



HAL
open science

Évaluation de techniques d'apprentissage profond pour la cartographie par télédétection multi-capteur de l'occupation et l'usage des sols en milieu rural au centre et au nord du Bénin

Sabi Bruno Bio Nikki Sare

► To cite this version:

Sabi Bruno Bio Nikki Sare. Évaluation de techniques d'apprentissage profond pour la cartographie par télédétection multi-capteur de l'occupation et l'usage des sols en milieu rural au centre et au nord du Bénin. Sciences de l'environnement. 2023. dumas-04510809

HAL Id: dumas-04510809

<https://dumas.ccsd.cnrs.fr/dumas-04510809v1>

Submitted on 19 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

AgroParisTech 

université
PARIS-SACLAY

 cirad

**Évaluation de techniques d'apprentissage profond pour la
cartographie par télédétection multi-capteur de l'occupation et
l'usage des sols en milieu rural au centre et au nord du Bénin**

Sabi Bruno BIO NIKKI SARE

MASTÈRE SPÉCIALISÉ
SYSTÈMES D'INFORMATION LOCALISÉS POUR L'AMÉNAGEMENT DES TERRITOIRES (SILAT)
MANAGER DE PROJETS GÉOMATIQUES

SILAT

Tuteur(s) université :
Roberto INTERDONATO

Tuteur(s) entreprise :
Raffaele GAETANO

Yvon C. HOUNTONJJI



Octobre 2023



L'établissement AgroParisTech diffuse ce mémoire et se dégage de toute responsabilité sur le contenu proposé.

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition.

L'obligation de citation et de référencement lors de l'utilisation de ce document doit être respectée selon la licence Creative Commons apposée

Ce document est soumis aux conditions d'utilisation "Paternité – Pas d'utilisation commerciale – Pas de modification 4.0 France disponible en ligne : <https://creativecommons.org/licenses/by-nc-nd/4.0/>



Résumé

L'utilisation des séries temporelles d'images satellites (*SITS*) s'est avérée indispensable pour une cartographie précise et à jour de l'occupation des sols sur de vastes étendues. Bien que la plupart des études portant sur les *SITS* se soient concentrées sur l'emploi d'algorithmes de classification traditionnels tels que les *Random Forests (RF)*, les algorithmes d'apprentissage profond ont connu un succès remarquable pour les tâches supervisées, notamment dans le cas de données présentant des structures entre les attributs, telles que l'espace ou le temps. Dans cette optique, cette étude propose une analyse comparative approfondie de quatre modèles de classification de l'occupation des sols. Pour ce faire, nous avons fait usage d'une série temporelle d'images Sentinel-2 sur l'année 2022 (400 variables, soit 16 bandes × 25 estampilles temporelles) et une base de données vérité terrain. Les résultats obtenus attestent de la bonne capacité des réseaux de neurones convolutifs à saisir des motifs complexes et à produire des prédictions en étroite concordance avec les observations réelles, démontrant ainsi leur pertinence pour les tâches complexes de classification basées sur les *SITS* surtout dans le contexte spécifique de notre zone d'étude caractérisée par un paysage assez fragmenté, la petite taille des exploitations agricoles et l'incomplétude de la série temporelle d'images (fortes nébulosités) dû à la présence de nuages. De plus, il ressort que le modèle *TempCNN* qui exploite principalement les corrélations dans le temps à été meilleur par rapport aux approches intégrant une description du contexte spatial tels que *DuPLO* et *CNN2D*.

Mots clés : *séries temporelles d'images satellites, cartographie de l'occupation du sol, Sentinel-2, deep learning, Bénin.*

Abstract

The use of Satellite Image Time Series (SITS) is crucial for accurate and current land-use mapping across extensive regions. While many SITS studies have concentrated on conventional classification algorithms like Random Forests (RF), deep learning algorithms have demonstrated notable success, particularly in supervised tasks involving data structures such as spatial or temporal attributes. This research conducts a thorough comparative analysis of four land-use classification models. The dataset encompasses Sentinel-2 image time series for the year 2022, comprising 400 variables (16 bands × 25 time stamps) and a ground truth database. The results highlight the effectiveness of convolutional neural networks (CNNs) in capturing intricate patterns and generating predictions that align closely with real observations. This underscores their relevance for complex classification tasks based on SITS, particularly in our study area characterized by a fragmented landscape, small agricultural holdings, and incomplete temporal image series due to substantial cloud cover. Moreover, the TempCNN model, which primarily leverages temporal correlations, outperformed approaches that incorporate a spatial context description, such as DuPLO and CNN2D.

Key words : *satellite image time series, land cover mapping, Sentinel-2, deep learning, Benin.*

Table des matières

1	Introduction	4
2	Étude du projet	5
2.1	Présentation et cadrage du projet	5
2.2	Lettre de mission	6
2.3	Bénéficiaires	8
2.4	Modalités d'organisation du projet	8
2.5	Analyse de Risques et Opportunités	9
2.6	Planning Directeur	11
3	État de l'art	12
3.1	Généralités sur l'intelligence artificielle (IA)	12
3.2	<i>Machine learning et Deep learning</i>	12
3.3	Apprentissage / entraînement des modèles	13
3.4	Métriques d'évaluation des modèles	14
3.5	Apprentissage automatique et classification d'images satellites	15
4	Approche méthodologique et mise en oeuvre du projet	17
4.1	Présentation du site de l'étude	17
4.2	Méthodologie générale	18
4.3	Collecte et préparation des données	19
4.3.1	Collecte des données terrains et post-traitements	19
4.3.2	Acquisition des images satellites et mise en forme des données	21
4.4	Échantillonnage des données, apprentissage et évaluation des modèles	23
4.5	Présentation, paramétrage et implémentation des techniques	23
4.5.1	Présentation et paramétrage de la baseline <i>Random Forest</i>	24
4.5.2	Présentation et paramétrage de <i>TempCNN</i>	24
4.5.3	Présentation et paramétrage de <i>CNN2D</i>	25
4.5.4	Présentation et paramétrage de <i>DuPLO</i>	26
4.6	Production des cartes	27
4.7	Matériels et outils techniques utilisés	27
5	Résultats obtenus et Discussions	29
5.1	Test de performances des techniques implémentées	29

5.2	Étude comparative des modèles testés	33
5.2.1	Comparaison des Métriques entre les Modèles évalués	33
5.2.2	Analyse comparative par classe d'occupation du sol	34
5.2.3	Analyse visuelle des cartes d'occupation du sol produites	35
5.3	Discussions et recommandations	36
5.3.1	Discussions	36
5.3.2	Recommandations et proposition d'orientations techniques	37
6	Conclusion et perspectives	38
7	Remerciements	38
8	Références bibliographiques	39
9	Annexes	41
9.1	Annexe 1 : carte produite avec la <i>baseline Random Forest</i>	41
9.2	Annexe 2 : carte produite avec le modèle <i>TempCNN</i>	42
9.3	Annexe 3 : carte produite avec le modèle <i>CNN2D</i>	43
9.4	Annexe 4 : carte produite avec le modèle <i>DuPLO</i>	44

1 Introduction

Les systèmes de production agricole du Centre et du Nord du Bénin ont subi des transformations rapides et significatives au cours des dernières décennies. Ces régions ont été profondément affectées par le développement croissant des cultures de coton, de soja et de noix de cajou, ce qui a entraîné une expansion considérable des surfaces agricoles. Cependant, cette expansion a également engendré des changements profonds dans les pratiques agricoles et pastorales, ainsi que dans les écosystèmes et la structure des paysages.

L'une des avancées majeures dans cette ère de transformation est la disponibilité croissante de données satellitaires précises et accessibles en Afrique de l'Ouest. Ainsi, cette abondance de données combinée aux progrès réalisés dans le domaine de la télédétection multi-capteur et des techniques d'apprentissage automatique offre des opportunités de développer de nouvelles applications dans les domaines de l'agriculture et de l'environnement afin de caractériser et de suivre l'évolution des paysages en relation avec les pratiques agricoles.

C'est dans ce contexte que s'inscrit le projet OBSYDYA (Observatoire Pilote des Paysages et Dynamiques Agricoles du Bénin), qui vise à exploiter les images satellites disponibles pour développer de nouvelles applications dans les domaines de l'agriculture et de l'environnement. Dans le cadre de ce projet, le présent mémoire de fin d'étude du Mastère Spécialisé SILAT (Systèmes d'Information Localisés pour l'Aménagement des Territoires) se concentre spécifiquement sur l'évaluation des techniques d'apprentissage automatique, en particulier le *deep learning* (apprentissage profond), pour la cartographie de l'occupation et de l'usage des sols au centre et au nord du Bénin.

L'objectif principal de cette étude est de tester les performances des méthodes de machine learning les plus innovantes, notamment celles issues de la littérature récente en *deep learning*, pour la classification de l'occupation et de l'usage des sols à partir de séries temporelles d'images satellitaires. L'étude explorera les capacités de ces techniques avancées afin de fournir une cartographie précise et détaillée des différentes classes d'occupation des sols dans une sélection de sites d'étude situés dans les régions du Centre et du Nord du Bénin.

Ce mémoire sera structuré en différentes sections. La première section présentera l'étude du présent projet de fin d'étude, ses objectifs, ses enjeux et le contexte dans lequel il s'inscrit. Ensuite, l'état de l'art examinera brièvement les travaux de recherche existants sur la classification de l'occupation des sols avec le *deep learning*. La troisième section abordera l'approche méthodologique et la mise en œuvre concrète du projet, détaillant les choix techniques, la collecte des données satellitaires, les pré-traitements et les méthodes de classification utilisées. Les résultats obtenus à partir de l'application de ces techniques seront présentés dans la quatrième section. Enfin, les discussions en cinquième section permettront d'analyser et d'interpréter ces résultats.

2 Étude du projet

2.1 Présentation et cadrage du projet

Le présent projet de fin d'étude du Mastère Spécialisé SILAT (MS SILAT) porte sur l'évaluation de techniques d'apprentissage profond pour la cartographie par télédétection multi-capteur de l'occupation et l'usage des sols en milieu rural au centre et au nord Bénin. Ce projet fait partie intégrante du Résultat 1 / Thèse 1 d'un grand projet dénommé OBSYDYA - BENIN (Observatoire Pilote des Paysages et des Dynamiques Agricoles du Bénin) comme le montre la figure 2 ci après :

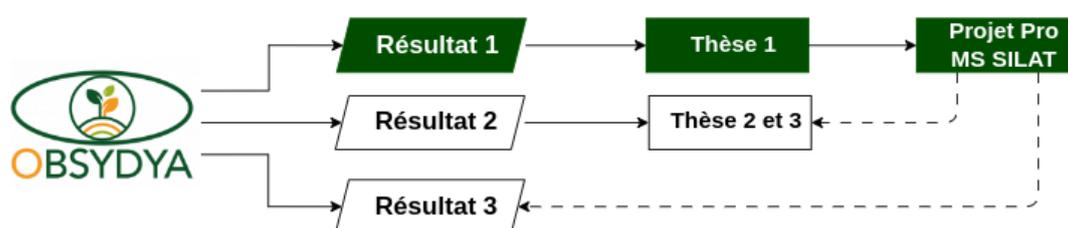


FIGURE 2 – Situation du projet MS SILAT par rapport au grand projet OBSYDYA

En effet, les objectifs du projet OBSYDYA peuvent être divisés en deux sous-objectifs complémentaires : l'objectif scientifique consiste à explorer et à opérationnaliser les possibilités offertes par l'offre actuelle d'images satellites pour produire de nouveaux indicateurs (réguliers et peu coûteux) et des cartes de la dynamique des systèmes agraires et des paysages du Bénin au niveau central et septentrional ; l'objectif opérationnel est de mettre en place un observatoire pilote des systèmes agraires et des paysages, sous la forme d'un collectif de personnes partageant une plateforme de services dédiée à la capitalisation des informations spécialisées provenant principalement de données satellitaires. Ladite plateforme vise à renforcer les capacités des acteurs agricoles nationaux et locaux dans la planification des infrastructures, les conseils agricoles et la coordination au sein des chaînes de valeur. Le projet OBSYDYA vise trois résultats majeurs :

- R1 : une base de données cartographique des paysages, de l'usage des sols et de leurs évolutions est constituée à l'échelle de toute la zone Nord et Centre du Bénin et à l'échelle de zones représentatives des systèmes agraires en évolution, puis des chaînes de traitements automatisés permettant sa mise à jour régulière seront proposées.
- R2 : des indicateurs permettant de caractériser les systèmes agraires et les paysages ainsi que leurs dynamiques et leur durabilité environnementale, sont co-construits avec les acteurs locaux sur des paysages-types représentatifs des principales dynamiques en cours (développement du coton, développement des plantations arborées, évolution du couvert forestier, fronts pionniers, etc.). Ces indicateurs font l'objet de tests méthodologiques afin de les rendre opérationnels.

- R3 : une plateforme informatique pilote (serveur de données + outils de traitement + services informationnels) de traitement, partage et diffusion de l'information spatialisée est mise en place.

Comme on peut le constater sur la figure 2, notre projet professionnel MS SILAT participe de façon générale à l'atteinte des objectifs du Projet OBSYDYA. Plus précisément, ce projet fournira directement des résultats pour la thèse 1 et indirectement aux résultats 2 et 3.

2.2 Lettre de mission

Dans cette sous partie, nous faisons une présentation complète de notre projet de fin d'étude à travers la lettre de mission. Cette lettre de mission décrit en quelques lignes le contexte et le périmètre de la mission que l'auditeur s'engage à réaliser et sur laquelle le commanditaire s'engage à concentrer l'activité de l'auditeur (objectifs, livrables attendus).

Intitulé du projet : Projet OBSYDYA - Évaluation de techniques d'apprentissage profond pour la cartographie par télédétection multi-capteur de l'occupation et l'usage des sols en milieu rural au centre et au nord Bénin.

Échéance du projet : 06 avril au 31 septembre 2023

Équipe projet :

TABLE 1 – Équipe projet

Commanditaire	Tuteur	Chef du Projet
GAETANO Raffaele CIRAD	INTERDONATO Roberto CIRAD	BIO NIKKI Bruno MS - SILAT

Contexte scientifique

Financé par l'Union Européenne dans le cadre de l'initiative *DeSIRA (Development Smart Innovation through Research in Agriculture)*, le projet OBSYDYA Bénin (Observatoire pilote des paysages et dynamiques agricoles du Bénin) a comme objectif la spécification et le prototypage d'un observatoire des systèmes et paysages agricoles du Nord et Centre Bénin, sous forme d'une plateforme de services dédiés à la capitalisation d'information spatiale dérivée principalement d'imagerie satellite. Cette plateforme vise à renforcer les capacités des acteurs en lien avec les activités agricoles du pays à l'échelle locale et nationale dans la planification d'infrastructures, le conseil agricole et la coordination des filières de production. Pour offrir des réponses à cet enjeu opérationnel, l'objectif scientifique est d'explorer et opérationnaliser les possibilités mises à disposition par l'offre actuelle en imagerie satellite pour la production de cartographies des systèmes agricoles du Nord et Centre Bénin, ainsi que des nouveaux indicateurs spatiaux, de façon régulière et à moindre coût. Une composante du projet est en particulier dédiée à l'identification des outils et chaînes de traitement pour la production de cartes de zonage agro-systémique de la zone d'étude,

ainsi que de cartes d'occupation du sol et usage des sols agricoles aux échelles nationale et globale, à partir de données de télédétection multi-capteur (optique à haute et très haute résolution, radar). Suivant une démarche par zones pilotes, sur lesquelles un recueil de données de terrain est prévu annuellement pendant la durée du projet (jusqu'en 2025), la composante s'occupera, d'un côté, de fournir aux autres composantes des produits « frais » obtenus par les techniques de l'état de l'art (comme les chaînes *iota2* [1] et *Moringa* [*moringa*]), ainsi que, dans une posture plus prospective, tester et développer des nouvelles techniques issues de l'intelligence artificielle pour améliorer la précision et la profondeur descriptive de ces produits (par ex., en raffinant les nomenclatures d'usage agricole des terres, notamment par rapport à des filières spécifiques – anacardier, manguiers, coton,...).

Objectif et finalités

L'objectif de ce stage est, en général, de dresser un bilan des performances et des limites des techniques basées sur l'apprentissage profond pour la classification de l'occupation et l'usage des sols, de parution récente dans l'état de l'art, dans la zone d'étude du projet OBSYDYA. L'attention sera portée en particulier sur les techniques pouvant intégrer de l'imagerie multi-capteur (optiques HR – Sentinel-2 - et THR - SPOT6/7, radar – Sentinel-1).

Activités prévues

- une revue bibliographique des techniques d'apprentissage profond pour la classification de l'occupation du sol, avec une attention particulière aux méthodes basées sur les séries temporelles d'images HR (ex. TempCNN [2], DuPLO [3]) et sur le couplage de séries temporelles et imagerie à très haute résolution spatiale (ex. M3Fusion [4]);
- la sélection et l'implémentation en Python (à l'aide de la librairie PyTorch pour l'apprentissage profond) de certaines de ces techniques;
- la préparation et le prétraitement d'un jeu de données d'images multi-capteur sur deux des zones pilotes du projet (entre la ville de Parakou et la forêt de N'dali);
- la mise en forme d'une base de données de référence pour l'apprentissage et la validation, basée sur les données de terrain collectées dans le cadre du projet OBSYDYA;
- l'apprentissage des modèles et la quantification de leurs performances à l'aide de la BD construite;
- la rédaction d'un rapport scientifique dans lequel des pistes d'améliorations méthodologiques sont préconisées.

Livrables

- un rapport scientifique dans lequel des pistes d'améliorations méthodologiques sont préconisées;
- des scripts python, modèles d'apprentissage, chaînes de traitement et base de données implémentés au cours de l'étude pour l'atteinte des résultats obtenus.

2.3 Bénéficiaires

Les différents bénéficiaires du projet et l'intérêt pour chacun se présentent comme suit :

- **Équipe projet OBSYDYA Bénin** : contribution à l'atteinte du Résultat 1 du projet et des thèses ;
- **Institutions de recherches au Bénin** : Mise à disposition de Modèles et Base de données de référence ;
- **MAEP, ATDA, PDA, Collectivités territoriales et groupement paysans** : Mise à disposition de Base de données de référence et cartes thématiques ;
- **L'auditeur SILAT** : Validation du diplôme, Initiation du travail de la thèse.

2.4 Modalités d'organisation du projet

Dans le but d'assurer le bon déroulement du projet et surtout pour sa mise en oeuvre, nous avons défini les instances opérationnelles et stratégiques du projet comme suit :

Comité de projet :

- Chef projet : coordonner, organiser et assurer la mise en oeuvre des activités du projet ;
- Commanditaire : superviser et valider les résultats, apporter un appui technique pour l'atteinte des objectifs.

Comité de projet élargi :

- Équipe du projet OBSYDYA – BÉNIN : apprécier les résultats obtenus
- Université de Parakou et autres partenaires : apprécier les résultats obtenus

Groupes de travail :

- Groupe 1 : Acquisition, traitement des données et implémentation des techniques d'apprentissage automatique retenues : Bruno et Raffaele ;
- Groupe 2 : Validation et appréciation des résultats obtenus : Yvon , Raffaele et autres membres du Projet OBSYDYA – BENIN.

Nous avons également opté pour une décomposition de notre projet en cinq (05) grands lots comme le montre le tableau ci - dessous.

TABLE 2 – Décomposition du projet en lots

N°	Lots	Responsables
Lot 1	Acquisition des images satellites – THRS et collecte des données sur le terrain	Bruno et Raffaele
Lot 2	Pré-traitement des données collectées et mise au point d'un script de préparation des données.	Bruno et Raffaele
Lot 3	Identification des techniques, implémentions, tests	Bruno et Raffaele
Lot 4	Appréciation les résultats obtenus et recommandations	Raffaele et Yvon
Lot 5	Rédaction du rapport de fin d'étude et du rapport technique	Bruno

2.5 Analyse de Risques et Opportunités

Le Tableau 3 montre une synthèse de l'analyse des risques et opportunités du projet.

TABLE 3 – Analyse des risques et opportunités

N° R/O	Famille	Description	Gravité	Probabilité	Criticité
R1	Technique	Fortes nébulosité sur les images disponibles pour la constitution du jeu de données	Peu grave	Probable	Important
R2	Technique	Mauvaise qualité des données collectées sur le terrain	Peu grave	Peu Probable	Faible
R3	Management	Difficulté de respect des délais au vu de la complexité des techniques ML	grave	Peu Probable	Important
O1	Management	Disponibilité d'une équipe pluridisciplinaire pour le projet	-	Très Probable	-
O1	Technique	Couverture de la zone d'étude et disponibilité des données	-	Très Probable	-
O3	Technique	Disponibilité de la documentation et des détails sur les techniques ML à tester	-	Très Probable	-

Pour chacun des risques et opportunités énumérés précédemment, nous avons également identifié les méthodes de traitement et de gestion à travers des actions de maîtrise comme le montre le Tableau 4 ci dessous :

TABLE 4 – Traitement et gestion des risques (R) et Opportunités (O)

N° R/O	Action de maîtrise	Pilote	Date de levée R/O	Coût(ou gain)	Fondement du coût / gain
R1	Identifier plusieurs sources de données exploitables	Bruno et Raffaele	Dès le démarrage	Fiabilité	Analyse de la disponibilité
R2	Surveiller correctement la campagne terrain / faire des réajustements le cas échéant	Équipe projet	Lors et après les travaux de terrain	Fiabilité	Disponibilité de l'équipe projet
R3	Faire une bonne revue de littérature et prendre les techniques les plus pertinentes	Bruno et Raffaele	Début de la mission	Faisabilité	Les techniques ont été identifiées depuis
O1	Échanges et séances régulières avec l'équipe	Bruno	Début de la mission	Évolution	-
O1	Multiplication de la source de données	Bruno et Raffaele	Début de la mission	Temps	-
O3	Documentation régulière sur les techniques	Bruno	Début de la mission	Temps	-

2.6 Planning Directeur

Activités / Mission	Avril				Mai				Juin				Juillet				Août				Septembre			
	S1	S2	S3	S4	S1	S2	S3	S4	S1	S2	S3	S4	S1	S2	S3	S4	S1	S2	S3	S4	S1	S2	S3	S4
Revue de littérature et balisage méthodologique																								
Revue bibliographique des techniques d'apprentissage profond pour la classification de l'occupation du sol																								
Sélection des techniques et choix méthodologiques																								
Collecte, préparation et prétraitement des données																								
Collecte des images satellitaires (S2 - THRS)																								
Préparation et prétraitement des données (script Python + OTB)																								
Organisation et mise au point d'un jeu de données exploitable.																								
Méthodes d'apprentissage automatique, implémentation et test des techniques retenues																								
Préparation de l'environnement de travail pour l'implémentation des méthodes et appropriation aux outils de travail																								
Récupération de la base de données de référence issue de la mission terrain sur le site d'étude du projet OBSYDYA																								
Mise en forme d'une base de données de référence pour l'apprentissage et la validation, basée sur les données de terrain collectées dans le cadre du projet OBSYDYA																								
Implémentation des techniques retenues en Python																								
Apprentissage des modèles et quantification de leurs performances à l'aide de la BD construite																								
Évaluation et validation des résultats obtenus																								
Rédaction des rapports de fin d'étude																								
Rédiger un rapport scientifique dans lequel des pistes d'amélioration méthodologiques sont préconisées																								
Rédiger un rapport d'étude et de la thèse professionnelle du MS SILAT																								
Soumission des rapports au commanditaire																								
Prise en compte des observations																								
Dépôt des rapports finaux																								
Réunions du projet OBSYDYA / Séance avec le commanditaire																								

FIGURE 3 – Planification du projet

3 État de l'art

3.1 Généralités sur l'intelligence artificielle (IA)

L'intelligence artificielle (IA) est un domaine de recherche multidisciplinaire qui vise à développer des systèmes informatiques capables d'imiter certaines capacités intellectuelles humaines. Selon RUSSELL et NORVIG [5], l'IA englobe des domaines tels que l'apprentissage automatique, la vision par ordinateur, le traitement du langage naturel et la robotique.

3.2 *Machine learning et Deep learning*

L'apprentissage automatique (*machine learning*) et l'apprentissage profond (*deep learning*) sont des domaines de l'intelligence artificielle qui ont révolutionné de nombreux aspects de la science des données et de la résolution de problèmes complexes.

L'**apprentissage automatique** consiste à développer des modèles et des algorithmes qui permettent aux machines d'apprendre à partir de données et de prendre des décisions ou de faire des prédictions sans être explicitement programmées. Selon SHRESTHA et SOLBERG [6], l'apprentissage automatique se concentre sur la construction de modèles statistiques et d'algorithmes qui permettent aux machines d'extraire des informations utiles à partir des données. Il est utilisé dans de nombreux domaines tels que la vision par ordinateur, la reconnaissance vocale, la recommandation de produits et la détection de fraudes.

Le **deeplearning**, quant à lui, est une branche spécifique de l'apprentissage automatique qui utilise des réseaux de neurones artificiels profonds pour résoudre des problèmes complexes. Comme l'expliquent GOODFELLOW, BENGIO et COURVILLE [7], les réseaux de neurones artificiels sont des modèles couramment utilisés dans l'apprentissage automatique, s'inspirant du fonctionnement des neurones dans le cerveau humain. Ainsi, les réseaux de neurones profonds, composés de nombreuses couches de neurones, sont capables d'apprendre des représentations hiérarchiques des données, ce qui leur permet de capturer des modèles et des caractéristiques plus complexes. Leur utilisation s'est révélée particulièrement efficace dans des domaines tels que la vision par ordinateur et le traitement du langage naturel.

Dans leur revue comparative, VO, WALDNER, CAO et al. [8] soulignent que le *deeplearning* a permis des avancées significatives dans la classification des données hyperspectrales, en exploitant les caractéristiques spectrales profondes pour des performances améliorées. Le *deeplearning* a également été largement utilisé dans l'analyse d'images satellites, permettant la détection d'objets, la segmentation sémantique et l'estimation de paramètres environnementaux.

Par ailleurs, dans le domaine de l'apprentissage automatique, la qualité du jeu de données est d'une importance primordiale pour garantir des performances optimales des modèles. Habituellement, le jeu de données est divisé en trois parties distinctes, comme illustré dans la Figure 4 ci - après :

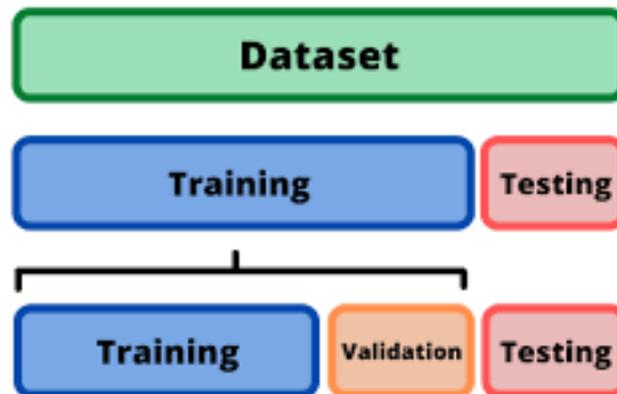


FIGURE 4 – organisation des données en apprentissage automatique

- **L'ensemble d'entraînement (*Training set*)** : il s'agit d'un sous-ensemble de données utilisé pour entraîner le modèle. Le modèle apprend à partir de ces données en ajustant ses paramètres pour minimiser l'erreur entre les prédictions et les étiquettes réelles associées aux exemples d'entraînement.
- **L'ensemble de validation (*Validation set*)** : il s'agit d'un sous-ensemble de données utilisé pour évaluer les performances du modèle pendant le processus d'apprentissage. Il est utilisé pour ajuster les hyperparamètres du modèle, tels que la taille du réseau neuronal, le taux d'apprentissage, etc., afin d'obtenir les meilleures performances possibles sur des données inconnues.
- **L'ensemble de test (*Testing set*)** : il s'agit d'un autre sous-ensemble de données qui est utilisé pour évaluer les performances finales du modèle, une fois que celui-ci a été complètement exécuté et optimisé. Cet ensemble permet d'estimer la capacité du modèle à généraliser et à effectuer des prédictions précises sur de nouvelles données qui n'ont pas été utilisées lors de l'entraînement ou de la validation.

Cette division en ensembles d'entraînement, de validation et de test est essentielle pour évaluer correctement les performances d'un modèle d'apprentissage automatique et pour éviter le surapprentissage (*overfitting*), où le modèle s'adapte trop étroitement aux données d'entraînement et ne généralise pas bien sur de nouvelles données.

3.3 Apprentissage / entraînement des modèles

Le processus d'apprentissage des modèles en apprentissage automatique et en *deep learning* peut être résumé en quatre étapes principales : la collecte et la préparation des données, la création et la formation du modèle, l'évaluation des performances du modèle et, enfin, l'utilisation du modèle pour effectuer des prédictions.

1. **Collecte et préparation des données** : tout d'abord, il est essentiel de collecter des données pertinentes pour résoudre un problème spécifique. Ces données peuvent provenir de diverses sources telles que des bases de données, des fichiers texte, des images, etc. Ensuite, les données doivent être préparées en les nettoyant, les transformant et en les adaptant à l'objectif de l'apprentissage. Cela peut inclure des tâches comme la normali-

sation des données, le traitement du texte, le redimensionnement des images, etc.

2. **Création et entraînement du modèle** : une fois les données préparées, il est temps de créer un modèle d'apprentissage automatique ou de *deep learning*. Ce modèle est une représentation mathématique qui apprend des motifs et des relations à partir des données. Il existe différents types de modèles, tels que les réseaux de neurones, les arbres de décision, les machines à vecteurs de support, etc. Le modèle est ensuite entraîné en lui fournissant les données d'entraînement. Ainsi, l'entraînement d'un modèle consiste à ajuster ses paramètres en utilisant des données d'entraînement afin de minimiser l'erreur entre ses prédictions et les valeurs réelles. Dans le cas des réseaux de neurones, cela se fait par itérations, où les données sont passées dans le modèle, l'erreur est calculée, les paramètres sont mis à jour en fonction des gradients de l'erreur, et le processus est répété. Les hyperparamètres du modèle sont ajustés périodiquement en fonction de la performance sur des données de validation. L'objectif est d'obtenir un modèle capable de généraliser et de faire des prédictions précises sur de nouvelles données.
3. **Évaluation des performances du modèle** : une fois que le modèle a été entraîné, il est évalué pour estimer sa performance et sa capacité à généraliser à de nouvelles données. Pour cela, des données de test, distinctes des données d'entraînement, sont utilisées. Les métriques d'évaluation courantes seront présentées dans la section suivante. Si le modèle ne répond pas aux critères de performance souhaités, il peut être nécessaire de revenir à l'étape précédente et d'ajuster le modèle ou les données.
4. **Utilisation du modèle pour effectuer des prédictions** : suite à l'entraînement et l'évaluation du modèle, ce dernier peut être utilisé pour effectuer des prédictions sur de nouvelles données. Le modèle est alimenté avec de nouvelles données en entrée et il produit des prédictions en utilisant les connaissances acquises pendant l'apprentissage. Ces prédictions peuvent être utilisées pour résoudre des problèmes spécifiques, comme la classification d'images, la reconnaissance vocale, etc.

3.4 Métriques d'évaluation des modèles

En *machine learning*, il existe plusieurs métriques d'évaluation qui permettent de mesurer les performances et la qualité des modèles. Voici une brève synthèse des métriques d'évaluation couramment utilisées :

- **Matrice de confusion (*Confusion Matrix*)** : la matrice de confusion est un tableau qui résume les résultats de classification d'un modèle. Elle montre le nombre de prédictions correctes et incorrectes pour chaque classe, permettant ainsi d'évaluer plus en détail les performances du modèle.
- **Précision (*Precision*)** : la précision mesure la proportion d'instances positives prédites correctement parmi toutes les instances prédites comme positives. Elle est particulièrement utile lorsque l'accent est mis sur la minimisation des faux positifs. Une valeur élevée de précision indique que le modèle a une faible propension à classer à tort des instances négatives comme positives.
- **Rappel (*Recall*)** : le rappel, également connu sous le nom de sensibilité ou taux de vrais positifs, mesure la proportion d'instances positives prédites correctement parmi toutes les instances réellement positives. Il est particulièrement utile lorsque l'accent est mis sur la minimisation des faux négatifs. Une valeur élevée de rappel indique que le modèle a une faible propension à classer à tort des instances positives comme négatives.

- ***F1-score*** : il s'agit d'une mesure qui combine à la fois la précision et le rappel en une seule valeur. Il est particulièrement utile lorsque les classes sont déséquilibrées et que l'on souhaite évaluer à la fois la capacité du modèle à minimiser les faux positifs et les faux négatifs. Le *F1-score* est calculé comme la moyenne harmonique de la précision et du rappel, ce qui pénalise les écarts entre ces deux valeurs par rapport à une moyenne arithmétique. Une valeur élevée de *F1-score* implique donc un équilibre entre la précision et le rappel du modèle.
- **Exactitude (*Accuracy*)** : l'exactitude mesure la proportion d'instances correctement classées par rapport au nombre total d'instances. C'est une métrique générale qui donne une indication globale des performances du modèle, mais elle peut être trompeuse lorsque les classes sont déséquilibrées en termes de nombres d'échantillons disponibles pour le test.
- **Courbe ROC (*Receiver Operating Characteristic*)** : la courbe *ROC* représente la performance d'un modèle de classification en traçant le taux de vrais positifs (rappel) par rapport au taux de faux positifs sur différents seuils de classification. Elle fournit une indication de la capacité du modèle à discriminer entre les classes.

Il convient de choisir les métriques d'évaluation appropriées en fonction du problème de classification spécifique et des objectifs du modèle. Cependant, l'utilisation de plusieurs métriques permet d'obtenir une vision plus complète des performances du modèle et de prendre des décisions plus éclairées.

3.5 Apprentissage automatique et classification d'images satellites

La combinaison de l'intelligence artificielle (IA) et du traitement des images satellites a ouvert de nouvelles possibilités pour l'analyse et l'interprétation des données à grande échelle. Les avancées dans le domaine de l'apprentissage automatique ont permis de développer des algorithmes de détection et de classification automatique d'objets dans les images satellites [6]. Ces techniques basées sur l'IA ont démontré leur efficacité dans des domaines tels que la détection automatique des changements environnementaux (déforestation, expansion urbaine, changements climatiques, etc.), la cartographie et la classification automatisées des caractéristiques géographiques, telles que les rivières, les lacs, les routes, les bâtiments, les cultures, etc. [8]. L'utilisation de l'apprentissage automatique a permis d'extraire des informations précieuses à partir des images satellites, facilitant ainsi la prise de décisions dans des domaines allant de la gestion des ressources naturelles, la surveillance de la végétation, la prévision des récoltes agricoles à la réponse aux catastrophes [6] [8]. Cette combinaison de l'IA et du traitement des images satellites continue d'évoluer, offrant de nouvelles perspectives pour une utilisation plus précise et efficace des données géospatiales. Ainsi, dans la littérature scientifique, les principales techniques d'apprentissage automatique pour la classification et l'analyse de série temporelle d'images satellites se présentent comme suit :

- ***Temporal Convolutional Neural Network for the Classification of Satellite Image Time Series*** [2] : cette étude propose un réseau neuronal convolutif temporel (*TempCNN*) pour la classification des séries temporelles d'images satellites. Le *TempCNN* capture à la fois les caractéristiques temporelles et spatiales pour une classification précise de l'occupation des sols.
- ***DuPLO : A Dual-Viewpoint Deep Learning Architecture for Time Series Classification*** [3] : cette recherche présente *DuPLO*, une architecture d'apprentissage en profondeur à double point de vue pour la classification des séries temporelles. *DuPLO* exploite des caractéristiques spatio-temporelles extraites à partir de deux points

de vue différents pour améliorer la classification de l'occupation des sols.

- ***M3Fusion : A Deep Learning Architecture for Multiscale Multimodal Multitemporal Satellite Data Fusion*** [4] : cette recherche introduit une architecture d'apprentissage en profondeur pour la fusion de données satellites multiéchelles, multimodales et multitemporelles provenant de différentes sources. *M3Fusion* utilise des techniques de fusion pour améliorer la classification de l'occupation des sols.
- ***Operational High-Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series*** [1] : cette recherche propose une méthode opérationnelle de production de cartes de couverture terrestre à haute résolution à l'échelle d'un pays en utilisant des séries temporelles d'images satellites. Elle met l'accent sur l'analyse des séries temporelles et l'utilisation des informations pour la classification de l'occupation des sols.
- ***The MORINGA Processing Chain : Automatic Object-Based Land Cover Classification of Tropical Agro-systems Using Multi-Sensor Satellite Imagery [moringa]*** : cette recherche présente une chaîne de traitement pour la classification automatique de la couverture terrestre basée sur des objets dans les agrosystèmes tropicaux en utilisant des images satellites provenant de différents capteurs. Elle utilise des techniques d'apprentissage automatique pour extraire des caractéristiques et classifier l'occupation des sols.
- ***Deep Learning-Based Classification for Land Use and Land Cover Mapping*** [9] : cette recherche explore l'utilisation de l'apprentissage en profondeur, en particulier des réseaux de neurones convolutifs (*CNN*), pour la classification de l'occupation et de l'usage des sols à partir de séries temporelles d'images satellites.
- ***Ensemble Learning Approaches for Land Cover Classification of Multitemporal Satellite Images*** [10] : Cette étude propose l'utilisation d'approches d'apprentissage en ensemble, telles que les forêts aléatoires, les machines à vecteurs de support (SVM) et les réseaux de neurones profonds, pour la classification de l'occupation et de l'usage des sols dans les séries temporelles d'images satellites. L'ensemble des modèles permet de combiner les prédictions de plusieurs classificateurs pour améliorer les performances.
- ***Spatiotemporal Fusion for Land Cover Classification Using Multitemporal Satellite Images*** [11] : Cette recherche propose des techniques de fusion spatio-temporelle pour la classification de l'occupation et de l'usage des sols à partir de séries temporelles d'images satellite. Elle explore des méthodes telles que la fusion des caractéristiques extraites à partir de chaque instant temporel, la fusion des prédictions de chaque instant ou la fusion des modèles de classification.
- ***Transfer Learning for Land-Use Classification in Multispectral Satellite Images*** [12] : Cette étude utilise le transfert d'apprentissage pour la classification de l'occupation et de l'usage des sols à partir de séries temporelles d'images satellites. Elle exploite des modèles pré-entraînés sur des tâches similaires ou des ensembles de données plus vastes, puis adapte ces modèles à la classification spécifique des images satellite.

Comme on peut le constater, la plupart des techniques explorées se base sur une architecture de réseaux de neurones convolutifs (*CNN*) avec des configurations différentes. Dans ce travail, nous nous concentrerons sur quelques unes des architectures dont le *TempCNN*, une méthode basée sur les réseaux convolutifs en 2D dans la dimension spatiale (*CNN2D*) puis un couplage des deux suivant l'approche proposée par *DuPLO*.

4 Approche méthodologique et mise en oeuvre du projet

4.1 Présentation du site de l'étude

Le site d'étude est un des sites pilotes du projet OBSYDYA Bénin couvrant une zone de 50 X 50 Km² autour de la ville de Parakou. Parakou est une ville du département du Borgou située à 415 km au Nord de Cotonou avec une superficie totale d'environ 441 km², dont 30 km² sont urbanisés. Limitée au Nord par la commune de N'Dali, au Sud, à l'Est et à l'Ouest par celle de Tchaourou, la ville de Parakou se trouve à 9°21' de latitude Nord, à 2°36' de longitude Est à une altitude moyenne de 350 m. Elle bénéficie d'un climat tropical de type soudanien, caractérisé par une saison sèche de cinq (05) mois, d'octobre à avril, et une saison pluvieuse le reste de l'année. Cette zone est marquée par des fluctuations de température avec une pluviométrie moyenne de 1200 mm par an. Le réseau hydrographique est constitué de cours d'eaux, de ruisseaux ou de marigots. Le couvert végétal observé à Parakou est dominé par la savane arborée et se caractérise par la présence du néré (*Parkia biglobosa*), du faux acajou (*Blighia sapinda*), de bois d'ébène (*Diospyros mespilifounis*), de karité (*Butyrosperum paradoxum*), etc. Les basfonds sont des prairies marécageuses de savanes, des buissons de bambous (*Bambusa arundinacca*). La région de Parakou se distingue par ses sols à texture légère et une épaisseur importante, résultant de faibles niveaux d'érosion et d'un lessivage en profondeur. La population de Parakou a connu une croissance moyenne annuelle de 2,1 %, passant de 103 577 habitants¹ en 1992 à 255 478 habitants² en 2013, dont une grande partie est concentrée dans la zone urbanisée.

Par ailleurs, dans les localités périphériques de la commune de Parakou, l'agriculture est l'activité prédominante, occupant la quasi-totalité des ménages. Les terres cultivées, gérées par les chefs de ménage, sont généralement héritées de père en fils. Les principales cultures vivrières développées sont le soja, le maïs, le manioc, le sorgho et le niébé, suivies de l'igname et du riz. En parallèle, les cultures de rente comme le coton, l'anacarde, l'arachide et le teck contribuent également à l'économie locale. Cependant, en raison des problèmes écologiques et financiers, la culture du coton est en déclin, laissant place à l'essor des cultures de soja, plantations de tecks, d'anacardiens et de manguiers qui offrent une double opportunité : protéger les sols tout en générant des revenus importants pour les agriculteurs.

La figure 5 montre la carte du site (OBSYDYA) de Parakou sur lequel l'étude a été menée.

1. RGPH3 (<https://instad.bj/statistiques/enquetes-et-recensements>)

2. RGPH4 (<https://instad.bj/statistiques/enquetes-et-recensements>)

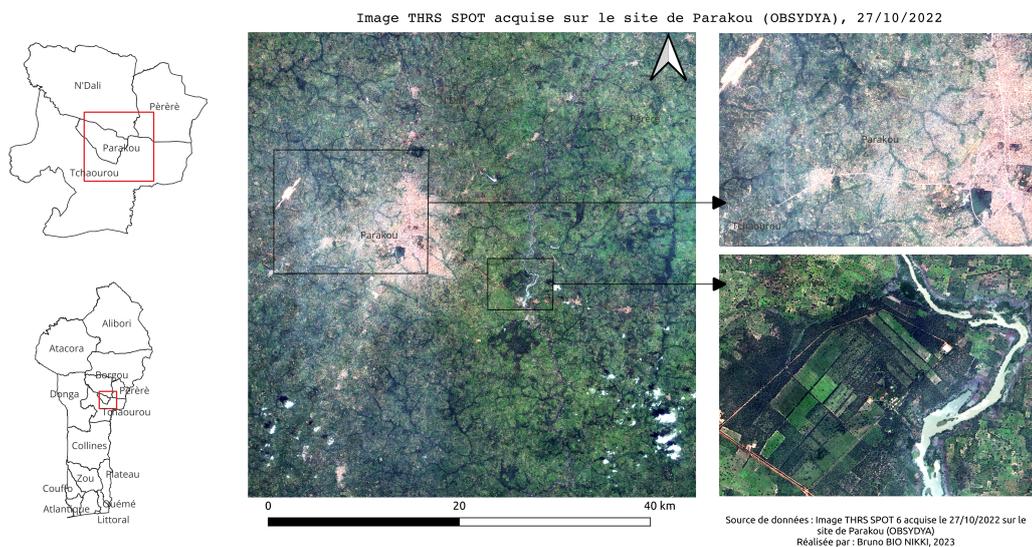


FIGURE 5 – Carte du site d'étude de Parakou (OBSYDYA)

4.2 Méthodologie générale

L'approche méthodologie utilisée dans ce travail a consisté en sa subdivision en cinq (05) grandes étapes complémentaires comme l'illustre la figure 6 ci - après. La première étape du processus de traitement consiste à la collecte des données de terrains pour constituer la base de données vérité terrain, puis l'acquisition des images satellites nécessaires à l'étude. Après cette étape, suit celle de la préparation et la mise en forme des données afin de constituer un jeux de données exploitable par les algorithmes et modèles développés aux étapes suivantes. Lesdites étapes consistent à la mise en place d'une baseline *Random Forest* comme algorithme de référence puis respectivement au test des méthodes par pixel (*TempCNN*) puis par patchs (*CNN2D et DuPLO*). A la suite de ces tests de performance, une évaluation et une étude comparative permettront de mieux évaluer objectivement chacune des méthodes.



FIGURE 6 – Méthodologie générale

4.3 Collecte et préparation des données

4.3.1 Collecte des données terrains et post-traitements

Le protocole d'acquisition des données terrains suit les lignes directrices mises en place par l'équipe du projet OBSYDYA Bénin. Pour cela, une campagne de terrain a eu lieu courant novembre 2022 sur le site d'étude de Parakou avec des adaptations pour tenir compte des caractéristiques de l'agriculture tropicale, notamment la petite taille des exploitations agricoles et les difficultés d'accessibilité. Pour faciliter la localisation des zones d'échantillonnage et l'acquisition à distance des points *GPS*, les opérateurs de terrain étaient équipés de tablettes *GPS* (*Samsung*) leur donnant accès à un projet *QGIS* avec des images à très haute résolution spatiale (*VHSR*) (images *SPOT 6/7* commandées juste avant les enquêtes). Ces équipements ont permis l'enregistrement *in situ* des attributs de 1883 points *GPS* représentant notre échantillon. Pour chaque point *GPS*, des attributs tels que l'ID, la date du relevé terrain, les identifiants et nom des classes au niveau 0 (échelle grossière) et au niveau 1 (échelle plus fine) ont été enregistrés. Sur le formulaire d'entrée de données, les ID sont enregistrés de façon automatique alors que des listes déroulantes ont été utilisées pour les autres attributs afin d'éviter les erreurs de saisie. Une fois toutes les données collectées sur le terrain, elles ont été compilées et transférées de manière sécurisée vers une base de données centrale, où les points de terrain ont été spatialisés à l'échelle de parcelles par photo-interprétation de l'acquisition à *THR*. Cette digitalisation facilite l'analyse ultérieure des données et leur utilisation dans des applications cartographiques ou d'études agricoles, permettant ainsi une meilleure compréhension des éléments d'occupation du sol et des exploitations agricoles dans la région étudiée. Les détails sur la nomenclature utilisée dans ce travail (qui correspond au niveau 0 dans le protocole OBSYDYA), ainsi que sur le nombre de polygones par classe et leurs correspondants en termes de pixels *Sentinel-2*, sont présentés dans le tableau 5.

TABLE 5 – Détails sur la nomenclature utilisée

Classes	Nombre de polygones	Nombre de pixels
<i>Cropland</i>	1158	49076
<i>Tree Crops</i>	479	27978
<i>Natural Woody Vegetation</i>	91	19951
<i>Grassland / Shrubland</i>	86	8602
<i>Bare Soil / Built-up Surfaces</i>	53	1564
<i>Water</i>	16	3585

L'analyse de la figure 7 montre un déséquilibre sur la représentativité des classes dans les données de référence. On note par exemple certaines classes majoritaires comme les '*Cropland*' et '*Tree Crops*' qui représentent respectivement 61,5% et 25,4% des polygones de la base de données de référence.

Statistiques par classe

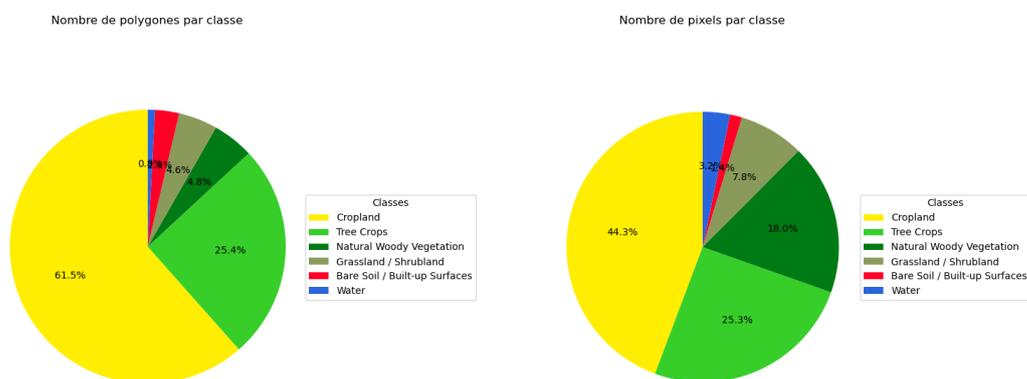


FIGURE 7 – Répartition des polygones et pixels par classe sur les données de références

Par ailleurs, afin d'avoir des informations dynamiques sur l'évolution de la végétation pour chaque classe sur l'année 2022 (tendances de croissance ou de déclin) et de mieux comprendre la distinction entre les classes ayant des motifs similaires en fonction des valeurs spectrales des pixels, nous avons étudié le profil temporel du *NDVI* par classe sur la période d'étude comme le montre la figure 8.

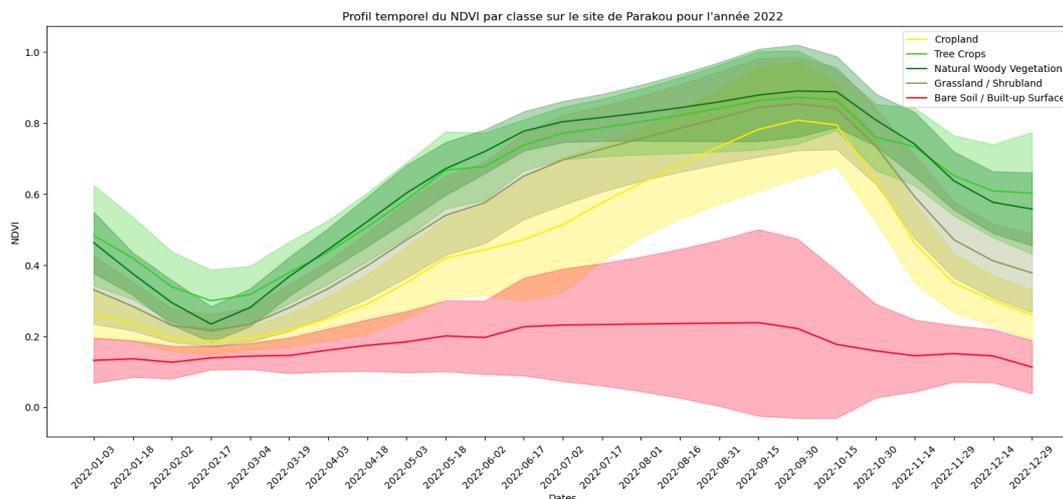


FIGURE 8 – Profil temporel du *NDVI* par classe sur le site de Parakou pour l'année 2022

L'analyse du profil temporel du *NDVI* par classe montre clairement que la valeur de *NDVI* est fortement corrélée à la pluviométrie de la saison au cours de l'année pour toutes les classes caractérisées par un couvert végétal. On remarque également que malgré le fait que certaines classes telles que le *Cropland* produit moins de biomasse verte que les milieux naturels (*Natural Woody Vegetation*), il y a une variabilité qui fait que les classes sont assez mélangées entre elle et se confondent quelques fois. Il se pose alors là un problème de séparabilité des classes d'occupation du sol. Ce qui motive à explorer les méthodes pointues d'analyse qui permettront de mieux explorer et fouiller les données acquises.

4.3.2 Acquisition des images satellites et mise en forme des données

Dans cette étude, nous exploitons les produits *Sentinel-2* (*S2*) dérivés des satellites jumeaux *S2A* et *S2B*, fournissant des images multispectrales et multi-temporelles de haute résolution accessibles sur la plateforme de données et services pour les surfaces continentales *Theia Land* [13]. Les produits *S2* se composent de treize (13) bandes spectrales avec une résolution spatiale variant entre 10m, 20m et 60m, selon la bande spécifique considérée.

En effet, les expériences ont été menées sur une série temporelle de 25 images *Sentinel-2* extraites suivant les limites du site d'étude de Parakou. Chaque image est composée de dix bandes spectrales avec une résolution spatiale de 10 mètres (les six (06) bandes à 20 mètres ont été interpolées spatialement à 10 mètres). De même, pour chaque image, six (06) indices radiométriques (voir figure 9) ont été calculées et concaténées aux dix (10) bandes initiales ; ce qui ramène à seize (16) bandes par image. Notons que la série commence en Janvier 2022 et se termine en Décembre 2022.

Normalized Difference Vegetation Index	$NDVI = \frac{B8 - B4}{B8 + B4}$	[Rouse et al., 1974]
Normalized Difference Water Index	$NDWI = \frac{B3 - B8}{B3 + B8}$	[McFEETERS, 1996]
Brightness Index	$BRI = \sqrt{B2^2 + \dots + B12^2}$	[Inglada et al., 2017]
Modified NDWI	$MNDWI = \frac{B3 - B11}{B3 + B11}$	[Xu, 2006]
Short Wave NDVI	$SWNDVI = \frac{B11 - B8}{B11 + B8}$	[Gao, 1996]
Normalized Difference Red Edge Index	$NDRE = \frac{B8 - B5}{B8 + B5}$	[Barnes et al., 2000]

FIGURE 9 – Indices radiométriques calculés

Par ailleurs, les données ont été initialement préparées grâce à un script écrit en *python* et utilisant l’outil de traitement des images de télédétection *Orfeo ToolBox (OTB)* [14]. Les différentes étapes de ce processus sont illustrées par la figure 10 ci - après :

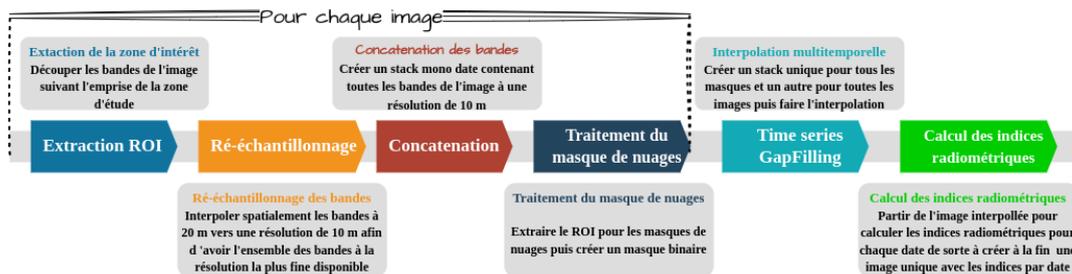


FIGURE 10 – Processus de préparation des données

Les effets atmosphériques, d’adjacence et de pente sont corrigés en utilisant la chaîne de traitement *MACCS/MAJA*. Sur un total de 77 images initialement acquises, les 25 images utilisées ont été obtenues en faisant une interpolation temporelle linéaire sur une grille temporelle régulière avec un intervalle de quinze (15) jours afin de combler les vides créés par la présence de nuages sur certaines images de la série temporelle [1].

A la fin de tout ce processus, nous obtenons une base de données représentant toute la série temporelle d’image S2 sur l’année 2022 (400 variables, soit 16 bandes \times 25 estampilles temporelles) et une base de données vérité terrain utilisable pour tester les différentes méthodes retenues. La section suivante décrit la technique d’échantillonnage des données.

4.4 Échantillonnage des données, apprentissage et évaluation des modèles

Afin de créer un cadre logique de comparaison et de test des modèles, nous avons procédé à un échantillonnage des données permettant ainsi de construire un même jeu de données d'apprentissage et de test pour tous nos modèles. Cet échantillonnage a été fait "par polygones" de telle sorte que tous les pixels/patches d'un polygone vont soit à l'apprentissage soit au test pour éviter les biais d'apprentissage dus à l'autocorrélation spatiale. Ainsi, les données de références ont été rasterisées pour former le raster des labels et celui des groupes indivisibles (ici, les identifiants uniques des polygones de la BD) utiles pour le processus d'apprentissage des modèles et la séparation des données en jeux de données de d'apprentissage et de test. Ces rasters sont ensuite transformés en tableau de données multidimensionnelles grâce à la bibliothèque *Rasterio*. Par la suite, pour la construction de notre jeu de données, les données brutes de la série temporelle sont chargées de façon séquentielle (par lot) sous forme d'un tableau multidimensionnelle en tenant compte uniquement des pixels qui sont annotées (les pixels pour lesquels une vérité terrain existe). Au cours de ce processus, chaque séquence de données est normalisée (normalisation de type *Min/Max*) en utilisant des valeurs statistiques des percentiles à 2% et à 98%. Ensuite, les données sont divisées en ensembles d'entraînement et de test à l'aide d'une approche de validation croisée stratifiée par groupes (*StratifiedGroupKFold* de la bibliothèque *Scikit-Learn* [15]), assurant ainsi une distribution équilibrée des classes dans chaque *fold* (cinq (05) *folds* au total) permettant ainsi d'avoir une validation cohérente des différentes techniques. Pour chaque *fold*, un nouveau modèle est créé, puis entraîné sur les données d'entraînement spécifiques à ce *fold*. Le modèle entraîné est ensuite évalué sur les données de test pour calculer les scores F1 pour chaque classe. Ces scores et d'autres métriques sont imprimés et sauvegardés pour chaque *fold*. À la fin de la validation croisée, le modèle entraîné est exporté pour une utilisation ultérieure. L'analyse finale agrège les prédictions et étiquettes des différents modèles créés par *fold* pour évaluer les performances globales du modèle.

Par ailleurs, pour le cas spécifique des méthodes *CNN2D* et pour l'approche style *DuPLO*, l'échantillonnage se fait par patches centrés sur chaque pixel.

Concernant la stratégie d'apprentissage, du fait de la disponibilité déjà limitée de données de référence nous n'avons pas utilisé un jeu de données de validation tels que décrit en Section 4.5 pour choisir le nombre d'époques d'apprentissage, mais nous avons décidé de figer ce nombre en regardant l'évolution des *Loss* (*pertes*) calculées pendant l'apprentissage. Au bout d'un certain nombre d'époque on constate que la fonction de perte ne change plus trop. Ce qui nous a permis de fixer le nombre d'époques. De même, pour les modèles de *deep learning TempCNN*, *CNN2D* et *DuPLO*, les paramètres de poids et de biais sont optimisés en utilisant *Adam* (*Optimisation du moment adaptatif*) avec ses valeurs de paramètres par défaut. La taille du lot (*batch size*) est fixée à 32, et le nombre d'époques à 100.

4.5 Présentation, paramétrage et implémentation des techniques

Comme indiqué dans l'approche méthodologique générale, après la préparation et la mise en forme des données, nous mettons en place un cadre de comparaison des techniques retenues à travers une technique de référence (le *Random Forest*); nous testons d'abord les méthodes par pixels et ensuite celles par patches. Dans cette partie de notre travail, nous faisons une description succincte de chacune de ces techniques.

4.5.1 Présentation et paramétrage de la baseline *Random Forest*

Random Forest est un puissant algorithme d'apprentissage automatique qui appartient à la famille des méthodes d'ensemble. Il a été proposé par BREIMAN [16] en 2001 et est largement utilisé pour des tâches de classification et de régression dans divers domaines. Cela inclut la classification d'images satellites en raison de sa robustesse, de sa précision et de sa capacité à gérer des données à haute dimension. En combinant plusieurs arbres de décision et en effectuant des sélections aléatoires de caractéristiques, *Random Forest* produit des résultats fiables pour la classification d'images satellites, même dans des environnements complexes [17]. Les images satellites contiennent souvent un grand nombre de bandes spectrales et de pixels, mais l'algorithme est capable de gérer efficacement ces données sans surajustement. De plus, *Random Forest* peut découvrir des interactions non linéaires entre les caractéristiques, ce qui est essentiel pour capturer les relations complexes présentes dans les images satellites [17]. En somme, *Random Forest* constitue un choix judicieux comme algorithme de référence pour la classification d'images satellites, offrant des performances solides et des informations interprétables pour les applications de télédétection et de surveillance environnementale [17] [18].

Dans le cadre de la mise au point de la *baseline* sur l'algorithme *Random Forest*, nous avons opté pour une implémentation dans la bibliothèque *Scikit-Learn* [15] en *Python*. Le modèle *Random Forest* a été configuré avec des paramètres soigneusement sélectionnés pour garantir des performances optimales. Plus précisément, nous avons choisi d'utiliser 500 arbres pour former notre modèle, en utilisant le critère de partitionnement '*gini*' pour évaluer la qualité des séparations. De plus, nous avons laissé la profondeur maximale des arbres illimitée, tout en définissant le nombre minimum d'échantillons requis pour diviser un nœud à 2 et le nombre minimum d'échantillons requis dans une feuille à 1 pour assurer un ajustement précis et éviter le surajustement. Afin de sélectionner les caractéristiques de manière pertinente, nous avons opté pour la méthode '*sqrt*' qui choisit le nombre de caractéristiques à tester pour une subdivision de nœud égal à la racine carrée du nombre total de caractéristiques.

4.5.2 Présentation et paramétrage de *TempCNN*

TempCNN (*Temporal Convolutional Neural Network for the Classification of Satellite Image Time Series*) [2] est une architecture de réseau neuronal convolutif (*CNN*) spécialement conçue pour l'analyse des séries temporelles. Cette approche novatrice utilise des couches de convolution, qui sont plus communément utilisés pour une analyse dans la dimension spatiale, pour plutôt extraire efficacement les caractéristiques temporelles des données d'entrée en travaillant sur les dimensions spectrales et temporelles. En agrégeant ces caractéristiques, *TempCNN* les propage ensuite à travers des couches entièrement connectées, permettant ainsi de capturer les relations temporelles importantes dans les séries d'images [2]. Grâce à sa capacité à capturer les dépendances temporelles, ce modèle est capable de prendre des décisions basées sur l'évolution de la séquence d'images, ce qui en fait un outil puissant pour l'analyse des données temporelles complexes. La combinaison des couches de convolution avec la propagation de caractéristiques dans des couches entièrement connectées confère à *TempCNN* une puissance d'apprentissage profond, facilitant ainsi l'identification de motifs significatifs dans les séries temporelles pour des applications variées.

L'architecture mise en œuvre dans le cadre de cette étude est un modèle de réseau neuronal convolutif (*CNN*) inspiré du **TempCNN** proposé par PELLETIER, WEBB et PETITJEAN [2] qui comprend trois (03) couches de convolution *1D*, chacune suivie d'une couche de normalisation par lots (*BatchNorm1d*), d'une activation *ReLU* et d'un taux de

désactivation (*dropout*) pour prévenir le surajustement. Les sorties des couches de convolution sont ensuite aplaties et passées à travers une couche entièrement connectée (*fully connected layer*) suivie d'une autre couche de normalisation par lots et d'une activation *ReLU*. Finalement, une couche de sortie linéaire (*softmax*) est utilisée pour générer les probabilités de classification pour les différentes classes. Ce modèle a été conçu avec une fonction utilitaire pour la création des couches de convolution et une autre pour les couches entièrement connectées, permettant une définition modulaire et réutilisable des blocs de construction. Le modèle utilise également des hyperparamètres tels que le taux de régularisation *L2*, le taux de désactivation et le nombre d'unités dans les couches de convolution et les couches entièrement connectées pour réguler et ajuster la complexité du modèle.

La figure 11 montre l'architecture du *TempCNN* mise en oeuvre.

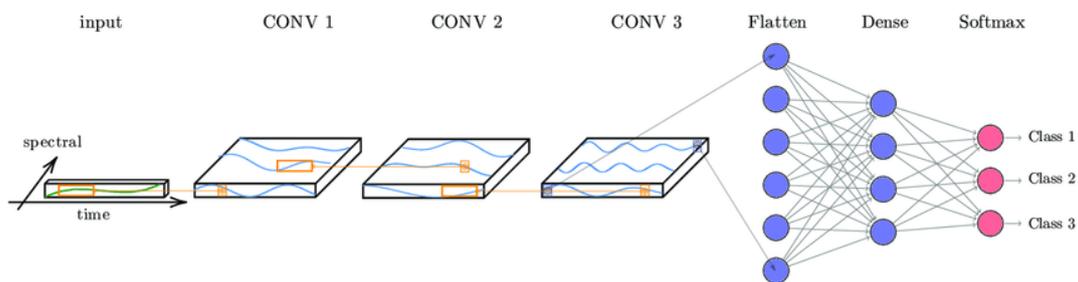


FIGURE 11 – Architecture de *TempCNN*

4.5.3 Présentation et paramétrage de *CNN2D*

CNN2D est une architecture de réseau neuronal convolutif traditionnelle adaptée à l'analyse de séries temporelles d'images [19]. Inspirée de la vision humaine, cette approche permet de prendre en compte les corrélations spatiales, les contextes, les voisinages, les structures et les textures présentes dans chaque image considérée comme un instant temporel. En utilisant des couches de convolution, *CNN2D* capture efficacement les informations spatiales dans chaque image et analyse l'évolution des caractéristiques extraites à travers la séquence pour saisir simultanément les informations spatiales et temporelles dans les séries temporelles d'images [19]. Malgré l'absence d'une modélisation explicite des caractéristiques temporelles (bandes et estampilles temporelles sont traités au même niveau, contrairement à *TempCNN*, cette utilisation astucieuse de l'architecture convolutive classique fait de *CNN2D* un outil puissant pour l'exploration et l'exploitation des informations dans des séries d'images multi-bandes chronologiques.

Le modèle "*CNN2D*" implémenté dans le cadre de cette étude est un réseau neuronal convolutif (*CNN*) composé de trois couches de convolution *2D*, chacune suivie d'une normalisation par lots (*BatchNorm2d*), d'une activation *ReLU* et d'un taux de désactivation (*dropout*) pour prévenir le surajustement. Les sorties des couches de convolution sont ensuite aplaties et passées à travers une couche entièrement connectée (*fully connected layer*) suivie d'une autre couche de normalisation par lots et d'une activation *ReLU*. Finalement, une couche de sortie linéaire est utilisée pour générer les logits (scores de classe) pour chaque image d'entrée. Le modèle utilise également des hyperparamètres tels que le nombre d'unités dans les couches de convolution et les couches entièrement connectées, le taux de désactivation et la taille des noyaux de convolution pour réguler et ajuster la complexité du modèle.

La figure 12 présente l'architecture du *CNN2D* mise en oeuvre.

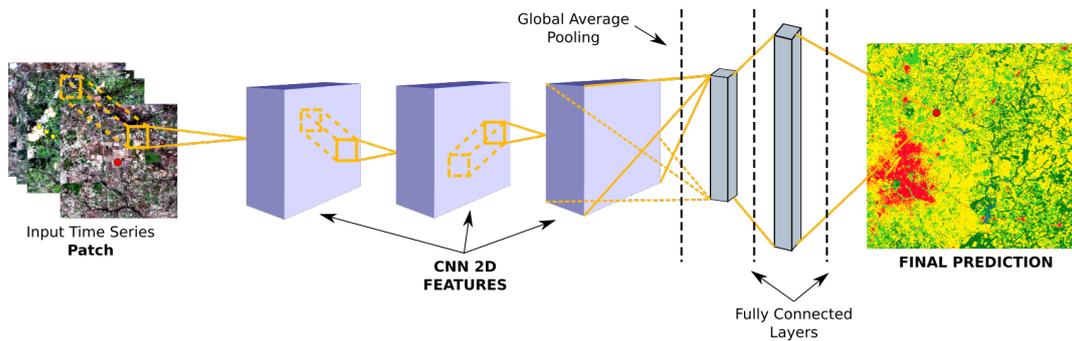
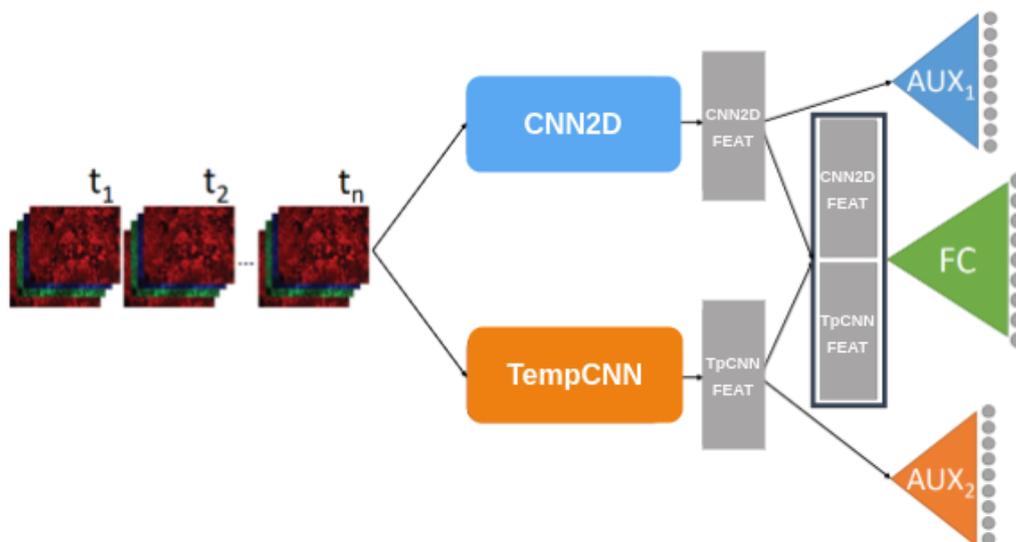


FIGURE 12 – Architecture de CNN2D

4.5.4 Présentation et paramétrage de *DuPLO*

Initialement implémentée par INTERDONATO, IENCO, GAETANO et al. [3], *DuPLO* (*A Dual ViewPoint deep Learning architecture for time series classification*) est une architecture *deeplearning* qui adopte une double perspective pour la classification de séries temporelles d'images, intégrant les motifs locaux et les tendances globales afin d'améliorer la précision de la classification. En intégrant ces deux perspectives, *DuPLO* vise à améliorer la précision et la robustesse de la classification des séries temporelles. Pour ce faire, cette approche implique des réseaux neuronaux récurrents (*RNN*) et des réseaux neuronaux convolutifs (*CNN*) adaptés pour capturer efficacement les dépendances temporelles ainsi que les motifs spatiaux au sein des données de séries temporelles.

Inspirées de cette approche, dans ce travail, nous avons implémenté une architecture similaire combinant des informations provenant de deux perspectives distinctes : le *TempCNN* d'un côté et le *CNN2D* de l'autre côté dont nous avons fait cas respectivement dans les sections 4.5.2 et 4.5.3. Dans cette architecture "*DuPLO*", la partie "*CNN2D*" consiste en une couche de réseau neuronal convolutif (*CNN*) conçue pour traiter des données en deux dimensions. La partie "*TempCNN*" est dédiée à un autre *CNN* spécialement conçu pour des données temporelles. Afin de mettre en place cette approche combinant les deux architectures, dans la propagation avant (*forward*), les données *2D* sont d'abord passées à travers le *CNN2D* pour obtenir des descripteurs associés. Ensuite, les pixels centraux de chaque *patch 2D* sont extraits et remodelés pour correspondre à la forme d'entrée attendue par le réseau temporel *TempCNN*. Ces données temporelles sont ensuite traitées par ce dernier pour générer des descripteurs temporelles. En plus des deux branches principales, il y a deux classifieurs auxiliaires qui sont responsables de générer des prédictions par sources à partir des caractéristiques extraites par les deux branches principales, afin de forcer la complémentarité dans le calcul de ces derniers. Les descripteurs auxiliaires sont ensuite concaténés et envoyés à une couche entièrement connectée finale qui produit les prédictions de sortie du modèle tenant compte à la fois des informations spatiales et temporelles apprises. Cette architecture vise à capturer des informations complémentaires des deux perspectives, améliorant ainsi la précision globale de la classification des séries temporelles. La figure 13 montre l'architecture du modèle *DuPLO* mise en oeuvre.


 FIGURE 13 – Architecture de *DuPLO*

4.6 Production des cartes

Le processus de production des cartes géospatiales à partir des séries temporelles d'images est réalisé à l'aide d'un *script Python* pour chacune des techniques implémentées. Tout d'abord, les images de la série temporelle sont chargées à partir d'un répertoire spécifique. Ensuite, les modèles pré-entraînés concernés (5 au total soit un modèle par *fold*) sont chargés pour effectuer les prédictions sur toute la série temporelle. Pour faciliter le traitement de ce grand volume de données, les données sont découpées en tuiles, grâce à la fonction *Window* de la bibliothèque *Rasterio*. Chaque tuile est ensuite normalisée (normalisation de type *Min/Max*) en utilisant des valeurs statistiques (les percentiles à 2% et 98%) obtenues préalablement à partir du jeu d'échantillons issu de la base de données de référence. Le modèle est ensuite utilisé pour prédire les étiquettes de classe pour chaque tuile de l'image de la série temporelle. Afin de garantir de meilleurs résultats, nous effectuons des prédictions avec les cinq (05) modèles pour chaque ensemble de données de la série temporelle, puis choisissons la meilleure prédiction au travers d'un vote majoritaire avant de l'enregistrer dans l'image de sortie. Ce processus de traitement est répété pour chaque tuile de l'image, garantissant la cohérence et l'intégrité des résultats. Une fois toutes les tuiles traitées, la carte de sortie finale créée permet d'avoir une cartographie très précise de l'occupation du sol à travers les données de la série temporelle d'images satellites.

4.7 Matériels et outils techniques utilisés

La mise en oeuvre de cette étude a nécessité l'utilisation d'un certain nombre de matériels et outils techniques dont :

- un ordinateur portable *HP EliteBook 830 G8 Notebook PC* équipé d'un processeur *11th Gen Intel® Core™ i7-1165G7 @ 2.80GHz × 8*, d'une carte graphique *Mesa Intel® Xe Graphics (TGL GT2)*, de 32 giga de RAM et 1 To de Stockage tournant sur une distribution linux *Ubuntu 22.04.3 LTS* qui a permis de gérer des tâches liées au traitement et à l'analyse des données géospatiales. Cet équipement a été un atout essentiel pour mener

à bien l'ensemble de l'étude.

- un ordinateur équipé d'un processeur *Intel (R) Xeon (R) E5-2667 v4@3.20Ghz avec 256 Go de RAM et quatre GPU TITAN X*, situé à Montpellier en France et accédé à distance depuis mon poste à Parakou au Bénin, utilisé pour accélérer les calculs intensifs liés à l'entraînement des modèles de *deep learning*. Son utilisation a permis d'accélérer le processus de formation des modèles et d'améliorer la performance globale des algorithmes implémentés.
- le langage *Python* : *Python* est un langage de programmation polyvalent et convivial qui a été largement utilisé dans cette étude pour la manipulation de données, l'implémentation des algorithmes, et l'analyse des résultats. Sa richesse en bibliothèques dédiées à la géospatialité et au *machine learning* en a fait un choix idéal pour cette recherche.
- l'environnement *Anaconda* [20] : *Anaconda* est une plateforme *open-source* qui facilite la gestion des environnements *Python* et l'installation de bibliothèques scientifiques essentielles. Grâce à *Anaconda*, la mise en place d'un environnement de développement cohérent et compatible a été simplifiée, assurant ainsi la reproductibilité des résultats.
- les Bibliothèques *Numpy*, *Rasterio*, *Scikit-Learn*, *PyTorch*[21] : ces bibliothèques *Python* sont essentielles pour l'analyse géospatiale et l'apprentissage automatique. *Numpy* offre des fonctionnalités avancées de calcul numérique, *Rasterio* permet de lire et de traiter les données raster, *Scikit-Learn* propose des algorithmes d'apprentissage automatique puissants et *PyTorch* est une bibliothèque de *deep learning* qui facilite l'implémentation des modèles de réseaux neuronaux.
- *Orfeo ToolBox (OTB)* [14] : l'*Orfeo Toolbox* est une bibliothèque *open-source* spécialisée en télédétection, offrant des outils avancés pour le traitement et l'analyse des images satellites. Elle a été utilisée dans cette étude pour des opérations spécifiques liées à la préparation et la mise en forme des données.
- *EODAG* [22] (*Earth Observation Data Access Gateway*) : *EODAG* est un outil *open-source* qui facilite l'accès et la recherche de données d'observation de la Terre provenant de différentes sources et plateformes. Il a été utilisé pour collecter des données géospatiales sur la plateforme *Theia Land* [13].
- *Visual Studio Code (VSCode)* : *Visual Studio Code* est un éditeur de code très apprécié pour sa légèreté et ses fonctionnalités avancées, adapté au développement Python. Son utilisation dans le cadre de cette étude a permis une programmation efficace et un débogage interactif pour faciliter le développement et l'optimisation du code. Il a également permis de faciliter l'accès au code sur le ordinateur en SSH.
- le logiciel *QGIS (Quantum GIS)* : *QGIS* est un logiciel *open-source* de système d'information géographique (SIG) utilisé pour la visualisation, l'analyse et la manipulation des données géospatiales. Son utilisation a été essentielle pour préparer, visualiser et interpréter les données géographiques utilisées et obtenues dans l'étude.

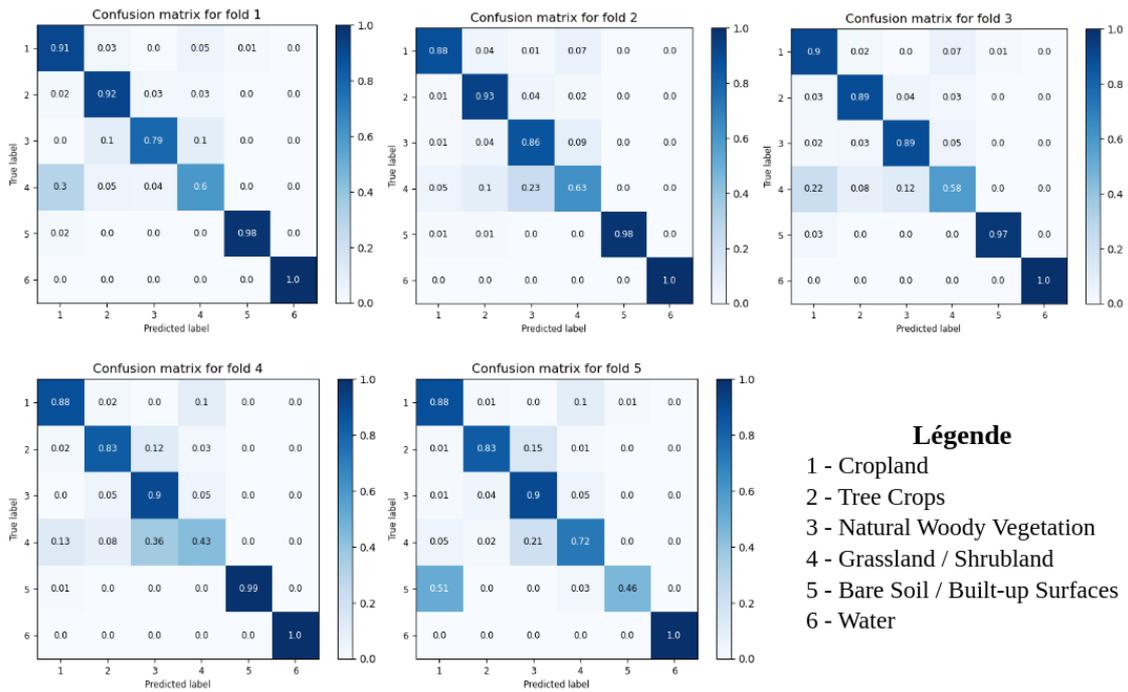
5 Résultats obtenus et Discussions

A l'issu des différentes expérimentations réalisées au cours de cette étude, nous sommes parvenu à ressortir les performances de chacune des techniques implémentées sur le jeu de données spécifiques à notre site d'étude. Dans cette section, nous présentons les performances obtenues avec chaque modèle puis nous faisons une étude comparative de ces modèles.

5.1 Test de performances des techniques implémentées

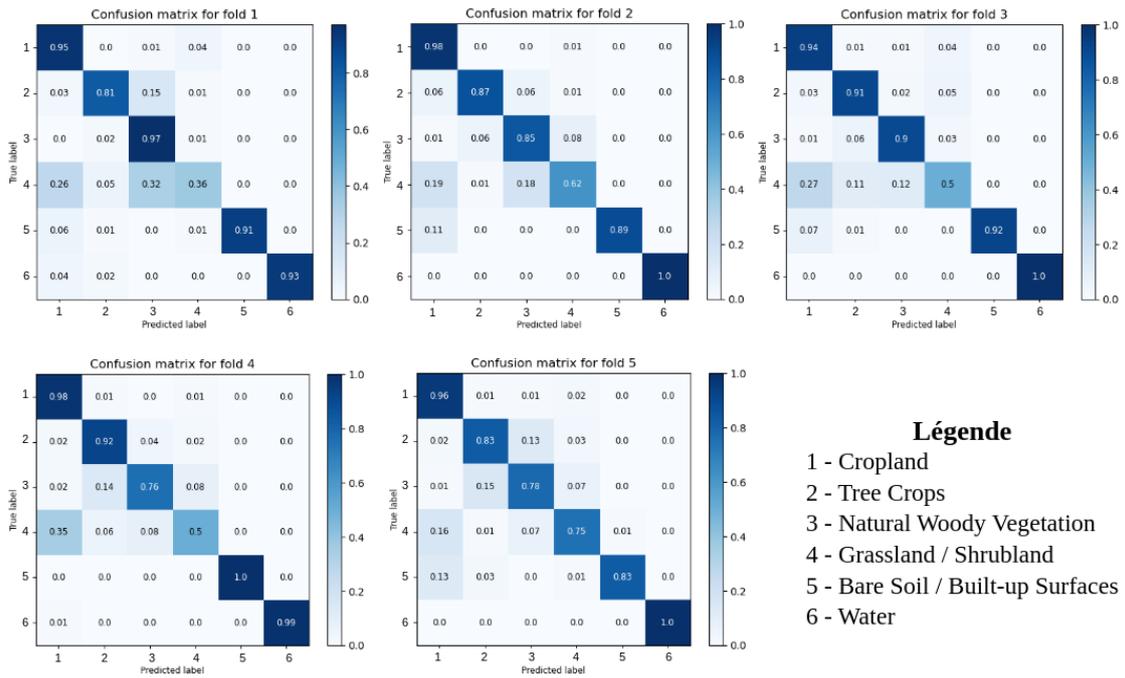
Les figures 14, 15, 16, et 17 présentent respectivement les matrices de confusions par *Fold* pour chacune des techniques *Random Forest*, *TempCNN*, *CNN2D* et *DuPLO*. De l'analyse de chacune de ces matrices de confusion, il ressort qu'il y a une cohérence entre les différents *folds* (les matrices ne se contredisent pas) : ce qui atteste la fiabilité des résultats obtenus. De même, on constate globalement que la classe *Grassland / Shrubland* est celle plus "faible" pour tous les *folds* et pour tous les modèles. Etant très variée, quelques confusions sont à noter sur cette classe : elle est tantôt classée dans les ligneux (*Natural Woody Vegetation*), tantôt dans le *cropland*, tantôt dans les sols nus et surfaces artificialisées (*Bare Soil / Built-up Surfaces*).

Par ailleurs, quelques petites différences se remarquent également d'un modèle à un autre : par exemple, bien que le modèle *TempCNN* soit globalement le meilleur, on remarque une meilleure concentration des matrices de confusion autour de leur diagonales pour le modèle *DuPLO*.



Average Accuracy : **87.3 %** Average Cohen's Kappa Coefficient : **81.44 %**

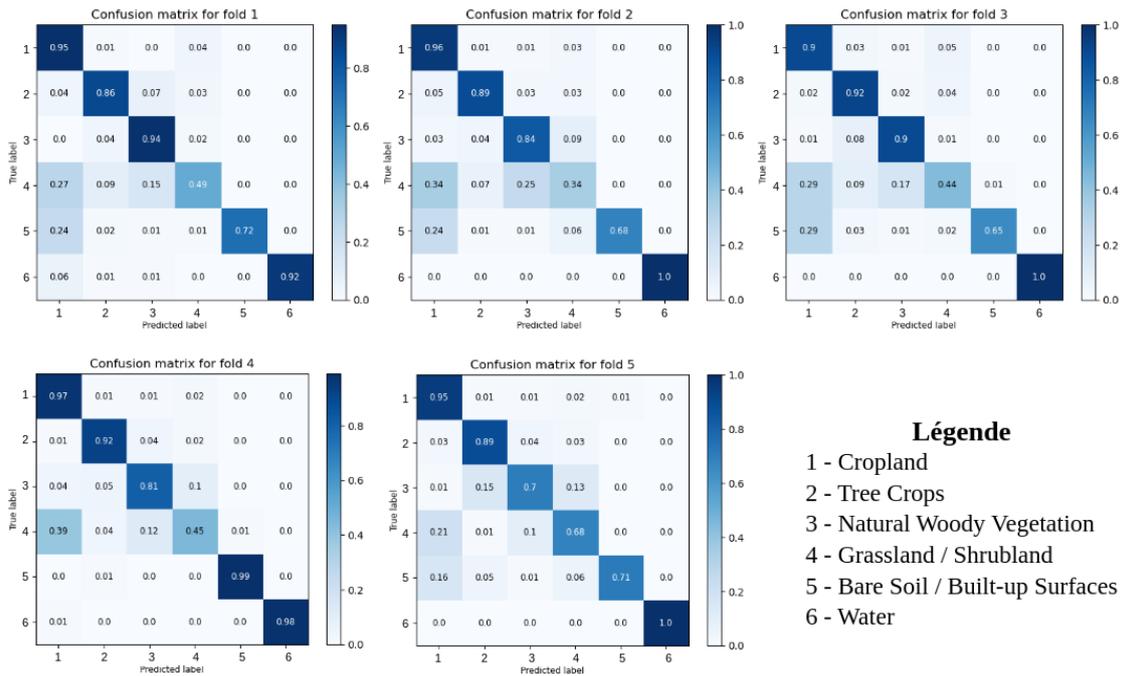
FIGURE 14 – Matrices de confusion par fold pour le modèle *Random Forest*



Average Accuracy : **88.61 %**

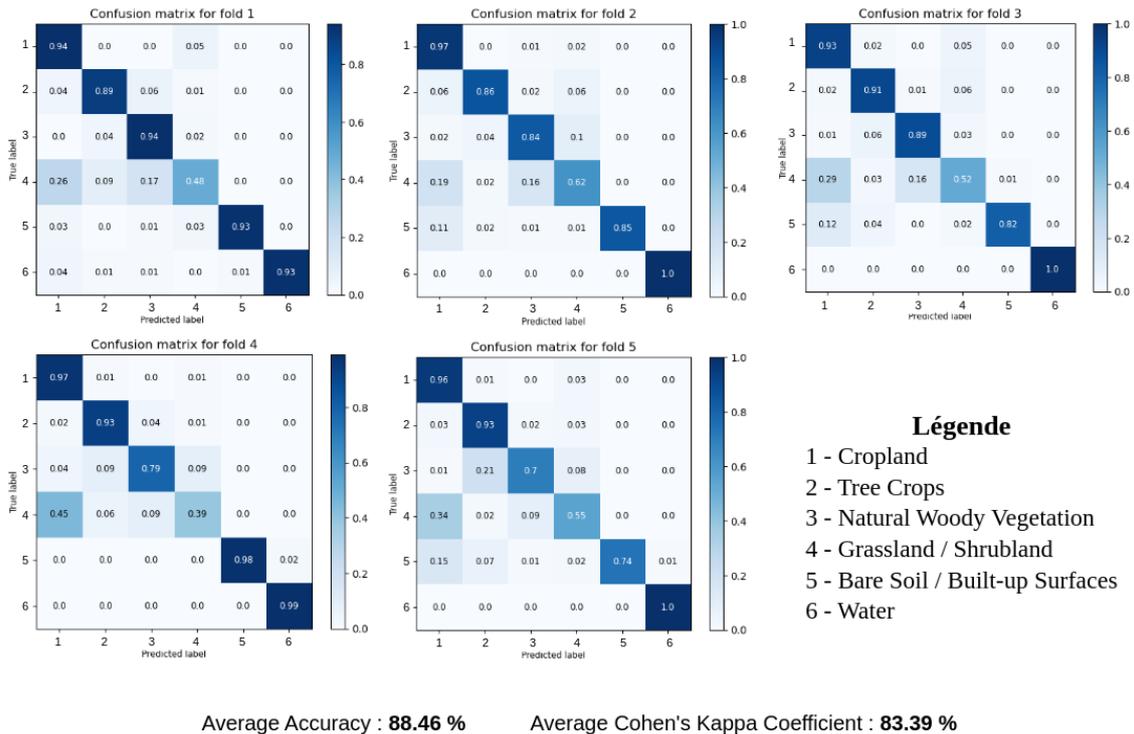
Average Cohen's Kappa Coefficient : **83.62 %**

FIGURE 15 – Matrices de confusion par fold pour le modèle *TempCNN*



Average Accuracy : **87.59 %** Average Cohen's Kappa Coefficient : **82.2 %**

FIGURE 16 – Matrices de confusion par fold pour le modèle *CNN2D*


 FIGURE 17 – Matrices de confusion par fold pour le modèle *DuPLO*

5.2 Étude comparative des modèles testés

5.2.1 Comparaison des Métriques entre les Modèles évalués

Dans cette étude comparative approfondie des modèles de classification de l'occupation du sol, quatre (04) méthodes ont été évaluées en utilisant la précision moyenne et le coefficient Kappa moyen comme mesures de performance. Les résultats obtenus (voir figure 18) révèlent que le modèle *TempCNN* se distingue en obtenant la meilleure précision moyenne (88,61%) et le coefficient Kappa moyen le plus élevé (83,62%), suggérant une capacité supérieure à saisir des motifs complexes et à produire des prédictions en forte adéquation avec les observations réelles. Les modèles *DuPLO* et *CNN2D* suivent de près avec des performances comparables, affichant des précisions moyennes respectives de 88,46% et 87,59%, et des coefficients Kappa moyens de 83,39% et 82,2%. Le modèle *Random Forest* quant à lui affiche les performances légèrement inférieures, atteignant une précision moyenne de 87,3% et un coefficient Kappa moyen de 81,18%.

En somme, on note que, les modèles basés sur les réseaux de neurones convolutifs, en particulier le *TempCNN*, semblent être les plus performants pour cette tâche de classification complexe de l'occupation du sol sur notre site d'étude.

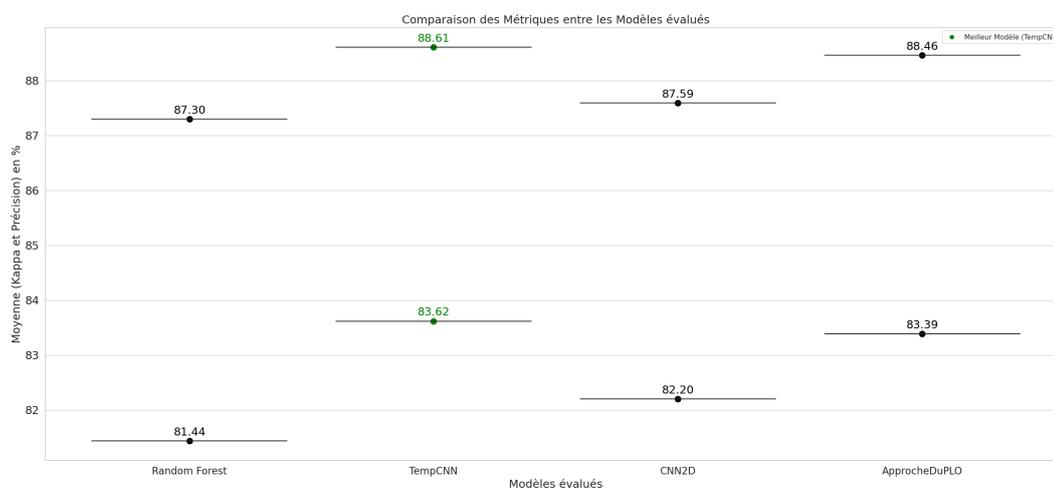


FIGURE 18 – Comparaison des Métriques entre les Modèles évalués

TABLE 6 – Comparaison des modèles

Modèles	Précision moyenne (%)	Kappa moyen (%)
<i>Random Forest</i>	87.3	81.44
<i>TempCNN</i>	88.61	83.62
<i>CNN2D</i>	87.59	82.2
<i>DuPLO</i>	88.46	83.39

5.2.2 Analyse comparative par classe d'occupation du sol

Pour mieux comprendre la capacité de chaque modèle à prédire les classes d'occupation du sol, nous avons également fait une analyse des *f1-score* par classe comme le montre la figure 19 ci - après. Cette analyse des performances des modèles par classe d'occupation du sol montre que le modèle *TempCNN* arrive à mieux prédire les classes '*Cropland*', '*Grassland / Shrubland*' et '*Bare Soil / Built-up Surfaces*' comparativement aux autres modèles. Pour ces classes, ce modèle est suivi des modèles *DuPLO* et *CNN2D*. Cependant, bien que ce modèle présente aussi de bonnes performances pour les classes '*Tree Crops*' et '*Natural Woody Vegetation*', les modèles *DuPLO* et *CNN2D* restent tout de même les meilleurs pour ces deux classes. Aussi, faut - il noter que tous les modèles testés restent à performances égales pour la classe '*Water*'.

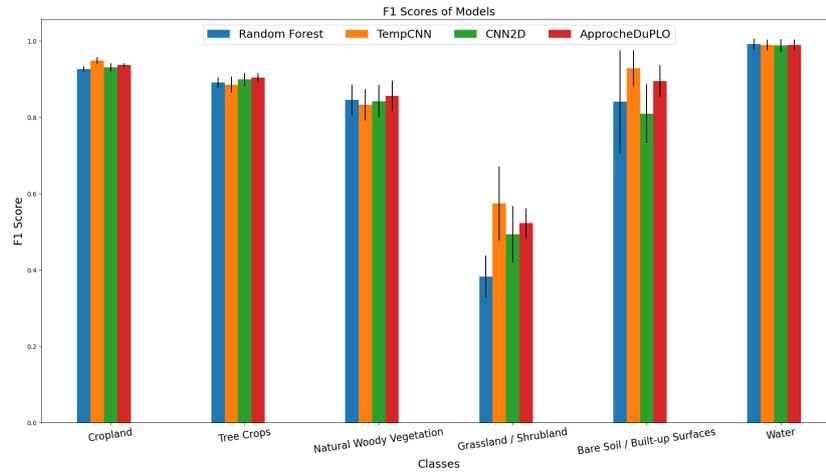


FIGURE 19 – Comparaison des F1 - score des modèles

5.2.3 Analyse visuelle des cartes d'occupation du sol produites

Dans la figure 20, nous apportons quelques détails permettant d'apprécier qualitativement les cartes produites par nos modèles respectifs à la fois en zone urbaine et dans deux différentes zones mixtes.

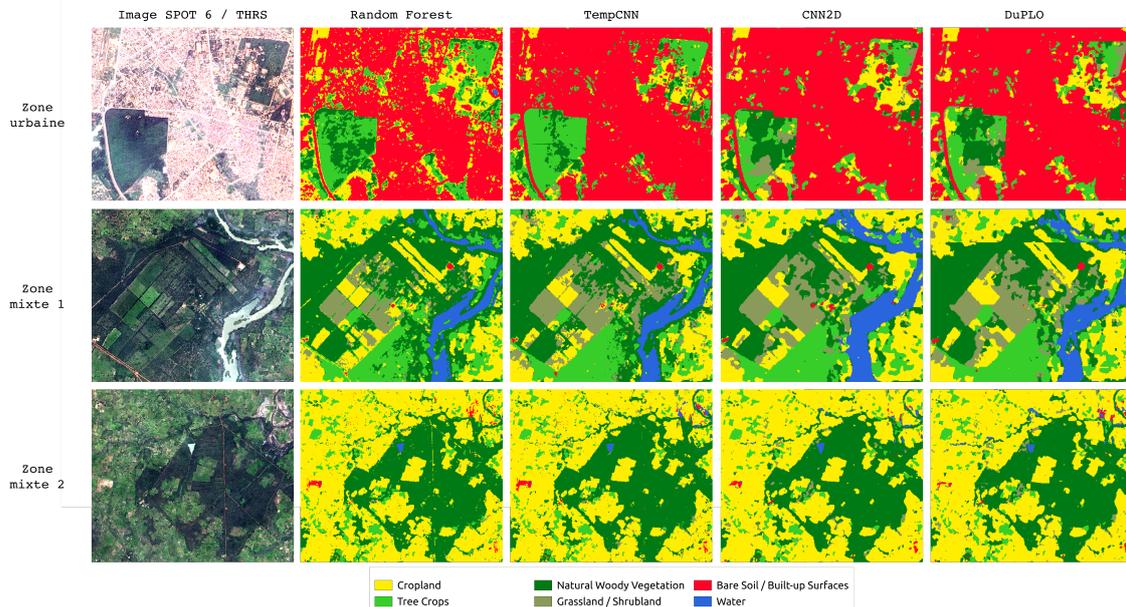


FIGURE 20 – Analyse visuelle des cartes d'occupation du sol produites

De cette analyse visuelle on constate que :

- les méthodes basées sur le *Deep learning* (*TempCNN*, *CNN2D* et *DuPLO*) apportent toutes une amélioration par

- rapport à la *baseline Random Forest*, avec des cartes plus régulières et moins de bruits à l'échelle locale ;
- les méthodes s'appuyant sur les descripteurs spatiaux (*CNN2D* et *DuPLO*) produisent une régularisation plus forte de la carte (elle est plus "lisse"). Ceci est bénéfique pour les classes arborées (plus de texture et de variabilité spatiale) mais peuvent dégrader les détails spatiaux par ailleurs ;
 - *TempCNN* préserve mieux les détails à petite échelle mais présente globalement plus de bruits, ce qui peut rendre plus difficile son exploitation en SIG.

5.3 Discussions et recommandations

5.3.1 Discussions

En s'appuyant sur les résultats de cette étude comparative approfondie des modèles de classification de l'occupation du sol, il est évident que l'application de réseaux de neurones convolutifs, en particulier le modèle *TempCNN*, représente une avancée significative dans le domaine de la télédétection. Avec une précision moyenne de 88,61% et un coefficient Kappa moyen de 83,62%, *TempCNN* établit un nouveau standard en matière de classification d'occupation du sol sur la zone du Nord Bénin, en cohérence avec ce qui est mis en évidence en général dans l'article de référence de PELLETIER, WEBB et PETITJEAN [2]. Cette avancée repose sur la capacité exceptionnelle de *TempCNN* à saisir les motifs complexes présents dans la dimension temporelle des SITS.. Cependant, il est important de noter que d'autres modèles, tels que *DuPLO* et *CNN2D*, ont également présenté des performances solides, ce qui souligne la diversité des approches possibles dans le domaine. L'étude comparative de ces modèles, en parallèle avec des références telles que l'article [3] de R. Interdonato et al. sur *DuPLO*, illustre l'évolution rapide des méthodes de classification d'occupation du sol grâce à l'apprentissage profond. Ces avancées offrent des opportunités prometteuses pour des applications variées, allant de la gestion des ressources naturelles à l'urbanisme, en fournissant des cartographies de l'occupation du sol plus précises et pertinentes, conformément aux recommandations énoncées dans l'article [4] de P. Benedetti et al. sur la fusion de données multimodales et multitemporelles.

De plus, l'analyse des performances par classe d'occupation du sol révèle des informations utiles pour orienter l'utilisation de ces modèles dans des contextes spécifiques. Par exemple, *TempCNN* excelle dans la prédiction des classes '*Cropland*', '*Grassland/Shrubland*', et '*Bare Soil/Built-up Surfaces*', ce qui en fait un choix privilégié pour des projets axés sur ces types d'occupation du sol. Cependant, pour les classes '*Tree Crops*' et '*Natural Woody Vegetation*', *DuPLO* et *CNN2D* restent les modèles de prédilection, ce qui est du fort probablement à la présence plus significative de motifs spatiaux aidant leur discrimination. Ces nuances dans les performances par classe soulignent l'importance de considérer la nature hétérogène des données d'occupation du sol et la diversité des objectifs de cartographie lors du choix du modèle approprié. En fin de compte, cette étude offre une perspective précieuse pour orienter la sélection des modèles dans des applications pratiques de cartographie de l'occupation du sol sur des zones d'études proche de celle analysée (et récurrentes dans le cadre du projet OBSYDYA), en tenant compte à la fois des performances globales et des spécificités des classes d'intérêt.

5.3.2 Recommandations et proposition d'orientations techniques

Au terme de ce travail, nous avons pu découvrir la possibilité d'améliorer les résultats obtenus en explorant de nouveaux axes de recherche dont :

- la variation approfondie des paramètres d'apprentissage, l'augmentation du nombre de couches de convolution ; et pour les méthodes par *patches*, la variation de la taille du *patch* ;
- l'intégration de données provenant de divers capteurs, notamment des images satellites optiques de résolutions différentes (séries temporelles S2 et données à Très Haute Résolution Spatiale) et des données issues de capteur Radar à ouverture de synthèse (SAR) tels que celles acquises par la mission Sentinel-1, ainsi que d'autres données spatialisés qui pourraient conduire à une classification plus robuste de l'occupation du sol en prenant en compte davantage de caractéristiques.

Ces différentes pistes de recherche dont nous venons de faire cas feront l'objet d'un travail plus approfondi qui se fera dans le cadre de la Thèse 1, Résultat 1 du projet OBSYDYA qui portera sur les aspects évoqués.

6 Conclusion et perspectives

Dans ce travail de fin d'étude, il a été question d'évaluer les performances de quatre (04) méthodes d'apprentissage automatique pour la classification de l'occupation du sol dans la zone du nord Bénin en utilisant des jeux de données spécifiques à cette région puis en faire une étude comparative. Cette étude comparative approfondie des modèles de classification de l'occupation du sol a mis en lumière l'efficacité remarquable des réseaux de neurones convolutifs, en particulier le modèle *TempCNN*, dans la cartographie précise de l'occupation du sol à partir de séries temporelles d'images satellites. Les résultats obtenus, avec une précision moyenne de 88,61% et un coefficient Kappa moyen de 83,62%, confirment que *TempCNN* est une référence pour cette tâche complexe de classification de l'occupation du sol basée sur les données spécifiques à notre zone d'étude. Néanmoins, cette étude a également montré que d'autres modèles tels que l'*DuPLO* et le *CNN2D* offrent des performances comparables, soulignant la nécessité de sélectionner le modèle en fonction des spécificités de la tâche et des classes d'occupation du sol concernées. Ces découvertes prometteuses ouvrent de vastes horizons pour l'amélioration de la cartographie de l'occupation du sol et la surveillance environnementale en combinant l'apprentissage profond avec des données de télédétection de haute qualité.

Par ailleurs, il faut noter que ce projet de fin d'étude a permis de mettre en application les connaissances acquises au cours du Mastère Spécialisé SILAT aussi bien dans l'approche méthodologique utilisée que dans le volet conduite de projet. De plus cette étude préliminaire nous a permis d'avoir aussi un premier recul pour la conception de méthodes nouvelles qui feront l'objet d'une thèse de doctorat dans le cadre du projet OBSYDYA.

7 Remerciements

Ce travail a été financé par l'Union Européenne à travers le projet OBSYDYA - Bénin dans le cadre de l'initiative *DeSIRA (Development Smart Innovation through Research in Agriculture)*. Il a été soutenu par l'ensemble de l'équipe projet composée des chercheurs des parties prenantes dont le CIRAD (France), l'INRAB (Bénin), l'Université de Parakou (Bénin), l'IRD (France) et ULiège (Belgique).

8 Références bibliographiques

Références

- [1] J. INGLADA, A. VINCENT, M. ARIAS, B. TARDY, D. MORIN et I. RODES, “Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series,” *Remote Sensing*, t. 9, n° 1, p. 95, 2017. DOI : 10.3390/rs9010095.
- [2] C. PELLETIER, G. I. WEBB et F. PETITJEAN, “Temporal Convolutional Neural Network for the Classification of Satellite Image Time Series,” *Remote Sensing*, t. 11, n° 5, p. 523, 2019. DOI : 10.3390/rs11050523.
- [3] R. INTERDONATO, D. IENCO, R. GAETANO et K. OSÉ, “DuPLO : A DUal view Point deep Learning architecture for time series classificatiOn,” *ISPRS Journal of Photogrammetry and Remote Sensing*, t. 149, p. 91-104, 2019. DOI : 10.1016/j.isprsjprs.2019.01.011.
- [4] P. BENEDETTI, D. IENCO, R. GAETANO, K. OSÉ, R. PENSA et S. DUPUY, “M3Fusion : A deep learning architecture for multiscale multimodal multitemporal satellite data fusion,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, t. 11, n° 2, p. 4939-4949, 2018. DOI : 10.1109/JSTARS.2018.2876357.
- [5] S. J. RUSSELL et P. NORVIG, *Artificial Intelligence : A Modern Approach*. Prentice Hall, 2016.
- [6] A. SHRESTHA et A. H. S. SOLBERG, “Recent advances in deep learning : An overview,” *Neurocomputing*, t. 330, p. 39-48, 2018.
- [7] I. GOODFELLOW, Y. BENGIO et A. COURVILLE, *Deep Learning*. MIT Press, 2016.
- [8] A. VO, F. WALDNER, H. X. CAO et T. T. PHAM, “Deep learning for classification of hyperspectral data : A comparative review,” *IEEE Geoscience and Remote Sensing Magazine*, t. 7, n° 2, p. 82-108, 2019.
- [9] W. LI, H. FU, L. YU, Q. LIU et X. SHAO, “Deep Learning-Based Classification for Land Use and Land Cover Mapping Using Multisensor and Multitemporal Satellite Imagery,” *Remote Sensing*, t. 10, n° 10, p. 1690, 2018.
- [10] X. GENG, H. SHI, X. XIE, J. XU et J. CHEN, “Ensemble Learning Approaches for Land Cover Classification of Multitemporal Satellite Images,” *Remote Sensing*, t. 12, n° 9, p. 1483, 2020.
- [11] S. LI, X. X. ZHU, X. WANG et J. CHEN, “Spatiotemporal Fusion for Land Cover Classification Using Multitemporal Satellite Images,” *IEEE Transactions on Geoscience and Remote Sensing*, t. 57, n° 9, p. 6986-6999, 2019.

- [12] E. MAGGIORI, Y. TARABALKA, G. CHARPIAT et P. ALLIEZ, “Transfer Learning for Land-Use Classification in Multispectral Satellite Images,” *IEEE Transactions on Geoscience and Remote Sensing*, t. 55, n° 7, p. 3, 2017.
- [13] *Theia Land*, URL : <https://www.theia-land.fr/>.
- [14] *Orfeo ToolBox (OTB)*, URL : <https://www.orfeo-toolbox.org/>.
- [15] *Scikit-Learn*, URL : <https://scikit-learn.org/stable/>.
- [16] L. BREIMAN, “Random forests,” *Machine Learning*, t. 45, n° 1, p. 5-32, 2001.
- [17] M. BELGIU et L. DRAGUT, “Random forest in remote sensing : A review of applications and future directions,” *ISPRS Journal of Photogrammetry and Remote Sensing*, t. 114, p. 24-31, 2016.
- [18] G. MOUNTRAKIS, J. IM et C. OGOLE, “Support vector machines in remote sensing : A review,” *ISPRS Journal of Photogrammetry and Remote Sensing*, t. 66, n° 3, p. 247-259, 2011.
- [19] B. ZHAO, H. LU, S. CHEN, Y. LIU et Y. ZHANG, “Convolutional Neural Networks for Time Series Classification,” *Data Mining and Knowledge Discovery*, t. 31, n° 3, p. 606-625, 2017. DOI : 10 . 1007/s10618-016-0483-9.
- [20] *Anaconda*, URL : <https://www.anaconda.com/>.
- [21] *Pytorch*, URL : <https://pytorch.org/>.
- [22] *EODAG*, URL : <https://eodag.readthedocs.io/en/stable/>.

9 Annexes

9.1 Annexe 1 : carte produite avec la *baseline Random Forest*

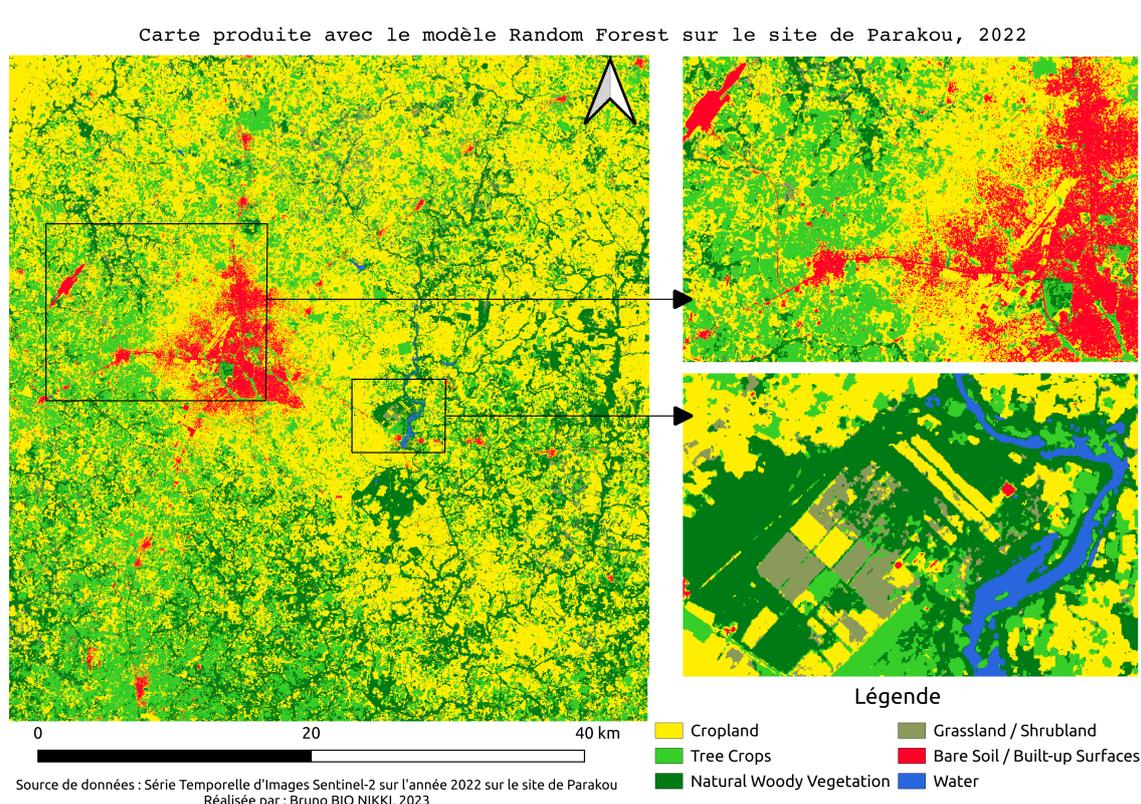


FIGURE 21 – Carte produite avec le modèle *Random Forest* sur le site de Parakou, 2022

9.2 Annexe 2 : carte produite avec le modèle *TempCNN*

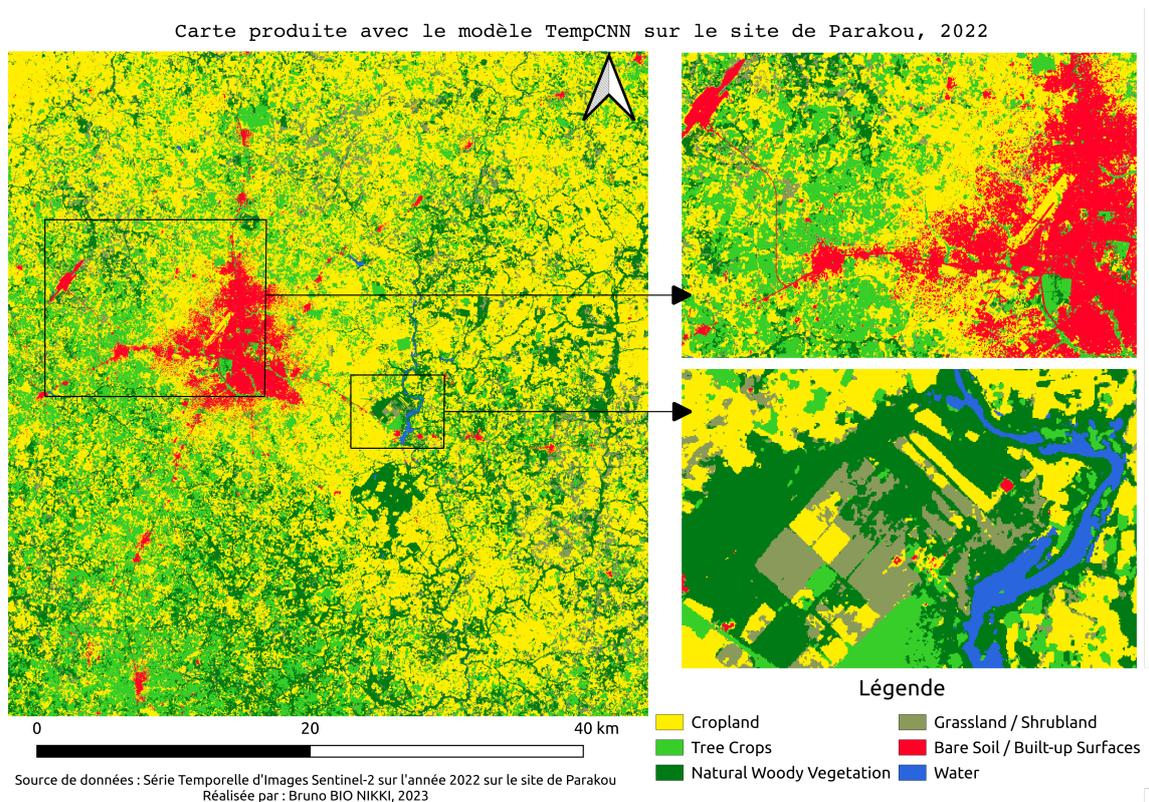


FIGURE 22 – Carte produite avec le modèle *TempCNN* sur le site de Parakou, 2022

9.3 Annexe 3 : carte produite avec le modèle *CNN2D*

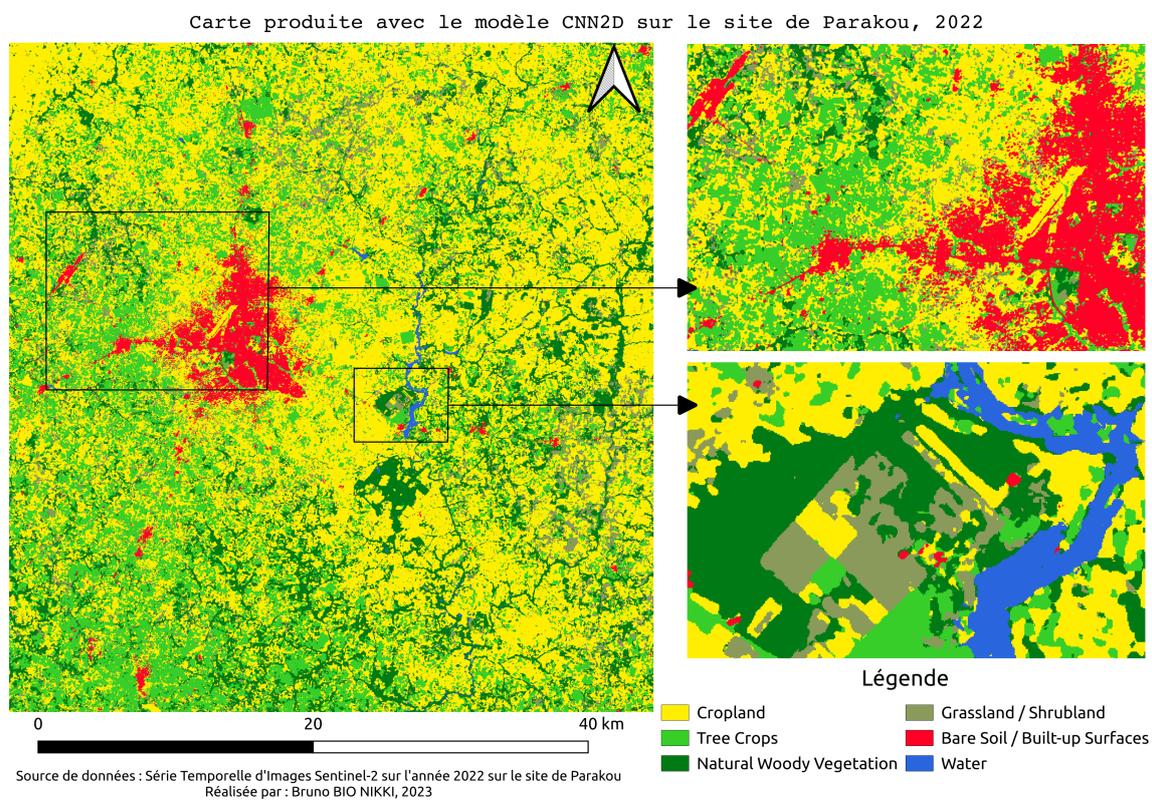


FIGURE 23 – Carte produite avec le modèle *CNN2D* sur le site de Parakou, 2022

9.4 Annexe 4 : carte produite avec le modèle *DuPLO*

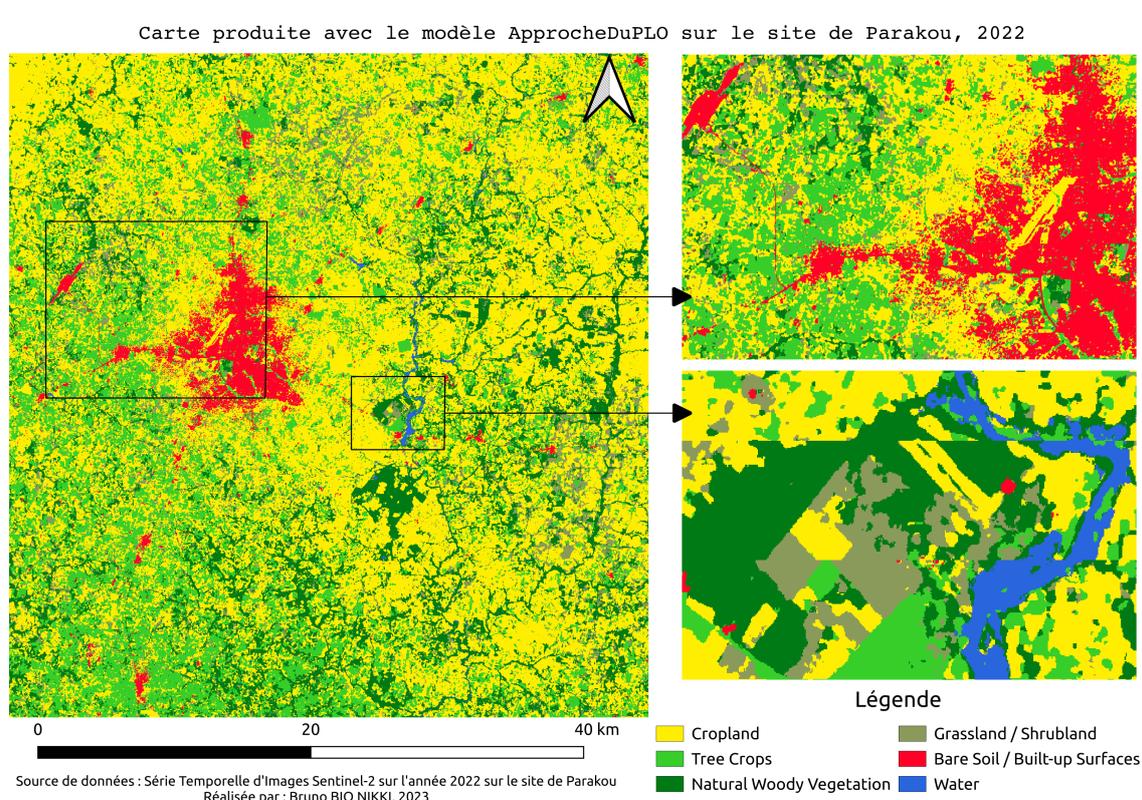


FIGURE 24 – Carte produite avec le modèle *DuPLO* sur le site de Parakou, 2022