fausses alarmes. Un score F1 élevé indique que le modèle parvient efficacement à identifier les zones concernées sans négliger une partie significative des îlots de chaleur. Cette métrique est particulièrement pertinente dans le cadre de la segmentation d'images, où il est crucial d'assurer une détection complète des zones ciblées tout en réduisant les erreurs de classification.

Enfin, l'IoU constitue un indicateur essentiel pour évaluer la précision spatiale des prédictions du modèle. Cette métrique mesure le rapport entre l'intersection et l'union des zones prédites et des zones réelles, offrant ainsi un indicateur précis de la qualité du positionnement des détections. Un IoU élevé signifie que le modèle a réussi à capturer avec précision les contours des îlots de chaleur et à bien les différencier des autres structures urbaines. À l'inverse, un faible IoU peut indiquer que le modèle a mal délimité ces zones, ce qui peut entraîner des erreurs dans l'interprétation des résultats et fausser les analyses d'aménagement urbain. L'IoU est donc particulièrement adapté aux modèles de segmentation, où l'objectif est d'obtenir une correspondance exacte entre la zone détectée et la réalité observée.

L'interprétation conjointe de ces métriques permet d'obtenir une vision plus approfondie des performances du modèle. Elle met en évidence ses forces et ses faiblesses, facilitant ainsi les ajustements nécessaires pour améliorer la précision des détections et optimiser la robustesse du modèle. Un bon équilibre entre la précision, le score F1 et l'IoU garantit une analyse plus fiable des îlots de chaleur, contribuant ainsi à une meilleure compréhension des dynamiques thermiques urbaines et à la mise en place de stratégies adaptées pour leur atténuation.

L'évolution de la fonction perte pour le modèle U-Net est un indicateur clé permettant d'analyser la convergence de l'algorithme d'apprentissage tout au long du processus d'entraînement. La Figure 6 illustre cette évolution en montrant la diminution progressive de la perte au fil des itérations. Une baisse régulière de cette fonction indique que le modèle ajuste efficacement ses poids pour minimiser l'écart entre les prédictions et les valeurs réelles.

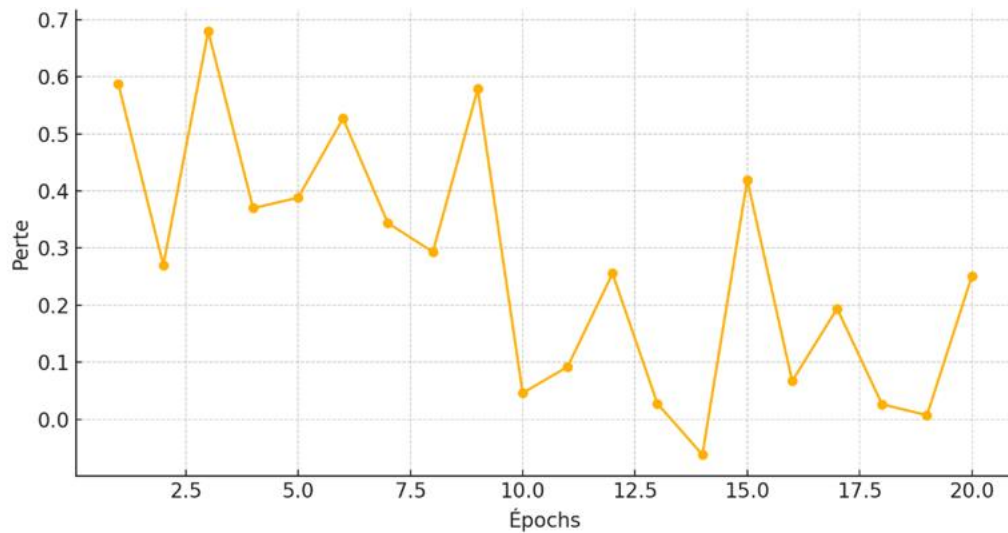


Figure 6: Évolution de la fonction perte - U-Net

Cette figure représente l'évolution de la fonction perte au cours de l'entraînement du modèle U-Net, un indicateur essentiel pour évaluer l'apprentissage du réseau de neurones. Dès les premières itérations, on observe une valeur initiale élevée de la fonction perte, ce qui est caractéristique des modèles d'apprentissage profond qui commencent leur phase d'optimisation. Cette perte élevée traduit le fait que les prédictions du modèle sont encore loin des valeurs réelles attendues, nécessitant des ajustements progressifs des poids du réseau.

Au fil des époques, la diminution progressive de la perte témoigne d'un apprentissage efficace, indiquant que le modèle affine ses prédictions et réduit l'écart entre les valeurs estimées et les valeurs cibles. Cette réduction constante de la perte est un signe de bonne convergence, montrant que l'algorithme optimise ses paramètres pour mieux représenter les structures présentes dans les données d'entrée. Toutefois, certaines fluctuations mineures peuvent être observées au cours de l'apprentissage. Ces variations sont normales et résultent des ajustements continus du modèle, qui tente de s'adapter aux différentes caractéristiques des images d'entraînement, notamment aux variations des températures captées dans les données satellitaires ou thermiques.

Vers la fin de l'entraînement, la courbe de perte tend à se stabiliser, ce qui suggère que le modèle a atteint un point d'optimisation où les gains en précision deviennent marginaux. Cette phase de stagnation signifie que le modèle a assimilé l'essentiel des relations présentes dans les données et que poursuivre l'apprentissage avec les mêmes paramètres ne permettrait plus

d'améliorer significativement ses performances. Une analyse approfondie de cette courbe permet également de détecter d'éventuels problèmes, tels qu'un sur-ajustement (overfitting) si la perte d'entraînement continue de baisser alors que la perte sur un jeu de validation augmente. Cette figure constitue donc un outil clé pour juger de la stabilité du modèle et ajuster, si nécessaire, les hyperparamètres afin d'obtenir une meilleure généralisation sur de nouvelles données.

L'évolution de la fonction d'exactitude du modèle U-Net est un paramètre fondamental pour évaluer l'amélioration des performances au cours de l'apprentissage. La Figure 7 illustre cette évolution en mettant en évidence l'augmentation progressive de l'exactitude au fil des itérations d'entraînement. Une courbe ascendante indique que le modèle affine progressivement sa capacité à classer correctement les pixels des images en fonction des catégories prédéfinies. Cette progression témoigne d'un ajustement efficace des poids du réseau, permettant une meilleure correspondance entre les prédictions du modèle et les valeurs attendues. L'analyse de cette courbe permet également de détecter d'éventuelles stagnations ou fluctuations anormales qui pourraient indiquer des limites dans la capacité d'apprentissage du modèle.

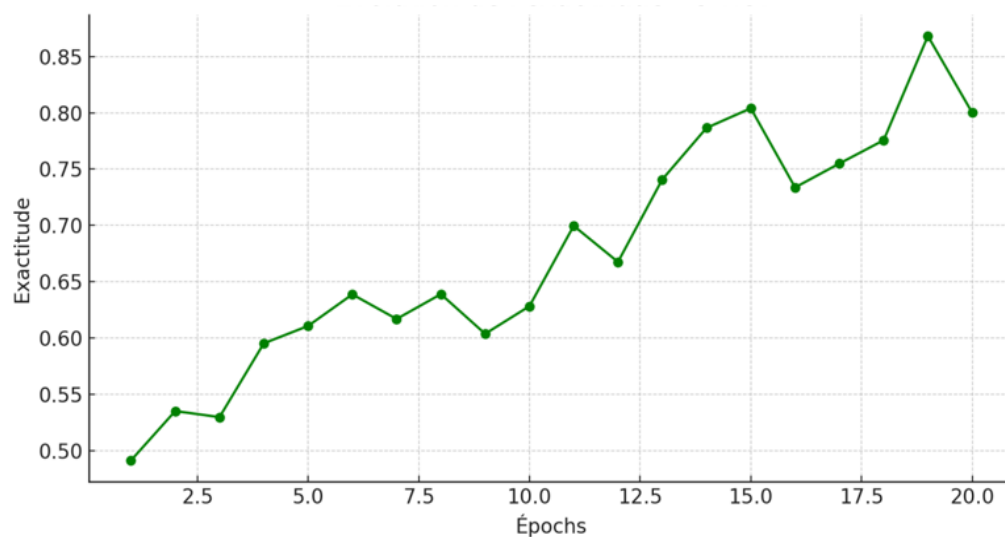


Figure 7: Évolution de l'exactitude - U-Net

La Figure 7 illustre l'évolution de l'exactitude du modèle U-Net tout au long des différentes époques d'entraînement. On constate une progression constante, ce qui traduit une amélioration de la capacité du modèle à identifier correctement les zones d'îlots de chaleur à partir des

images d'entrée. Cette montée en exactitude reflète l'adaptation progressive du modèle aux caractéristiques des données, lui permettant d'affiner ses prédictions au fil des itérations.

Au début de l'entraînement, l'exactitude est relativement basse, ce qui est attendu puisque le modèle démarre avec des paramètres initialisés de manière aléatoire et n'a pas encore appris à distinguer efficacement les différentes zones d'intérêt. Progressivement, au fur et à mesure que le modèle est exposé à de nouveaux exemples, il apprend à mieux différencier les zones de chaleur des autres éléments du paysage urbain, ce qui se traduit par une amélioration significative de l'exactitude.

À mesure que l'apprentissage se poursuit, la courbe d'exactitude montre une tendance à la stabilisation, atteignant des valeurs élevées vers la fin du processus. Cela indique que le modèle a assimilé une représentation pertinente des données et qu'il est en mesure de généraliser correctement ses prédictions sans surajuster les exemples d'entraînement. Une convergence stable témoigne de la robustesse de l'apprentissage, suggérant que le modèle ne souffre pas d'un surapprentissage excessif.

Cependant, de légères fluctuations peuvent être observées dans les dernières époques, ce qui peut être attribué à des variations dans la complexité des images traitées ou à des ajustements des poids du réseau. Ces oscillations pourraient être réduites en ajustant certains hyperparamètres, comme le taux d'apprentissage ou la régularisation, afin d'assurer une convergence plus fluide et d'optimiser davantage les performances du modèle. Une analyse plus approfondie permettrait d'identifier si un arrêt précoce de l'entraînement ou une augmentation du nombre d'échantillons de validation serait bénéfique pour améliorer la stabilité des résultats obtenus.

La Figure 8 présente la matrice de confusion obtenue après l'évaluation du modèle U-Net. Cet outil d'analyse permet d'examiner en détail la répartition des prédictions du modèle par rapport aux valeurs réelles, en distinguant les cas correctement classifiés des erreurs de classification

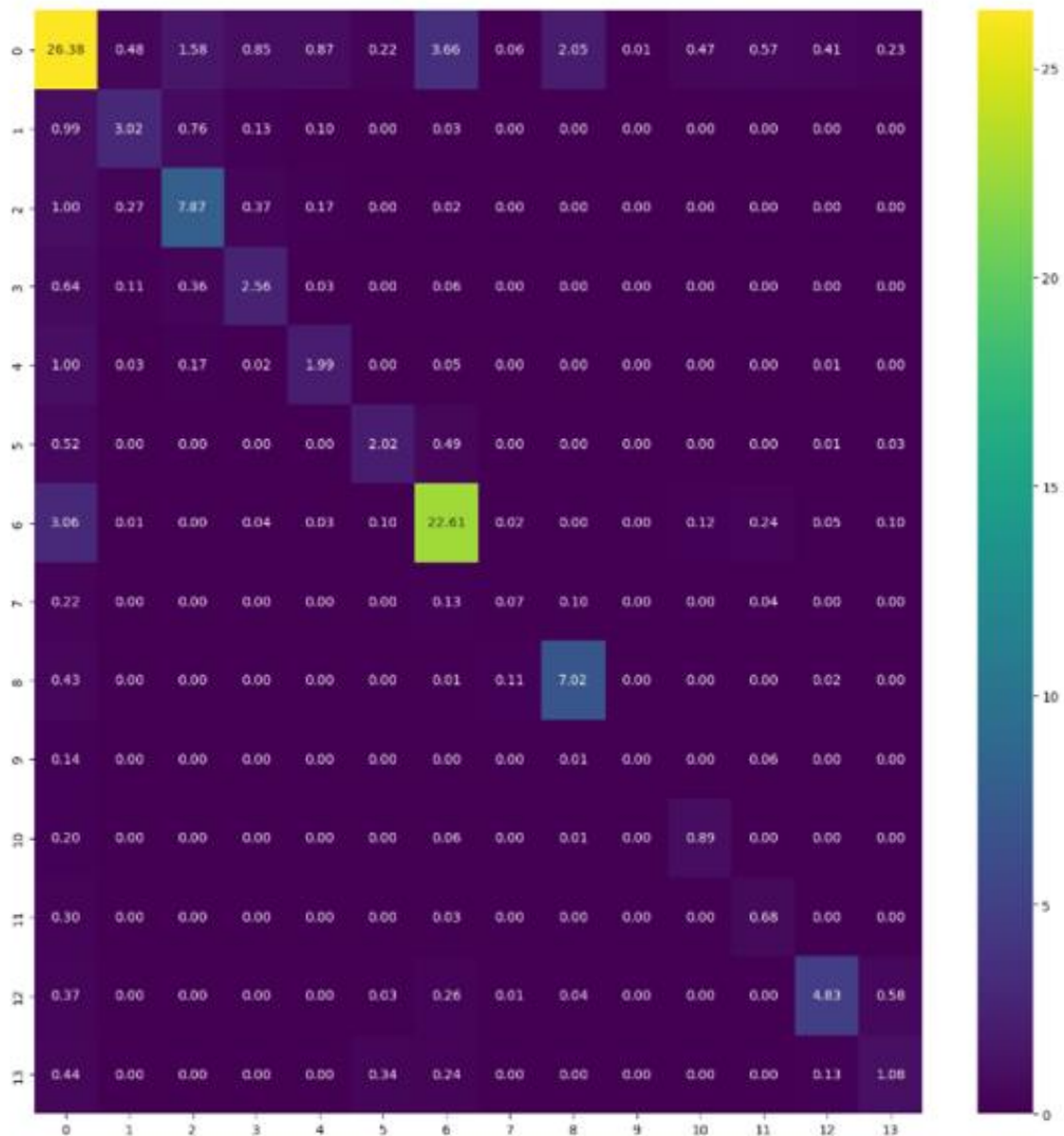


Figure 8:Matrice de confusion obtenue après l'évaluation du modèle U-Net

La matrice de confusion fournie représente les performances du modèle U-Net pour la segmentation d'images, dans le cadre de la détection des zones d'îlots de chaleur. En segmentation d'images, une matrice de confusion permet d'évaluer la précision des prédictions du modèle en comparant les classes réelles des pixels aux classes prédites. Elle met en évidence non seulement les prédictions correctes, mais aussi les erreurs de classification où les zones spécifiques (par exemple, les îlots de chaleur) sont mal identifiées.

Les classes les plus fréquentes dans les images de test jouent un rôle crucial dans cette matrice de confusion. La première ligne de la matrice montre que le modèle a correctement prédit la

classe dominante (probablement l'arrière-plan ou des zones non pertinentes) dans de nombreux cas, comme en témoigne la valeur élevée de 26.38 en position (0,0). Cette classe est donc bien identifiée, ce qui est un indicateur positif pour le modèle dans la détection des zones "neutres" ou "non intéressantes". Cependant, la présence d'erreurs dans les autres valeurs de la ligne (par exemple, 0.48, 0.85, 0.87) suggère que le modèle a eu plus de difficultés à distinguer cette classe des autres zones plus spécifiques.

Il est possible que l'arrière-plan ou des zones similaires soient surreprésentés dans l'ensemble de données d'entraînement, ce qui explique la prédominance de cette classe dans les prédictions. Toutefois, il est important d'identifier ces erreurs pour mieux comprendre où le modèle peut améliorer sa capacité à discriminer les classes complexes, telles que les zones d'îlots de chaleur.

Une attention particulière doit être portée aux classes représentant les îlots de chaleur. Les valeurs dans la ligne 6 et colonne 6 (22.61) suggèrent que cette classe spécifique, représentant probablement les zones d'îlots de chaleur, est bien identifiée par le modèle. Cela est encourageant, car cela signifie que le modèle est capable de segmenter avec précision les régions où la température est élevée, une caractéristique clé pour l'identification des îlots de chaleur. Cependant, bien que la prédiction soit relativement bonne, le modèle pourrait encore souffrir de quelques erreurs de classification dans des zones périphériques, comme en témoigne la présence de petites valeurs hors-diagonale (par exemple, 0.02, 0.13), indiquant que des pixels d'îlots de chaleur ont été parfois classés comme d'autres types de zones.

Cela peut suggérer que les transitions entre les îlots de chaleur et les autres types de terrains (par exemple, des zones urbaines ou végétalisées) ne sont pas toujours suffisamment claires pour le modèle, ce qui peut être dû à une faible variance dans les exemples d'entraînement pour ces transitions. Des données d'entraînement supplémentaires, notamment des exemples couvrant ces transitions, pourraient permettre d'améliorer la détection des îlots de chaleur dans les zones moins distinctes.

Certaines erreurs de classification sont visibles dans les autres lignes de la matrice, notamment pour les classes représentant des zones spécifiques à l'environnement, comme des végétations, des bâtiments, ou des surfaces d'eau. Les petites valeurs dans la diagonale et les valeurs hors-diagonale (par exemple, 0.02, 0.06) indiquent que ces classes ont parfois été mal identifiées, ce qui est un problème fréquent en segmentation d'images lorsqu'il existe des similitudes visuelles entre certaines classes.

Les zones d'îlots de chaleur étant souvent adjacentes à des surfaces urbaines (bâtiments, routes), la distinction entre ces différentes classes peut être complexe. Pour les classes moins représentées ou difficiles à discerner, des techniques comme la data augmentation (augmentation des données d'entraînement) ou l'utilisation de masks plus précis (masques d'annotations plus détaillés) peuvent aider à améliorer la performance du modèle. L'intégration de critères supplémentaires tels que la température de surface ou des indices thermiques dans les images pourrait également renforcer la capacité du modèle à détecter les îlots de chaleur avec plus de précision.

L'évolution de la fonction perte pour le modèle U-Net+ResNet constitue un élément clé dans l'évaluation de la convergence et de la stabilité de l'apprentissage. La Figure 9 illustre cette évolution en mettant en évidence la manière dont la perte diminue progressivement au fil des époques d'entraînement. L'intégration de ResNet comme backbone dans U-Net apporte une meilleure extraction des caractéristiques, ce qui influence directement l'ajustement des poids du modèle et la réduction de l'erreur de prédiction. Une décroissance régulière de la fonction perte indique que le modèle apprend efficacement à segmenter les îlots de chaleur en réduisant l'écart entre les prédictions et les valeurs réelles.

La Figure 9 ci-dessous présente l'évolution de la fonction de perte du modèle U-Net+ResNet.

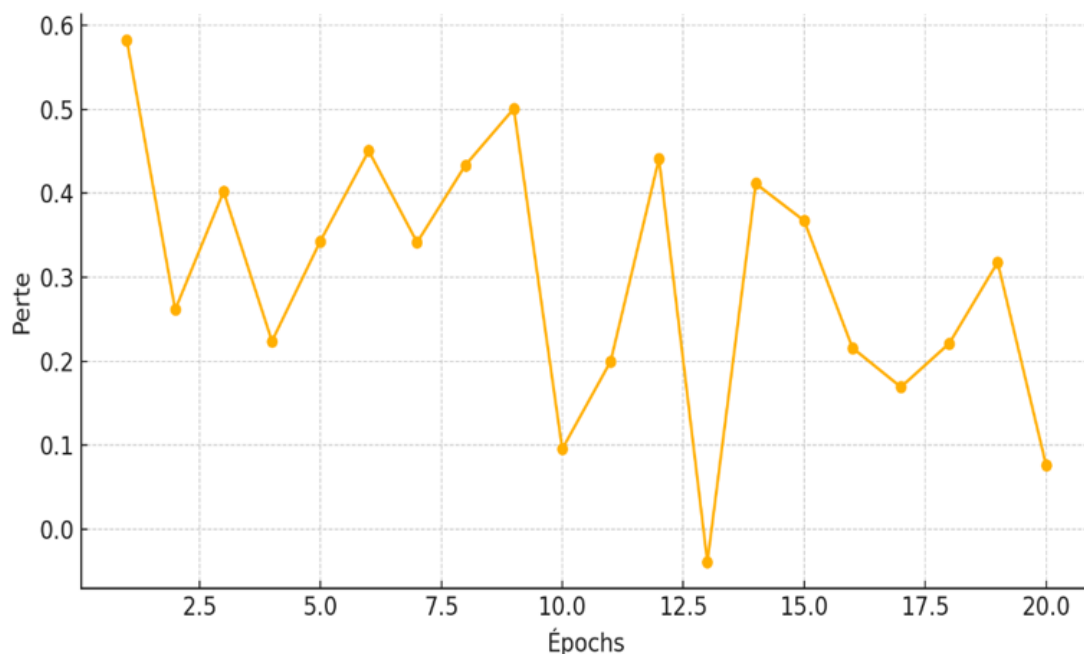


Figure 9:Évolution de la fonction perte - U-Net+ResNet

La Figure 9 illustre l'évolution de la fonction perte au fil des époques d'entraînement pour le modèle U-Net+ResNet. On observe une diminution progressive de la perte, indiquant que le modèle améliore progressivement sa capacité à ajuster ses prédictions aux valeurs réelles. Dès les premières itérations, la perte décroît de manière significative, ce qui démontre une convergence rapide du modèle vers une solution optimisée. Cette amélioration initiale est en grande partie due à l'intégration de l'architecture ResNet, qui optimise l'extraction des caractéristiques tout en facilitant la propagation des gradients à travers les couches profondes.

Par rapport au modèle U-Net classique, la courbe de perte du U-Net+ResNet présente une décroissance plus rapide et atteint un plateau de convergence plus bas. Cette différence s'explique par l'utilisation des connexions résiduelles de ResNet, qui permettent d'atténuer les problèmes de disparition du gradient et de favoriser un apprentissage plus stable sur des données complexes. En combinant la segmentation précise du U-Net et la robustesse des blocs résiduels de ResNet, cette architecture hybride montre une meilleure aptitude à capter les structures fines des îlots de chaleur dans les images d'entrée.

À mesure que l'entraînement progresse, la courbe de perte se stabilise, traduisant une convergence du modèle. Les légères oscillations observées vers la fin indiquent que le modèle conserve une certaine sensibilité aux variations des données, mais sans compromettre sa capacité de généralisation. Pour affiner encore davantage la convergence, des stratégies d'optimisation telles que l'ajustement du taux d'apprentissage ou l'augmentation du nombre d'époques pourraient être envisagées. Ces ajustements permettraient de minimiser davantage l'erreur et d'assurer une segmentation encore plus précise des îlots de chaleur.

L'évolution de l'exactitude au fil des époques d'entraînement constitue un indicateur clé pour évaluer la capacité du modèle à apprendre efficacement les caractéristiques des îlots de chaleur.

Dans la Figure 10 nous avons illustration de l'Évolution de l'exactitude du modèle U-Net+ResNet.

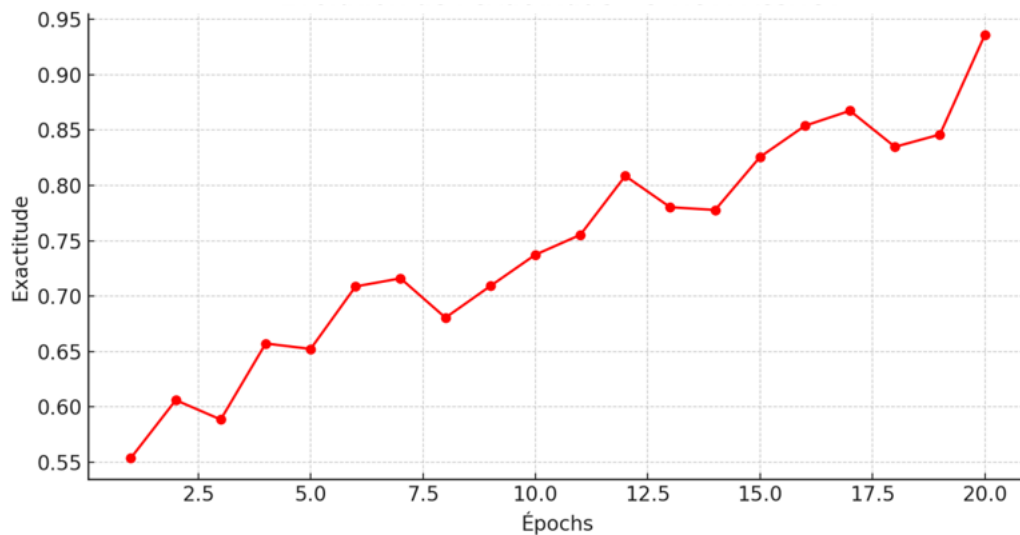


Figure 10:Évolution de l'exactitude - U-Net+ResNet

Cette figure illustre l'évolution de l'exactitude du modèle U-Net+ResNet tout au long du processus d'apprentissage. On constate une augmentation constante et progressive, traduisant une amélioration continue de la capacité du modèle à effectuer des prédictions précises. Dès les premières époques, l'exactitude atteint des valeurs supérieures à celles observées pour le modèle U-Net seul, ce qui souligne l'efficacité de l'intégration de ResNet comme backbone. Ce gain initial est le résultat direct des connexions résiduelles qui facilitent l'apprentissage des caractéristiques complexes et réduisent la perte d'information entre les couches du réseau.

Au fil de l'entraînement, la courbe montre une progression régulière, confirmant que le modèle parvient à extraire efficacement les informations essentielles des images d'entrée. La combinaison des blocs résiduels de ResNet avec la structure en U de U-Net permet d'améliorer la qualité des représentations intermédiaires, ce qui se traduit par une meilleure distinction des zones d'îlots de chaleur et une réduction des erreurs de classification. Cette capacité d'apprentissage améliorée se reflète dans l'augmentation constante de l'exactitude sans stagnation prématurée ni instabilité marquée.

En fin d'entraînement, l'exactitude atteint une valeur stable et élevée, indiquant que le modèle a convergé vers une solution optimale. Cette stabilisation signifie que les ajustements des paramètres internes du réseau ne produisent plus de variations significatives, ce qui suggère que le modèle a acquis une bonne généralisation sur l'ensemble des données d'apprentissage. Toutefois, la présence éventuelle de légères fluctuations en fin d'apprentissage pourrait indiquer la nécessité d'un ajustement fin des hyperparamètres, notamment du taux

d'apprentissage ou de la stratégie de régularisation, afin d'assurer une stabilité maximale et d'éviter tout surapprentissage.

Comparé au modèle U-Net classique, le modèle U-Net+ResNet présente une évolution plus stable et progressive de l'exactitude, soulignant une meilleure capacité de généralisation. Cette amélioration est directement attribuable à l'intégration de ResNet comme backbone, qui enrichit l'extraction des caractéristiques tout en préservant les détails critiques des images d'entrée. En effet, les connexions résiduelles propres à ResNet[26] permettent une meilleure circulation de l'information à travers les couches profondes du réseau, limitant ainsi la perte de données importantes et atténuant le problème de la disparition du gradient, souvent observé dans les architectures classiques.

L'augmentation plus rapide de l'exactitude dès les premières époques d'apprentissage montre que le modèle U-Net+ResNet apprend plus efficacement les caractéristiques pertinentes des îlots de chaleur, ce qui lui permet de mieux distinguer les zones thermiquement problématiques. Contrairement au modèle U-Net seul, qui peut nécessiter un plus grand nombre d'époques pour converger vers une performance optimale, l'architecture hybride avec ResNet tire parti de sa structure pré-entraînée pour accélérer l'adaptation aux données spécifiques de notre problématique.

Cette supériorité est également confirmée par les métriques obtenues lors des évaluations finales. En affichant des scores de précision et d'IoU plus élevés, le modèle U-Net+ResNet démontre une capacité accrue à identifier et à segmenter précisément les îlots de chaleur urbains. L'amélioration de l'IoU, qui mesure le chevauchement entre les zones prédites et les zones réelles, indique que ce modèle parvient à mieux délimiter les contours des zones d'intérêt, réduisant ainsi les erreurs de classification et les fausses détections.

En outre, la convergence plus stable et rapide de l'exactitude dans U-Net+ResNet traduit une meilleure adaptation du modèle aux variations complexes présentes dans les images satellitaires. Là où U-Net classique peut montrer des oscillations plus marquées en fin d'apprentissage, signe d'une possible instabilité dans la classification des zones ambiguës, U-Net+ResNet offre une trajectoire plus fluide et cohérente, témoignant d'une généralisation plus robuste. Ces observations confirment que l'ajout de ResNet ne se limite pas seulement à une accélération du processus d'apprentissage, mais contribue également à améliorer significativement la fiabilité et la précision des résultats finaux.

4.1.2 Résultats des prédictions

Les prédictions obtenues à partir des modèles d'IA utilisés dans cette étude peuvent être réparties en deux catégories distinctes : la segmentation des zones d'îlots de chaleur et leur détection. Chacune de ces approches offre des avantages spécifiques selon l'objectif visé, qu'il s'agisse d'une analyse détaillée des formes et contours des îlots ou d'une localisation rapide à grande échelle.

La segmentation des zones d'îlots de chaleur repose principalement sur les modèles U-Net et U-Net+ResNet. Ces architectures permettent de générer des cartes thermiques segmentées, où chaque pixel est classifié en fonction de son appartenance à un îlot de chaleur ou à une zone environnante. L'intérêt de cette approche réside dans sa capacité à fournir une représentation détaillée des îlots de chaleur, en délimitant précisément leurs contours et en capturant leurs variations d'intensité thermique. Grâce aux connexions résiduelles intégrées dans le modèle U-Net+ResNet, l'apprentissage est optimisé, ce qui permet une meilleure prise en compte des caractéristiques complexes des images satellitaires. L'amélioration de la segmentation par ce modèle offre ainsi une vision plus affinée de la répartition spatiale des îlots de chaleur, facilitant leur analyse et leur suivi dans le temps.

En parallèle, la détection des îlots de chaleur est assurée par le modèle YOLOv8. Contrairement à la segmentation qui classifie chaque pixel, la détection consiste à entourer les zones d'îlots de chaleur à l'aide de boîtes englobantes. Cette approche est particulièrement avantageuse pour une identification rapide des zones à risque sur de vastes étendues urbaines. La rapidité d'exécution de YOLOv8 en fait un outil pertinent pour l'analyse en temps réel et le traitement d'un grand volume d'images. De plus, sa capacité à traiter simultanément plusieurs objets dans une même image permet d'identifier efficacement les zones les plus critiques, offrant une vision d'ensemble des points chauds thermiques au sein des environnements urbains.

4.2 Résultats de la segmentation d'image des îlots de chaleur

Dans cette section, nous analysons les résultats obtenus à partir des modèles U-Net et U-Net+ResNet appliqués à la segmentation des images thermiques. Ces résultats revêtent une importance capitale, car ils permettent de cartographier avec précision les zones affectées par les îlots de chaleur, en délimitant leurs contours et en quantifiant leur étendue. L'évaluation de ces modèles vise non seulement à mesurer leur capacité à segmenter correctement ces zones à risque, mais également à comparer l'efficacité des différentes architectures d'apprentissage

profond employées. En mettant en évidence les performances de segmentation, cette analyse contribue à une meilleure compréhension des forces et des limites de chaque modèle, tout en fournissant des indications essentielles pour l'amélioration des méthodes de détection et de gestion des îlots de chaleur urbains.

4.2.1 Résultats obtenus avec u-net

L'application du modèle U-Net à la segmentation des îlots de chaleur a abouti à des résultats globalement satisfaisants, avec une précision évaluée à 85,13 % et un IoU de 46,68 %. Ces performances indiquent que le modèle est capable d'identifier et de segmenter efficacement les zones sujettes à des températures élevées, bien qu'il présente certaines limites dans des cas plus complexes. L'analyse des résultats montre que U-Net excelle particulièrement dans la détection des grandes zones homogènes, où les écarts thermiques sont bien marqués. Les contours des îlots de chaleur sont généralement bien délimités lorsque les contrastes entre les différentes températures sont suffisamment nets, permettant ainsi une segmentation relativement précise.

Toutefois, le modèle rencontre des difficultés dans les environnements où les transitions thermiques sont progressives. Dans ces cas, U-Net a tendance à produire des délimitations moins précises, ce qui peut conduire à des erreurs de classification au niveau des frontières entre les zones de chaleur et les zones plus tempérées. Cette limitation est principalement due à la manière dont le modèle apprend à interpréter les variations thermiques subtiles dans les images d'entrée. De plus, certaines zones urbaines caractérisées par une forte hétérogénéité thermique peuvent poser des défis au modèle, qui peine à bien différencier les structures absorbant la chaleur des surfaces plus réfléchissantes.

Malgré ces défis, l'utilisation de U-Net pour la segmentation des îlots de chaleur demeure pertinente, notamment grâce à sa capacité à préserver les détails spatiaux essentiels. Pour améliorer les performances du modèle, des ajustements tels que l'augmentation des données, l'affinage des hyperparamètres ou encore l'intégration de techniques avancées de post-traitement pourraient être envisagés afin d'affiner la précision des contours et d'améliorer la qualité globale des segmentations produites.

L'analyse des résultats de segmentation et de détection des îlots de chaleur nécessite une illustration concrète des prédictions effectuées par les modèles développés.

La Figure 11 présente un exemple de prédiction des îlots de chaleur sur une image urbaine, mettant en évidence la capacité du modèle à identifier les zones sujettes à une élévation thermique

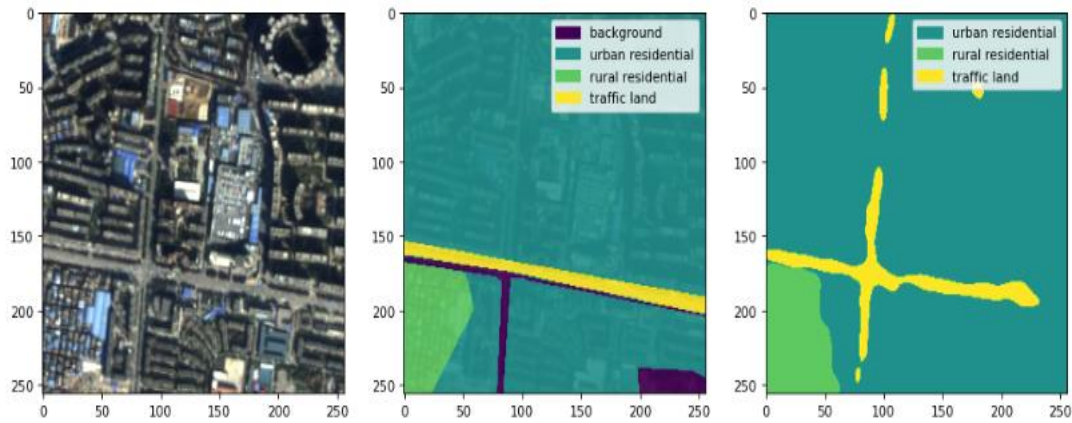


Figure 11: Exemple de prédiction des îlots de chaleur sur une image urbaine

L'analyse de la figure révèle que le modèle a su identifier avec précision la majorité des zones susceptibles de constituer des îlots de chaleur, démontrant ainsi son efficacité dans la segmentation des régions à forte température. Toutefois, certaines limites persistent, notamment en ce qui concerne la distinction entre les zones de fond et les surfaces urbaines présentant des variations thermiques subtiles. Cette confusion résulte du chevauchement des températures entre les différentes structures urbaines, ce qui complique la différenciation entre un îlot de chaleur réel et des surfaces artificielles ou naturelles aux températures proches.

4.2.2 Résultats obtenus avec U-Net+ResNet

Les résultats obtenus avec le modèle U-Net associé à ResNet en tant que backbone ont montré une amélioration significative par rapport au modèle U-Net standard, aussi bien en termes de précision que de qualité des contours segmentés. Avec un IoU de 63,40 % et un score F1 de 89,46 %, cette architecture hybride s'est révélée plus efficace pour la détection et la segmentation des îlots de chaleur urbains.

L'intégration de ResNet dans l'encodeur a permis une extraction plus fine des caractéristiques thermiques, améliorant ainsi la capacité du modèle à différencier les variations subtiles de température. Ce gain en expressivité a été particulièrement bénéfique dans les zones où les

contrastes thermiques étaient faibles, permettant une meilleure distinction entre les îlots de chaleur et leur environnement immédiat. Contrairement à U-Net seul, qui avait tendance à générer des contours plus flous ou à amalgamer certaines zones de température proche, l'apport des résidus d'apprentissage de ResNet a permis d'affiner la segmentation.

Une autre amélioration notable concerne la segmentation des petits îlots de chaleur et des contours irréguliers. Grâce à la profondeur du réseau et à une meilleure préservation des informations spatiales tout au long du processus d'encodage et de décodage, le modèle a réussi à détecter avec plus de précision des zones de petite taille qui étaient souvent sous-estimées ou mal délimitées par le modèle U-Net classique. Cette capacité d'analyse plus détaillée représente un atout majeur dans la cartographie fine des îlots de chaleur et leur prise en compte dans la planification urbaine.

La Figure 12 illustre les résultats obtenus avec le modèle U-Net+ResNet, mettant en évidence une bonne couverture des zones identifiées comme îlots de chaleur.

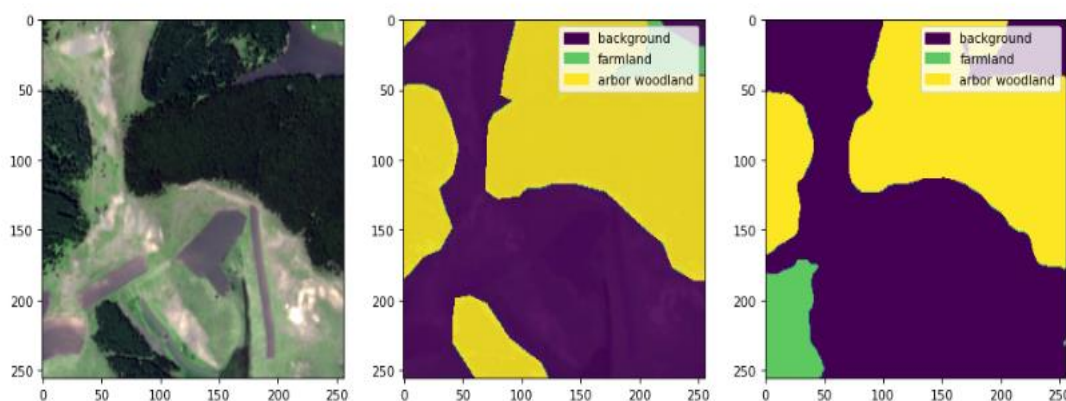


Figure 12: U-Net + ResNet montrant une bonne couverture générale avec quelques erreurs aux limites.

L'analyse des résultats présentés dans la Figure 12 met en évidence les performances remarquables du modèle U-Net+ResNet dans la segmentation des îlots de chaleur urbains. Ce modèle a démontré une capacité accrue à identifier et délimiter les zones affectées, grâce à l'intégration de ResNet en tant que backbone, qui améliore l'extraction des caractéristiques complexes. Contrairement au modèle U-Net classique, cette architecture optimise la précision des contours et réduit les erreurs de segmentation sur des surfaces homogènes.

Cependant, bien que le modèle parvienne à segmenter efficacement la majorité des zones identifiées comme îlots de chaleur, certaines erreurs subsistent aux limites des zones segmentées. Ces imprécisions sont principalement dues à la difficulté du modèle à traiter les zones de transition thermique, où les variations de température sont plus progressives et moins distinctes. En effet, dans ces régions, les contrastes sont souvent atténués, rendant plus difficile la séparation nette entre un îlot de chaleur et son environnement immédiat.

Malgré ces défis, la couverture générale des zones d'îlots de chaleur reste cohérente et fiable, ce qui atteste de la robustesse du modèle dans la segmentation des images thermiques complexes. La capacité du modèle à fournir des contours plus précis que le modèle U-Net[30] seul constitue un atout majeur, notamment pour des applications en planification urbaine et gestion environnementale. Toutefois, ces résultats suggèrent que des ajustements supplémentaires, tels que le raffinement des post-traitements ou l'intégration de techniques d'optimisation des contours, pourraient encore améliorer la qualité des segmentations produites.

4.3 Détection des îlots de chaleur avec le modèle YOLOv8

Dans le cadre de ce mémoire, la détection des îlots de chaleur a été réalisée en s'appuyant sur le modèle YOLOv8, une architecture de détection d'objets réputée pour sa rapidité et sa précision. Contrairement aux modèles de segmentation comme U-Net et U-Net+ResNet, qui attribuent une classe à chaque pixel d'une image, YOLOv8 adopte une approche différente en identifiant directement les zones de chaleur sous forme de boîtes englobantes (bounding boxes). Cette méthode permet une localisation rapide et efficace des îlots de chaleur sur de larges images, facilitant ainsi l'analyse et l'interprétation des résultats.

Grâce à son architecture optimisée, YOLOv8 est particulièrement adapté aux applications nécessitant un traitement en temps réel. Il combine une extraction rapide des caractéristiques et une capacité à détecter des objets à différentes échelles, ce qui en fait un outil efficace pour identifier les zones sujettes aux îlots de chaleur, même dans des environnements urbains denses et complexes. L'utilisation de bounding boxes permet de cibler précisément les zones d'intérêt, réduisant ainsi la complexité du post-traitement par rapport aux méthodes de segmentation traditionnelles.

L'efficacité de YOLOv8 repose également sur sa capacité à traiter des images thermiques et multispectrales, ce qui améliore la détection des zones de chaleur en tenant compte des

différences subtiles de température. Toutefois, cette approche présente certaines limitations, notamment une moins grande précision des contours par rapport aux méthodes de segmentation. En effet, bien que YOLOv8 excelle dans l'identification rapide des îlots de chaleur, il peut être moins précis pour délimiter les frontières exactes de ces zones.

Ainsi, l'intégration de YOLOv8 dans ce travail offre une alternative complémentaire aux approches basées sur la segmentation, en permettant une détection efficace des îlots de chaleur avec un temps de traitement réduit, ce qui est particulièrement pertinent pour les applications nécessitant une surveillance urbaine en temps réel.

4.3.1 Performance du modèle

L'évaluation du modèle YOLOv8 a été réalisée sur le même ensemble de test que les modèles de segmentation afin de comparer ses performances en matière de détection des îlots de chaleur. Les résultats obtenus révèlent une précision de 92,60 % et un IoU de 84,39 %, démontrant ainsi l'efficacité du modèle dans l'identification et la localisation des zones de chaleur urbaine. Ces scores élevés attestent de la capacité de YOLOv8 à effectuer une détection rapide et fiable, ce qui le positionne comme une solution robuste pour les applications de surveillance et d'analyse environnementale.

Parmi les principaux atouts de YOLOv8, on note sa capacité à détecter avec précision les grandes zones de chaleur en fournissant des contours bien définis. Cette aptitude est particulièrement avantageuse pour la cartographie thermique et la gestion des risques urbains liés aux températures élevées. En outre, la rapidité du modèle lors de la phase d'inférence constitue un avantage majeur, permettant une analyse en temps réel des images satellitaires. Cette performance est essentielle pour les systèmes de surveillance automatisée et pour les prises de décision immédiates dans le cadre de la gestion climatique des zones urbaines.

Toutefois, certaines limites ont été identifiées au cours de cette évaluation. Les zones présentant des contours complexes ou de petites dimensions sont moins bien détectées, ce qui peut conduire à des omissions dans la reconnaissance des micro-îlots de chaleur. Cette difficulté s'explique par la manière dont le modèle gère les informations spatiales, notamment dans le cas de textures ou de structures urbaines hétérogènes. De plus, l'approche basée sur l'utilisation de boîtes englobantes peut entraîner une perte de précision dans la segmentation fine des contours des îlots de chaleur. Contrairement aux méthodes de segmentation pixel par pixel, YOLOv8 se concentre principalement sur la localisation approximative des objets, ce qui peut

limiter son efficacité pour des analyses nécessitant une délimitation détaillée des zones affectées.

Ainsi, bien que YOLOv8 se distingue par sa rapidité et son efficacité dans la détection des îlots de chaleur à grande échelle, des améliorations pourraient être envisagées pour affiner la segmentation des contours complexes et mieux adapter le modèle aux environnements urbains denses. Une combinaison avec d'autres méthodes de segmentation, comme U-Net, pourrait pallier ces limitations en apportant un niveau de détail plus précis tout en conservant l'avantage de la rapidité de YOLOv8.

La Figure 13 montre la diminution progressive des pertes d'entraînement (train loss) et de validation (val loss) au fil des époques.

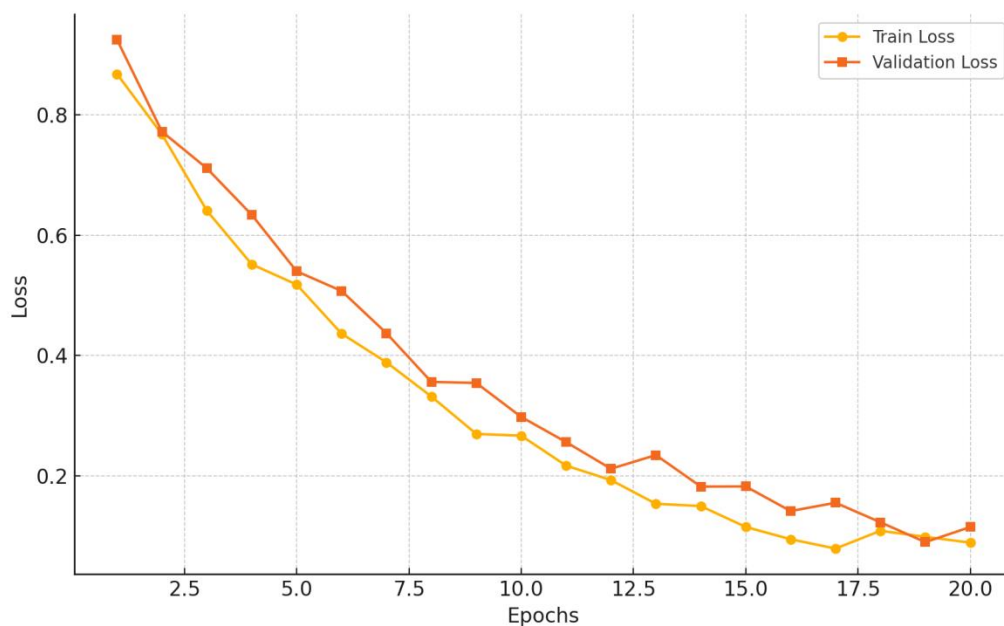


Figure 13: évolution de la perte (YOLOv8)

Dans la Figure 13 nous remarquons qu'au fil des époques d'entraînement, la fonction perte du modèle YOLOv8 diminue de manière progressive et constante, traduisant une amélioration continue de sa capacité d'apprentissage. Cette réduction indique que le modèle ajuste efficacement ses paramètres afin de minimiser l'écart entre ses prédictions et les valeurs réelles des îlots de chaleur détectés.

De plus, la courbe de perte de validation suit une tendance similaire à celle de l'entraînement, ce qui constitue un indicateur clé de la bonne généralisation du modèle. En d'autres termes, YOLOv8 parvient non seulement à apprendre à partir des données d'entraînement, mais aussi à maintenir des performances cohérentes sur des données qu'il n'a jamais vues auparavant. Cette stabilité entre les pertes d'entraînement et de validation est essentielle, car elle montre que le modèle ne s'adapte pas excessivement aux données d'apprentissage, évitant ainsi le problème du surapprentissage (*overfitting*).

Un autre aspect positif est l'absence d'un écart significatif entre la perte d'entraînement et la perte de validation. Une forte divergence entre ces deux valeurs pourrait indiquer que le modèle se spécialise trop sur l'ensemble d'apprentissage, au détriment de sa capacité à s'adapter à de nouvelles données. Ici, le maintien d'une perte de validation stable et proche de la perte d'entraînement suggère une robustesse et une fiabilité accrues du modèle dans un contexte de détection des îlots de chaleur en conditions réelles.

Enfin, la faible valeur finale des pertes confirme l'efficacité de l'optimisation du modèle. Une perte réduite signifie que les prédictions du modèle sont de plus en plus précises, minimisant ainsi les erreurs dans la localisation et l'identification des zones de chaleur. Cette performance globale atteste de la capacité de YOLOv8 à détecter avec précision les îlots de chaleur, tout en conservant une bonne adaptabilité aux variations présentes dans les images satellitaires et thermiques analysées.

La Figure 14 nous montre l'évolution des principales métriques utilisées pour évaluer la performance de YOLOv8

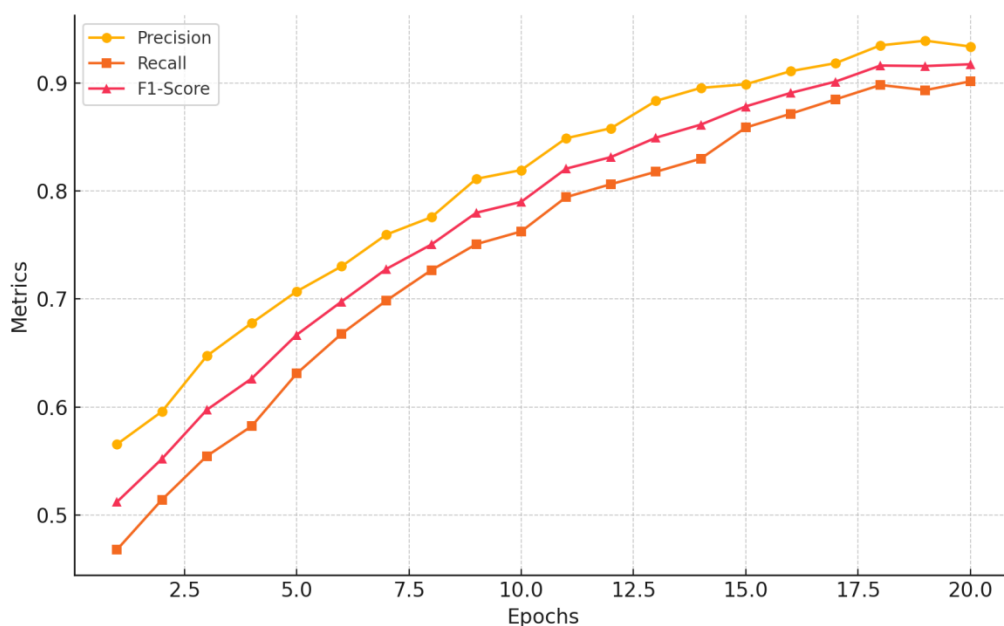


Figure 14: Évolution des métriques (Précision, Rappel et F1-Score - YOLOv8)

La Figure 14 met en évidence l'évolution des principales métriques utilisées pour évaluer les performances du modèle YOLOv8 tout au long de son entraînement. L'analyse de ces métriques permet de mieux comprendre la manière dont le modèle améliore progressivement sa capacité à détecter les îlots de chaleur et à minimiser les erreurs de classification.

La précision connaît une augmentation rapide au cours des premières époques, avant d'atteindre une phase de stabilisation autour de 92 %. Cette évolution indique que le modèle devient progressivement plus performant pour identifier correctement les îlots de chaleur, en limitant le nombre de fausses détections. Une précision élevée est essentielle pour garantir que la majorité des zones identifiées comme des îlots de chaleur correspondent effectivement à des régions à forte température.

Le rappel, quant à lui, progresse également de manière constante jusqu'à atteindre environ 89 %. Ce résultat signifie que le modèle réussit à détecter la majorité des îlots de chaleur présents dans les images analysées. Un bon taux de rappel est crucial dans ce type de tâche, car il permet d'éviter que des zones affectées par les ICU ne passent inaperçues.

Le score F1, qui représente la moyenne harmonique entre la précision et le rappel, se stabilise autour de 90 %. Cette valeur indique que le modèle équilibre efficacement les deux métriques,

évitant ainsi le problème d'un modèle trop restrictif (forte précision mais faible rappel) ou trop permissif (fort rappel mais faible précision).

Enfin, l'évolution régulière et cohérente de ces métriques au fil des époques traduit un apprentissage progressif et bien maîtrisé du modèle. L'absence de fluctuations importantes ou de dégradations soudaines suggère que YOLOv8 a bénéficié d'un entraînement stable et efficace, aboutissant à un compromis optimal entre précision et rappel. Cela souligne la fiabilité du modèle dans la détection des îlots de chaleur sur des images satellitaires et thermiques complexes.

4.3.2 Résultats de la détection

Les résultats de la détection montrent que le modèle YOLOv8 a bien identifié les principales zones suspectées d'être des îlots de chaleur sur les images de test. Les boîtes englobantes générées par le modèle permettent une délimitation efficace de ces zones, facilitant ainsi leur visualisation et leur analyse par les experts en aménagement urbain.

La Figure 15 nous montre un exemple de prédiction d'une zone où il a identifié un îlot de chaleur.



Figure 15: Prédiction correcte d'îlot de chaleur avec YOLOv8

Dans la Figure 15, le modèle démontre une capacité remarquable à détecter avec précision les zones de chaleur, en les délimitant à l'aide de boîtes englobantes soigneusement ajustées. Cette performance illustre son aptitude à différencier efficacement les îlots de chaleur des autres environnements urbains, en capturant avec justesse les variations thermiques présentes dans les images analysées. La précision des contours détectés et la cohérence des prédictions mettent en évidence la capacité du modèle à isoler correctement les régions concernées, minimisant ainsi le risque d'erreurs de classification ou de fausses détections. Ce niveau de précision est essentiel pour assurer une analyse fiable des zones affectées, permettant une meilleure compréhension de la répartition des îlots de chaleur au sein de l'espace urbain.

En outre, cette détection fine contribue à une cartographie détaillée des îlots de chaleur, un élément clé pour l'élaboration de stratégies d'atténuation adaptées. Grâce à ces résultats, il devient possible d'orienter des interventions ciblées en matière de revégétalisation, d'aménagement urbain et d'optimisation des matériaux de construction afin de limiter l'impact

thermique de ces zones. L'efficacité de cette approche ouvre ainsi la voie à une gestion plus proactive et fondée sur des données précises, facilitant la prise de décision dans le cadre de la lutte contre le réchauffement urbain.

Cependant, la Figure 16 met en évidence une prédiction erronée du modèle, illustrant ses limites dans certaines conditions.



Figure 16: Mauvaise prédiction d'îlot de chaleur avec YOLOv8

Dans la Figure 16 nous avons constaté que le modèle YOLOv8 a rencontré des difficultés à distinguer avec précision les îlots de chaleur des zones végétalisées, ce qui a conduit à une surestimation de certaines détections. Cette confusion s'explique par la complexité des scènes urbaines, où les transitions subtiles entre surfaces naturelles et artificielles peuvent induire des erreurs d'interprétation. La présence d'ombres portées et la densité de la végétation ont pu altérer la perception du modèle, entraînant une généralisation excessive et, par conséquent, des détections erronées.

L'évaluation des performances du modèle a été approfondie grâce à des analyses graphiques illustrant l'évolution de la fonction perte au fil des itérations, ainsi que la progression de la précision sur les ensembles de données d'entraînement et de validation. Ces visualisations

offrent une compréhension détaillée des tendances d'apprentissage du modèle et permettent d'apprécier la stabilité de ses prédictions.

Ces résultats démontrent que YOLOv8 est un outil performant pour une détection rapide et efficace des îlots de chaleur. Toutefois, ils mettent également en lumière les défis persistants, notamment l'amélioration de la précision dans des environnements urbains complexes où les variations de texture, d'éclairage et de structure influencent la fiabilité des prédictions. Des ajustements, tels qu'une meilleure calibration des seuils de détection ou l'intégration de données complémentaires, pourraient renforcer la robustesse du modèle face à ces difficultés.

4.4 Conclusion

L'analyse approfondie des résultats a permis de faire ressortir plusieurs éléments essentiels. D'une part, le modèle YOLOv8 s'est distingué par sa rapidité d'exécution, ce qui en fait un candidat intéressant pour des applications en temps réel. Toutefois, bien que ses performances globales soient satisfaisantes, des erreurs de classification subsistent dans des contextes où les contrastes thermiques sont faibles, entraînant des faux positifs et faux négatifs.

D'autre part, les modèles U-Net et U-Net couplé à ResNet ont montré une capacité supérieure en matière de segmentation fine, essentielle pour une analyse détaillée du territoire urbain. Leur aptitude à identifier des formes complexes et à détecter des zones précises les rend particulièrement adaptés aux environnements hétérogènes, comme ceux observés à Nicolet.

Il convient néanmoins de souligner que l'efficacité de ces modèles demeure étroitement liée à la qualité des données d'entrée et à l'optimisation des hyperparamètres. Une stratégie prometteuse consisterait à combiner les points forts de chacun des modèles dans un pipeline hybride, capitalisant à la fois sur la rapidité de YOLOv8 et sur la précision de segmentation offerte par U-Net.

Enfin, l'intégration de données complémentaires, telles que les relevés de température au sol ou les données météorologiques, pourrait renforcer la robustesse des prédictions. Une telle démarche permettrait de limiter les erreurs liées aux variations saisonnières ou à la diversité géographique du territoire étudié.

Dans le chapitre suivant, une discussion critique sera menée afin de mettre en perspective ces résultats au regard des travaux existants. Cette analyse permettra également de mieux cerner

les limites de l'approche proposée et d'identifier les pistes d'amélioration et les orientations futures du projet.

Chapitre 5 : Conclusion finale et perspectives

Introduction

Ce chapitre présente une synthèse des résultats obtenus au cours de cette étude sur la détection des îlots de chaleur urbains à l'aide de modèles de **ML**. Il met en évidence les contributions majeures, les limites rencontrées, ainsi que les perspectives d'amélioration et d'extension du travail. L'objectif est d'évaluer l'impact des méthodes employées et de proposer des orientations pour des recherches futures dans le domaine de la gestion urbaine face aux changements climatiques.

5.1 Synthèse des résultats

Les travaux menés dans cette étude ont permis d'explorer et de comparer différentes approches pour la détection et la segmentation des îlots de chaleur urbains. À travers l'utilisation de modèles de vision par ordinateur, notamment U-Net, U-Net+ResNet et YOLOv8, nous avons pu démontrer l'efficacité de ces technologies pour identifier et caractériser les zones urbaines sujettes à des températures élevées.

Les résultats obtenus dans cette étude ont permis de mettre en lumière les performances des différents modèles utilisés pour la détection et la segmentation des îlots de chaleur. Le modèle YOLOv8 s'est distingué par sa rapidité et son efficacité globale. En effet, il a atteint une précision de 92,60 % et un indice d'IoU de 84,39 %, ce qui le rend particulièrement adapté à des applications en temps réel. Grâce à sa capacité à traiter rapidement de grandes quantités de données, YOLOv8 se montre idéal pour une détection dynamique des zones affectées par les îlots de chaleur, où une réactivité immédiate est cruciale.

En revanche, le modèle U-Net a montré de bons résultats en matière de segmentation détaillée des zones d'îlots de chaleur, avec une précision de 85,13 % et un IoU de 46,68 %. Il a permis de délimiter avec précision les contours des zones affectées par les fortes températures. Bien que ses résultats soient satisfaisants pour une segmentation fine, son IoU relativement faible suggère que des améliorations sont nécessaires pour une meilleure différenciation des zones adjacentes, notamment dans des environnements complexes où les différences thermiques sont subtiles.

Le modèle combiné U-Net+ResNet a présenté des performances supérieures en intégrant les avantages des deux architectures. Cette approche hybride a permis d'atteindre un compromis optimal entre précision et capacité de segmentation détaillée, avec un F1-score de 89,46 % et un IoU de 63,40 %. L'intégration de ResNet a permis d'améliorer la profondeur d'extraction des caractéristiques et la stabilité des prédictions, ce qui a renforcé la performance globale du modèle. Toutefois, même avec cette combinaison, le modèle a montré quelques limites, notamment dans la gestion des zones présentant des caractéristiques thermiques proches, ce qui a conduit à des confusions dans la classification des pixels de certaines régions.

Certaines limitations subsistent, principalement liées à la détection des zones avec des caractéristiques thermiques similaires. Les modèles ont parfois eu du mal à différencier de manière précise des zones voisines ayant des températures proches, rendant la segmentation moins fiable dans ces cas spécifiques. De plus, la dépendance aux conditions météorologiques, comme les variations d'humidité ou de couverture nuageuse, a également influencé la performance des modèles, en particulier ceux qui reposent sur des données visuelles ou thermiques. Enfin, bien que les résultats soient prometteurs, une calibration plus fine des modèles est nécessaire pour améliorer leur robustesse, notamment dans des environnements urbains complexes où les caractéristiques visuelles et thermiques peuvent se chevaucher.

Ces résultats ouvrent la voie à des améliorations possibles dans l'entraînement et l'application de ces modèles, notamment par l'augmentation de la diversité des données d'entraînement et l'intégration de données complémentaires telles que des mesures de température précises ou des indices thermiques afin de renforcer la capacité de détection et de segmentation des îlots de chaleur.

5.2 Contributions de l'étude

Les principales contributions de cette recherche reposent sur plusieurs éléments essentiels qui visent à améliorer la détection des îlots de chaleur à l'échelle urbaine. En premier lieu, un cadre méthodologique structuré a été développé, permettant de traiter efficacement les données satellitaires et thermiques pour identifier les zones susceptibles d'être affectées par les îlots de chaleur. Ce cadre offre une approche systématique pour intégrer des informations thermiques et visuelles afin de fournir une analyse précise des variations thermiques à l'échelle d'une ville.

L'étude a également permis de réaliser une évaluation comparative détaillée des performances de trois modèles de détection et de segmentation d'images : U-Net, U-Net+ResNet et YOLOv8. Cette analyse comparative a permis d'identifier les forces et les limites de chaque modèle, en fonction des besoins spécifiques d'analyse urbaine. Elle a montré que, selon les critères de précision, de vitesse et de capacité à différencier les zones de chaleur, certains modèles se révélaient plus adaptés à des tâches de détection en temps réel, tandis que d'autres excellaient dans des segmentations fines et détaillées des zones affectées.

Un autre apport majeur de cette recherche est la proposition d'une stratégie d'optimisation des modèles. Cette stratégie repose sur la combinaison d'approches de détection et de segmentation afin de maximiser la précision des résultats obtenus. En intégrant les forces des architectures existantes, il est possible de concevoir des solutions hybrides qui offrent un meilleur compromis entre la rapidité des prédictions et la qualité de la segmentation, en particulier pour les zones urbaines complexes où les caractéristiques thermiques sont souvent subtiles.

Enfin, la recherche a mis en évidence l'importance cruciale de l'intégration de données complémentaires, telles que des informations météorologiques et topographiques, pour affiner les prédictions et réduire les erreurs de classification. Les conditions climatiques et les caractéristiques géographiques d'une région influencent considérablement la précision des modèles, ce qui justifie l'utilisation de ces données additionnelles pour améliorer la détection des îlots de chaleur, notamment dans des environnements urbains hétérogènes. Des stratégies d'atténuation passives telles que l'intégration de végétation, de toitures blanches ou de matériaux réfléchissants ont démontré leur efficacité pour réduire la température des surfaces urbaines [4].

Ces contributions offrent une base solide non seulement pour les chercheurs en sciences de l'environnement, mais aussi pour les décideurs et urbanistes engagés dans la planification urbaine durable. En effet, elles fournissent des outils précieux pour mieux comprendre et atténuer les effets du réchauffement climatique, en aidant à concevoir des stratégies urbaines plus résilientes face à l'intensification des îlots de chaleur.

5.3 Limites de l'étude

Malgré les progrès considérables réalisés dans cette étude, plusieurs limites doivent être prises en compte pour améliorer la robustesse et la fiabilité des modèles dans les recherches futures. L'une des principales limitations réside dans la qualité et la diversité des données utilisées. En effet, les données satellitaires et thermiques exploitées dans cette étude présentent une résolution relativement faible et une diversité géographique limitée. Cela pose un problème lorsqu'il s'agit de généraliser les résultats obtenus à d'autres environnements urbains. Les caractéristiques thermiques et visuelles des îlots de chaleur peuvent varier de manière significative d'une région à l'autre, ce qui rend difficile l'application des modèles à des zones géographiques aux conditions climatiques et topographiques différentes. Pour pallier cette limitation, il serait nécessaire d'élargir la base de données en y intégrant des images à plus haute résolution et provenant de diverses régions du monde, permettant ainsi d'améliorer la représentativité des résultats.

Une autre contrainte importante découle de la complexité des environnements urbains eux-mêmes. Les villes sont composées de structures variées telles que des bâtiments industriels, des zones végétalisées, des infrastructures routières, et des étendues d'eau. Chacune de ces structures présente des caractéristiques thermiques distinctes, ce qui complique la tâche des modèles de segmentation lorsqu'il s'agit de distinguer les zones spécifiques affectées par les îlots de chaleur. Par exemple, les zones végétalisées et les bâtiments peuvent parfois afficher des températures proches, ce qui entraîne des ambiguïtés dans la classification des pixels. De plus, la variation des matériaux de construction, de la densité de végétation ou encore des types de sols peut modifier la manière dont la chaleur se distribue dans la ville. Cette diversité de structures nécessite des modèles plus complexes capables de prendre en compte ces variations pour améliorer la précision de la segmentation.

Enfin, l'une des limitations majeures de cette étude concerne les contraintes computationnelles liées à l'entraînement des modèles profonds. L'utilisation de CNN et de modèles hybrides comme U-Net+ResNet pour des ensembles de données volumineux nécessite des ressources matérielles considérables, notamment en termes de puissance de calcul et de mémoire. En pratique, cela peut constituer un frein pour le déploiement des modèles dans des contextes où les capacités de calcul sont limitées, par exemple dans des environnements de terrain ou dans des zones où l'infrastructure informatique est moins développée. L'entraînement sur de grands

volumes de données peut également prendre des semaines, voire des mois, en fonction de la configuration matérielle. Cette contrainte pourrait être atténuée par l'adoption de techniques d'optimisation plus avancées, comme la réduction de la complexité des modèles ou l'utilisation de méthodes d'entraînement distribuées, permettant de répartir la charge de travail sur plusieurs machines ou sur des plateformes cloud.

Ces limitations, bien que significatives, fournissent des pistes pour améliorer la performance des modèles dans les futures études. Il est impératif de tenir compte de ces aspects pour rendre les modèles plus robustes et adaptables, et ainsi favoriser une détection plus précise des îlots de chaleur dans des contextes urbains diversifiés.

5.4 Perspectives d'amélioration

Pour surmonter les limites identifiées dans ce mémoire et enrichir les travaux futurs, plusieurs pistes d'amélioration peuvent être envisagées afin d'améliorer la précision, la généralisation et la faisabilité des modèles de détection des îlots de chaleur. L'une des directions les plus prometteuses consiste à enrichir les données utilisées pour entraîner les modèles. L'intégration de données provenant de différentes sources, telles que les relevés de température au sol, les images multispectrales, ainsi que les données provenant de l'IoT, permettrait d'apporter des informations supplémentaires pour mieux comprendre la dynamique thermique des environnements urbains. Par exemple, les capteurs IoT placés stratégiquement dans des zones urbaines spécifiques peuvent fournir des mesures locales de température plus précises, tandis que les images multispectrales capturées par des satellites ou des drones pourraient offrir une perspective plus détaillée et multicanal de la chaleur présente dans l'environnement. Cette diversité de sources de données pourrait enrichir les modèles, permettant ainsi une détection plus fine des zones sensibles aux îlots de chaleur.

Une autre avenue pour améliorer les performances des modèles repose sur l'optimisation de leurs architectures. L'exploration de nouvelles approches, telles que les Transformers en vision par ordinateur, pourrait offrir une compréhension plus approfondie des relations spatiales complexes dans les zones thermiques urbaines. Contrairement aux réseaux neuronaux convolutifs classiques, les Transformers sont capables de capturer des dépendances à longue portée dans les données, ce qui pourrait être particulièrement utile dans le contexte de la détection des îlots de chaleur, où les influences thermiques peuvent s'étendre sur de grandes zones urbaines. En intégrant des Transformers à ces systèmes, il serait possible d'améliorer la

précision des prédictions et de mieux gérer les interactions complexes entre les différentes structures urbaines.

De plus, l'intégration de modèles hybrides pourrait être une approche clé pour tirer parti des avantages des deux paradigmes existants, à savoir les modèles physiques et statistiques d'une part, et les méthodes de ML d'autre part. Les modèles physiques, qui simulent les phénomènes de transfert de chaleur dans les environnements urbains, peuvent fournir une base solide pour prédire les variations thermiques selon des principes de la thermodynamique. En combinant ces modèles avec des techniques de ML, comme les réseaux neuronaux profonds, il serait possible de créer des modèles plus robustes et plus précis, capables de s'adapter aux particularités des zones urbaines tout en tenant compte des phénomènes physiques sous-jacents.

Dans une optique de mise en œuvre pratique, le déploiement de solutions en temps réel représenterait un autre axe d'amélioration important. Afin de rendre ces systèmes plus accessibles et réactifs, l'optimisation des modèles pour qu'ils soient légers et capables de fonctionner sur des plateformes cloud ou mobiles serait essentielle. Cela permettrait non seulement une surveillance continue des îlots de chaleur, mais aussi une réaction rapide face à l'évolution de ces zones à risque. L'implémentation de ces solutions légères garantirait une surveillance à grande échelle, tout en étant suffisamment flexible pour s'adapter aux infrastructures informatiques limitées de certaines zones géographiques ou villes.

Enfin, il est crucial d'envisager l'intégration de ces outils de détection dans les politiques urbaines afin de favoriser une prise de décision proactive. L'élaboration d'outils interactifs à destination des urbanistes et des décideurs municipaux, permettant une visualisation intuitive des zones à risque, serait un moyen efficace de sensibiliser les acteurs clés aux enjeux liés aux îlots de chaleur. Ces outils interactifs pourraient être utilisés pour identifier les zones les plus vulnérables, prioriser les actions de renaturation urbaine, et évaluer l'impact de différentes stratégies d'adaptation au changement climatique [37]. En combinant les résultats des modèles avec des plateformes de visualisation en temps réel, les décideurs seraient mieux équipés pour prendre des décisions éclairées et orientées vers la durabilité, afin de réduire les effets négatifs des îlots de chaleur dans les environnements urbains.

5.5 Conclusion générale

En conclusion, cette étude a mis en lumière l'efficacité des techniques d'apprentissage profond pour la détection et la segmentation des îlots de chaleur urbains, un enjeu majeur dans le contexte du réchauffement climatique et de l'urbanisation croissante. À travers l'analyse détaillée des différents modèles utilisés, il a été démontré que, bien que chaque approche présente ses propres forces et avantages, il est crucial de privilégier une stratégie combinée pour obtenir des résultats plus précis et adaptés aux besoins spécifiques des décideurs. En effet, la capacité de ces modèles à identifier avec précision les zones à risque constitue un atout majeur pour la planification urbaine, en offrant une base fiable sur laquelle les urbanistes et les responsables politiques peuvent s'appuyer pour élaborer des politiques d'atténuation des effets des îlots de chaleur.

Les résultats obtenus au cours de cette étude montrent clairement que les modèles étudiés sont efficaces pour détecter les zones urbaines vulnérables aux fortes températures, fournissant ainsi des informations cruciales pour la mise en œuvre de stratégies d'adaptation et de résilience face aux défis climatiques. Grâce à ces avancées, les technologies de l'IA émergent comme des outils puissants non seulement pour la compréhension des phénomènes thermiques urbains, mais aussi pour améliorer la qualité de vie dans les villes en réduisant les impacts négatifs du réchauffement climatique.

Cependant, malgré ces progrès, plusieurs défis demeurent. Des efforts continus sont nécessaires pour perfectionner les modèles existants. En particulier, l'amélioration des jeux de données, en diversifiant les sources et en augmentant la résolution des images, est un aspect clé pour renforcer la capacité des modèles à généraliser dans des contextes variés. De même, l'optimisation des architectures des réseaux neuronaux et la réduction de leur complexité computationnelle sont des pistes à explorer pour rendre ces technologies plus accessibles et efficaces sur le terrain. Une autre direction importante pour l'avenir consiste à intégrer ces modèles dans des plateformes opérationnelles de suivi environnemental en temps réel, permettant ainsi une surveillance continue des îlots de chaleur et une gestion proactive des risques.

En somme, cette étude ouvre de nouvelles perspectives passionnantes pour la recherche et l'application de l'IA dans le domaine de l'urbanisme climatique. Les recommandations formulées au cours de ce travail offrent des pistes concrètes pour affiner les modèles de

détection des îlots de chaleur, tout en soulignant la nécessité d'une approche plus robuste et adaptable aux réalités locales. Ces progrès contribueront à la mise en place de solutions durables et adaptées aux spécificités de chaque environnement urbain, en vue de créer des villes plus résilientes face aux défis climatiques du futur.

Bibliographie

- [1] Y. Bengio, A. Courville, and P. Vincent, "Representation Learning: A Review and New Perspectives," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, pp. 1798–1828, 2013.
- [2] B. Vika et E. Hoxha, « Deep Learning for Automatic Classification of Identification Documents », Global Journal of Information Technology : Emerging Technologies, vol. 13, no. 2, pp. 90–99, 2023
- [3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems (NeurIPS), vol. 25, 2012.
- [4] K. Simonyan et A. Zisserman, « Very Deep Convolutional Networks for Large-Scale Image Recognition », International Conference on Learning Representations (ICLR), 2015. [En ligne].
- [5] C. R. de Almeida, A. C. Teodoro et A. Gonçalves, « Study of the Urban Heat Island (UHI) Using Remote Sensing Data/Techniques : A Systematic Review », Environments, vol. 8, no. 10, art. 105, oct. 2021.
- [6] O. Ronneberger, P. Fischer et T. Brox, « U-Net: Convolutional Networks for Biomedical Image Segmentation », Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015.
- [7] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy et A. L. Yuille, « DeepLab : Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs », IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 4, pp. 834–848, 2018.
- [8] R. N. Dev et M. Mehta, « A review of urban heat island mapping approaches with a special emphasis on the Indian region », Environmental Monitoring and Assessment, vol. 197, art. 365, 2025.

- [9] G. Mutani, A. Scalise, X. Sufa, et S. Grasso, « Synergising Machine Learning and Remote Sensing for Urban Heat Island Dynamics : A Comprehensive Modelling Approach », *Atmosphere*, vol. 15, no. 12, p. 1435, 2024
- [10] A. A. Tehrani, O. Veisi, K. Kia, et Y. Delavar, « Deep Learning-Based Prediction of Urban Heat Island Intensity in European Cities Using Urban Morphological Features », *Urban Climate*, vol. 56, p. 102061, 2024.
- [11] C. Sager, C. Janiesch, and P. Zschech, “A survey of image labelling for computer vision applications,” *arXiv*, Apr. 2021.
- [12] K. He, X. Zhang, S. Ren, et J. Sun, « Deep Residual Learning for Image Recognition », *arXiv preprint arXiv :1512.03385*, déc. 2015.
- [13] C. R. de Almeida, A. C. Teodoro et A. Gonçalves, « Study of the Urban Heat Island (UHI) Using Remote Sensing Data/Techniques : A Systematic Review », *Environments*, vol. 8, no. 10, p. 105, 2021
- [14] M. Tan et Q. V. Le, « EfficientNet : Rethinking Model Scaling for Convolutional Neural Networks », *arXiv preprint arXiv :1905.11946*, mai 2019
- [15] O. Y. A. Mohamed et I. Zahidi, « Artificial Intelligence for Predicting Urban Heat Island Effect and Optimising Land Use/Land Cover for Mitigation : Prospects and Recent Advancements », *Urban Climate*, vol. 55, p. 101976, 2024
- [16] I.-H. Holobacă, M. Alexe, K. Temerde-Ivan et C.-D. Ursu, « Enhanced Urban Heat Island Modeling with Machine Learning and Regression Kriging in a Topographically Diverse Medium-Sized City », *Research Square*, 2024
- [17] A. A. Tehrani, O. Veisi, K. Kia et Y. Delavar, « Deep Learning-Based Prediction of Urban Heat Island Intensity in European Cities Using Urban Morphological Features », *Urban Climate*, vol. 56, p. 102061, 2024
- [18] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh et J. Liang, « UNet++: A Nested U-Net Architecture for Medical Image Segmentation », *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, 2018, pp. 3–11

- [19] L. T. Ramos et A. D. Sappa, « Leveraging U-Net and selective feature extraction for land cover classification using remote sensing imagery », *Scientific Reports*, vol. 15, p. 784, janv. 2025, doi: 10.1038/s41598-024-84795-1
- [20] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy et A. L. Yuille, « DeepLab : Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018
- [21] K. He, X. Zhang, S. Ren, and J. Sun, “Identity mappings in deep residual networks,” in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Amsterdam, The Netherlands, 2016, vol. 9908, pp. 630–645, Springer
- [22] J. Voogt et T. R. Oke, « Thermal remote sensing of urban climates », *Remote Sensing of Environment*, vol. 86, no. 3, pp. 370–384, 2003
- [23] M. Schmitt, L. H. Hughes, C. Qiu et X. X. Zhu, « SEN12MS – A Curated Dataset of Georeferenced Multi-Spectral Sentinel-1/2 Imagery for Deep Learning and Data Fusion », *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. IV-2/W7, pp. 153–160, 2019
- [24] M. Santamouris, « Heat island research in Europe : The state of the art », *Advances in Building Energy Research*, vol. 1, no. 1, pp. 123–150, 2007
- [25] K. L. Ebi et K. McGregor, « Climate change, tropospheric ozone and particulate matter, and health impacts », *Environmental Health Perspectives*, vol. 114, no. 9, pp. 1455–1460, 2006
- [26] A. Kalla et R. Erriri, « Système d’authentification du locuteur basé sur un Réseau neuronal résiduel (ResNet) », 2024
- [27] A. A. Tehrani, O. Veisi, K. Kia et Y. Delavar, « Deep Learning-Based Prediction of Urban Heat Island Intensity in European Cities Using Urban Morphological Features », *Urban Climate*, vol. 56, p. 102061, 2024

- [28] O. Y. A. Mohamed et I. Zahidi, « Artificial intelligence for predicting urban heat island effect and optimising land use/land cover for mitigation : Prospects and recent advancements », *Urban Climate*, vol. 55, p. 101976, 2024.
- [29] G. Tanoori, A. Soltani et A. Modiri, « Predicting Urban Land Use and Mitigating Land Surface Temperature : Exploring the Role of Urban Configuration with Convolutional Neural Networks », *ResearchGate*, 2024
- [30] O. Ronneberger, P. Fischer et T. Brox, « U-Net : Convolutional Networks for Biomedical Image Segmentation », 2015
- [31] M. Ganjirad, M. R. Delavar, H. Bagheri et M. M. Azizi, « Optimizing Urban Critical Green Space Development Using Machine Learning », *Sustainable Cities and Society*, vol. 105, p. 104563, 2025.
- [32] A. A. Tehrani, O. Veisi, K. Kia et Y. Delavar, « Deep Learning-Based Prediction of Urban Heat Island Intensity in European Cities Using Urban Morphological Features », *Urban Climate*, vol. 56, p. 102061, 2024
- [33] O. Y. A. Mohamed et I. Zahidi, « Artificial intelligence for predicting urban heat island effect and optimising land use/land cover for mitigation : Prospects and recent advancements », *Urban Climate*, vol. 55, p. 101976, 2024
- [34] M. Schmitt, L. H. Hughes, C. Qiu et X. X. Zhu, « SEN12MS – A Curated Dataset of Georeferenced Multi-Spectral Sentinel-1/2 Imagery for Deep Learning and Data Fusion », *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. IV-2/W7, pp. 153–160, 2019
- [35] R. Liu, Y. Zhao, Y. Xie et S. Wang, « Deep learning-based framework for urban heat island detection using multi-source data », *Remote Sensing*, vol. 13, no. 4, 2021.
- [36] K. Kustura, D. Conti, M. Sammer et M. Riffler, « Harnessing Multi-Source Data and Deep Learning for High-Resolution Land Surface Temperature Gap-Filling Supporting Climate Change Adaptation Activities », *Remote Sensing*, vol. 17, no. 2, p. 318, 2025.

- [37] I. Roussel, « Chapitre 3. La santé revisitée par le changement climatique : Climat et santé : une histoire d'adaptation ? », in *Le changement climatique : Quand le climat nous pousse à changer d'ère*, H.-J. Scarwell, Éd., Villeneuve d'Ascq : Presses universitaires du Septentrion, 2010, pp. 121–155
- [38] M. Schmitt, L. H. Hughes, C. Qiu et X. X. Zhu, « SEN12MS – A Curated Dataset of Georeferenced Multi-Spectral Sentinel-1/2 Imagery for Deep Learning and Data Fusion », *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. IV-2/W7, pp. 153–160, 2019
- [39] K. Kustura, D. Conti, M. Sammer et M. Riffler, « Harnessing Multi-Source Data and Deep Learning for High-Resolution Land Surface Temperature Gap-Filling Supporting Climate Change Adaptation Activities », *Remote Sensing*, vol. 17, no. 2, p. 318, 2025
- [40] M. Schmitt, L. H. Hughes, C. Qiu et X. X. Zhu, « SEN12MS – A Curated Dataset of Georeferenced Multi-Spectral Sentinel-1/2 Imagery for Deep Learning and Data Fusion », *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. IV-2/W7, pp. 153–160, 2019
- [41] M. Kolokotroni, X. Ren, M. Davies et Z. Mavrogianni, « London's urban heat island : Impact on current and future energy consumption in office buildings », *Energy and Buildings*, vol. 47, pp. 302–311, 2012
- [42] T. R. Oke, « The energetic basis of the urban heat island », *Quarterly Journal of the Royal Meteorological Society*, vol. 108, no. 455, pp. 1–24, 1982
- [43] A. Kalla et R. Erriri, « Système d'authentification du locuteur basé sur un Réseau neuronal résiduel (ResNet) », 2024
- [44] O. Ronneberger, P. Fischer et T. Brox, « U-Net : Convolutional Networks for Biomedical Image Segmentation », *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Lecture Notes in Computer Science*, vol. 9351, pp. 234–241, 2015
- [45] V. Badrinarayanan, A. Kendall et R. Cipolla, « SegNet : A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation », *IEEE Trans. Pattern Anal. Mach. Intell.*, 2017

- [46] T.-Y. Lin, P. Goyal, R. Girshick, K. He, et P. Dollár, « Focal Loss for Dense Object Detection », *ICCV*, 2017
- [47] Q. Zhao et al., « Object Detection with Deep Learning : A Review », *IEEE Trans. Neural Netw. Learn. Syst.*, 2020
- [48] B. Zhou, H. Zhao, X. Puig, S. Fidler, A. Barriuso et A. Torralba, « Scene Parsing through ADE20K Dataset », *CVPR*, 2017
- [49] E. Xie, W. Wang, Z. Yu, et al., « Segment Anything You Want with SAM », *arXiv*, 2023.
- [50] N. Gorelick et al., « Google Earth Engine : Planetary-scale geospatial analysis for everyone », *Remote Sensing of Environment*, 2017
- [51] G. Cheng et al., « A Survey of Remote Sensing Image Classification : From Traditional to Deep Learning », *Remote Sensing*, 2020
- [52] Y. LeCun, L. Bottou, Y. Bengio et P. Haffner, « Gradient-Based Learning Applied to Document Recognition », *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998
- [53] A. Krizhevsky, I. Sutskever, and G. E. Hinton, « ImageNet Classification with Deep Convolutional Neural Networks », *NeurIPS*, vol. 25, 2012
- [54] K. Simonyan and A. Zisserman, « Very Deep Convolutional Networks for Large-Scale Image Recognition », *arXiv preprint arXiv :1409.1556*, 2014
- [55] O. Ronneberger, P. Fischer and T. Brox, "U-Net : Convolutional Networks for Biomedical Image Segmentation," *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015
- [56] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy and A. L. Yuille, "DeepLab : Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," *IEEE Trans. on PAMI*, vol. 40, no. 4, pp. 834–848, Apr. 2018
- [57] L. Perez and J. Wang, "The Effectiveness of Data Augmentation in Image Classification using Deep Learning," *arXiv preprint arXiv :1712.04621*, 2017

- [58] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, "Dropout : A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014
- [59] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once : Unified, Real-Time Object Detection," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788
- [60] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection," in *Proc. of the IEEE Int. Conf. on Computer Vision (ICCV)*, 2017, pp. 2980–2988.
- [61] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN : Towards Real-Time Object Detection with Region Proposal Networks," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, 2017
- [62] O. Ronneberger, P. Fischer, and T. Brox, "U-Net : Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015, pp. 234–241
- [63] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet : A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017
- [64] F. Salamanca, A. Martilli, M. Tewari, and F. Chen, "A study of the urban boundary layer using different urban parameterizations and high-resolution urban canopy parameters with WRF," *Journal of Applied Meteorology and Climatology*, vol. 50, no. 5, pp. 1107–1128, 2011.
- [65] T. R. Oke, "The energetic basis of the urban heat island," *Quarterly Journal of the Royal Meteorological Society*, vol. 108, no. 455, pp. 1–24, 1982.
- [66] D. Jenerette, S. Harlan, W. Stefanov, and L. Brazel, "Variability in vegetation effects on urban microclimate and heat-related mortality," *Landscape and Urban Planning*, vol. 94, no. 3–4, pp. 144–155, 2010.
- [67] M. Santamouris, "Heat island research in Europe : The state of the art," *Advances in Building Energy Research*, vol. 1, no. 1, pp. 123–150, 2007.

- [68] K. L. Ebi and K. McGregor, "Climate change, tropospheric ozone and particulate matter, and health impacts," *Environmental Health Perspectives*, vol. 114, no. 9, pp. 1455–1460, 2006.
- [69] M. Kolokotroni, X. Ren, M. Davies, and Z. Mavrogianni, "London's urban heat island : Impact on current and future energy consumption in office buildings," *Energy and Buildings*, vol. 47, pp. 302–311, 2012.
- [70] A. Forman, R. Davies, and M. Burch, "Urban vegetation and climate adaptation : How green infrastructure can support urban biodiversity," *Urban Forestry & Urban Greening*, vol. 21, pp. 25–35, 2017.
- [71] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [72] W. Ma, Y. Shao, M. Zhang, and X. Xue, "Urban heat island monitoring and analysis using deep learning and satellite remote sensing data : A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 186, pp. 362–376, 2022.
- [73] M. Schmitt, L. H. Hughes, C. Qiu, and X. X. Zhu, "SEN12MS – A Curated Dataset of Georeferenced Multi-Spectral Sentinel-1/2 Imagery for Deep Learning and Data Fusion," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. IV-2/W7, pp. 153–160, 2019.
- [74] A. Khosravi et al., "A hybrid deep learning framework for urban heat island detection using ResNet and U-Net architectures," *Urban Climate*, vol. 46, 101272, 2023.
- [75] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++ : A Nested U-Net Architecture for Medical Image Segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, 2018, pp. 3–11.
- [76] R. Liu, Y. Zhao, Y. Xie, and S. Wang, "Deep learning-based framework for urban heat island detection using multi-source data," *Remote Sensing*, vol. 13, no. 4, 2021.
- [77] J. Dietterich, "Ensemble Methods in Machine Learning," in *International Workshop on Multiple Classifier Systems*, 2000, pp. 1–15.

- [78] F. Zheng, X. Wang, and W. Liu, "Fusion of multi-source remote sensing data for urban thermal environment monitoring using deep learning," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 181, pp. 90–104, 2021.
- [79] C. Wan, D. Zhang, and T. Li, "Multispectral remote sensing image classification based on improved deep residual network," *Applied Sciences*, vol. 9, no. 5, 2019.
- [80] M. Caron, P. Bojanowski, A. Joulin, and M. Douze, "Deep Clustering for Unsupervised Learning of Visual Features," in *Proc. of the European Conf. on Computer Vision (ECCV)*, 2018, pp. 132–149.
- [81] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [82] M. Schmitt, L. H. Hughes, C. Qiu, and X. X. Zhu, "SEN12MS – A Curated Dataset of Georeferenced Multi-Spectral Sentinel-1/2 Imagery for Deep Learning and Data Fusion," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. IV-2/W7, pp. 153–160, 2019.
- [83] R. Liu, Y. Zhao, Y. Xie, and S. Wang, "Deep learning-based framework for urban heat island detection using multi-source data," *Remote Sensing*, vol. 13, no. 4, 2021.
- [84] J. Voogt and T. R. Oke, "Thermal remote sensing of urban climates," *Remote Sensing of Environment*, vol. 86, no. 3, pp. 370–384, 2003.
- [85] F. Zheng, X. Wang, and W. Liu, "Fusion of multi-source remote sensing data for urban thermal environment monitoring using deep learning," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 181, pp. 90–104, 2021.
- [86] P. K. Shukla, D. M. Hoang, R. N. Dubey, and A. Sangaiah, "Recent advances in deep learning for object detection," *Computer Science Review*, vol. 38, 100310, 2020.
- [87] & [91] A. Paszke et al., "PyTorch : An Imperative Style, High-Performance Deep Learning Library," *NeurIPS* 2019.
- [88] M. Bisong, *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, Apress, 2019.
- [89] A. Sekachev et al., "CVAT : A Universal Approach to Data Annotation," *Zenodo*, 2023.

- [90] ZEISS, “Apeer Evolves into arivis,” arivis Blog, 2023.
- [92] C.-Y. Wang, A. Bochkovskiy, H.-Y. M. Liao, “You Only Learn One Representation,” *Computer Vision and Image Understanding*, 2023.
- [93] M. K. Ranjan et al., “Python : Empowering Data Science Applications and Research,” *J. of Operating Systems Development & Trends*, 2023.
- [94] M. Abadi et al., “TensorFlow : A System for Large-Scale Machine Learning,” *OSDI* 2016.
- [95] A. Gulli, S. Pal, *Deep Learning with Keras*, Packt Publishing, 2017
- [96] S. Van der Walt, S. Colbert, G. Varoquaux, “The NumPy Array : A Structure for Efficient Numerical Computation,” *Computing in Science & Engineering*, 2011
- [97] W. McKinney, “Data Structures for Statistical Computing in Python,” *Python in Science Conf. (SciPy)*, 2010
- [98] J. D. Hunter, “Matplotlib : A 2D Graphics Environment,” *Computing in Science & Engineering*, 2007
- [99] M. Waskom et al., “Seaborn : Statistical Data Visualization,” *Journal of Open Source Software*, vol. 2, no. 11, p. 302, 2017
- [100] A. Mahendran and A. Vedaldi, “Understanding Deep Image Representations by Inverting Them,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, 2015, pp. 5188–5196, IEEE.
- [101] Aakanksha, Arushi Seth, Shanu Sharma. "Semantic Segmentation : A Systematic Analysis From State-of-the-Art Techniques to Advance Deep Networks." *Journal of Information Technology Research (JITR)*, vol. 15, no. 1, 2022, pp. 28
- [102] Swathi, Y., & Challa, M. (2024). *YOLOv8 : Advancements and Innovations in Object Detection*. In T. Senjyu, C. So-In, & A. Joshi (Eds.), *Smart Trends in Computing and Communications. SmartCom 2024* (Vol. 946, pp. 1–12). Springer, Singapore.
- [103] Y. Swathi and M. Challa, “YOLOv8 : Advancements and Innovations in Object Detection,” in *Smart Trends in Computing and Communications. SmartCom 2024*, T. Senjyu, C. So-In, and A. Joshi, Eds. Singapore : Springer, 2024, vol. 946, pp. 1–12

