

WalousMàJ

Itérations

Version	1
Révision	2
Projet	WalousMàJ
Date	28/09/2021

	Nom	Société	Signature
Auteur	Damien Draime	Aerospacelab	
Relecteur / Approbateur	Jonathan Denies	Aerospacelab	
Approbateur	Benoît Deper	Aerospacelab	<i>Benoît Deper</i>

Table des matières

Table des matières	2
Abréviations et acronymes	4
1. Objet du rapport	5
2. Sélection de l'approche initiale	5
3. Itérations	9
3.1. Itération 01	9
3.1.1. Objectifs	9
3.1.2. Implémentations	9
3.1.3. Observations et interprétations	11
3.1.4. Suggestions	13
3.1.5. Décisions	14
3.2. Itération 02	16
3.2.1. Objectifs	16
3.2.2. Implémentations	16
3.2.3. Observations et interprétations	18
3.2.4. Suggestions	24
3.2.5. Décisions	24
3.3. Itération 3	26
3.3.1. Objectifs	26
3.3.2. Implémentations	26
3.3.3. Observations et interprétations	26
3.3.4. Suggestions	28
3.3.5. Décisions	28
3.4. Itération 4	29
3.4.1. Objectifs	29
3.4.2. Implémentation	29
3.4.3. Observations et interprétations	31
3.4.4. Suggestions	35
3.4.5. Décisions	35
3.5. Itération 5	35
3.5.1. Objectifs	35
3.5.2. Implémentations	36
3.5.3. Observations et interprétations	38
3.5.4. Suggestions	46
3.5.5. Décisions	46
4. Récapitulatif des décisions influençant la solution retenue	48
5. Références	56

6. Annexes	57
6.1. Itération 01	57
6.1.1. Configuration – Implementation I	57
6.1.2. Configuration – Implementation II	63
6.2. Itération 02	64
6.2.1. Configuration – Implementation I	65
6.2.2. Configuration – Implementation II	71
6.2.3. Configuration – Implementation III	78
6.3. Itération 03	85
6.3.1. Configuration	85
6.4. Itération 04	92
6.4.1. Configuration – Pre-entrainement	92
6.4.2. Configuration – Affinement	99
6.5. Itération 05	106
6.5.1. Configuration – Implementation I	106
6.5.2. Configuration – Implementation II	113

Abréviations et acronymes

BN	Batch Norm (normalisation par batch)
CA	Comité d'Accompagnement
ETP	Equivalent Temps Plein
GN	Group Norm (normalisation par groupe)
GPU	Graphical Processing Unit (carte graphique)
LIDAR	Light Detection And Ranging
LPIS	Land Parcel Identification System (Parcellaire agricole anonyme)
PICC	Projet Informatique de Cartographie Continue
MCARI	Modified Chlorophyll Absorption in Reflectance Index
NDMI	Normalized Difference Moisture Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
OA	Overall Accuracy (Précision globale)
OCS	OCcupation du Sol
PA	Producer Accuracy (Précision producteur)
PICC	Projet Informatique de Cartographie Continue
SAR	Synthetic Aperture Radar
SOL_2018	Solution développée dans le cadre du projet WALOUS afin de générer la carte d'occupation du sol de la Wallonie pour 2018
SOL_2019/20	Solution proposée par AEROSPACELAB à l'issue de ce marché
SPW	Service Public de Wallonie
SyncBN	Synchronized Batch Norm (normalisation synchronisée par batch)
TEST_Set	Jeu de données servant à évaluer la solution finale durant la phase III
TRAIN_Set	Jeu de données servant à entraîner les modèles d'une solution intermédiaire
UA	User Accuracy (Précision utilisateur)
VAL_Set	Jeu de données servant à valider une solution intermédiaire durant la phase de prototypage. Inclut VAL_Set_2018 et, potentiellement, VAL_Set_2019/20
VAL_Set_2018	Jeu de données de validation se basant sur les données de 2018
VAL_Set_2019/20	Jeu de données de validation se basant sur les données de 2019 (ou 2020)
WAL_OCS_2018	Carte d'occupation du sol de la Wallonie pour 2018 générée par le projet WALOUS
WAL_OCS_IA_2019/20	Carte d'occupation du sol de la Wallonie pour 2019 (ou 2020) générée pour ce marché
WAL_OCS_CHA_IA_2018_2019/20	Produit vectoriel regroupant les changements survenus entre 2018 et 2019 (ou 2020) et labélisés

1. Objet du rapport

Ce rapport détaille les 5 itérations de la Phase II de Prototypage. Pour chacune de ces itérations, les objectifs et les raisons de celle-ci sont listés. Les solutions intermédiaires implémentées sont présentées ainsi que leur performance. Finalement, la dernière section de ce rapport présente un récapitulatif des différentes décisions prises lors de ces itérations permettant de mettre en lumière l'évolution de la solution durant toute la Phase II.

Pour rappel, la deuxième Phase de ce marché fut, en effet, consacrée à faire évoluer une solution intermédiaire à la suite de multiples itérations successives permettant d'aboutir à une solution répondant au cahier des charges. Le CA fut intégré dans ce processus itératif via les présentations de fin d'itération. Ces dernières, présentant l'évolution de la solution actuelle et ses limites, permettaient d'ouvrir la discussion sur les manières d'améliorer la solution actuelle pour l'itération suivante.

2. Sélection de l'approche initiale

AEROSPACELAB a considéré deux approches pour démarrer la Phase II. D'une part, une approche centrée autour de la détection des changements et de la segmentation de ceux-ci (Approche A), d'autre part, une approche en deux temps segmentant rapidement les zones triviales et affinant les zones plus difficiles (Approche B). La méthode de travail agile, avec laquelle le SPW et AEROSPACELAB ont décidé de travailler, permet de ne pas s'ancrer définitivement au choix de l'approche à explorer initialement. En effet, si l'approche initiale retenue venait à révéler, a posteriori, des faiblesses majeures, alors, avec l'accord du CA, une nouvelle approche aurait été considérée pour la prochaine itération de la Phase II.

L'Approche A repose sur la détection de changements, leur segmentation¹, et finalement leur fusion avec WAL_OCS_2018. Elle est illustrée à la Figure 1. La détection des changements se fait principalement sur base des orthophotos, mais l'utilisation des sources de données additionnelles n'est pas à exclure. Une fois les changements survenus entre 2018 et 2019/2020 identifiés, la segmentation de ceux-ci peut avoir lieu². Plusieurs sources de données peuvent également faire partie de cette étape de segmentation. Finalement, les changements segmentés sont intégrés à WAL_OCS_2018, via l'étape de Fusion, afin d'obtenir WAL_OCS_IA_2019/20.

Cette approche ne fut cependant pas retenue pour 3 raisons principales :

- *Absence de données labélisées pour une approche supervisée.*
Si un modèle supervisé est destiné à la détection des changements alors un large jeu de données labélisées doit être disponible. Or ce n'est pas le cas et le budget de ce marché ne permet pas l'alternative visant à sous-traiter, à un service tiers, la génération d'un tel jeu de données.
- *Pas de support dans la littérature pour une approche non-supervisée.*
Alternativement, un modèle non-supervisé pour la détection des changements peut également être envisagé. Cependant, l'expérience qu'a accumulée AEROSPACELAB dans ce domaine, ainsi que la littérature scientifique, ne préconise pas l'utilisation de ces modèles pour une telle résolution spatiale. En effet, cette tâche de détection de changements est loin d'être triviale, et ce même pour des modèles supervisés, puisque les conditions d'illumination, l'angle de prise de vue, la coregistration, les conditions saisonnières et météorologiques sont autant de facteurs qui peuvent avoir un impact important sur les performances du modèle à une si grande résolution spatiale.
- *Pas de correction des erreurs de WAL_OCS_2018.*
Ce type d'approches ne permet pas de corriger les erreurs de WAL_OCS_2018 puisque, pour qu'une modification soit apportée, il faut d'abord que celle-ci soit liée à un changement qui ait été détecté. Ainsi en prenant l'exemple d'un étang incorrectement

¹ On entend par le terme « segmentation » à la fois les étapes de clustérisations et de classification.

² Il s'agit bien ici de segmentation, et non de simple classification, puisque plusieurs changements de nature différente peuvent être regroupés (e.g. : construction d'un nouveau bâtiment et parking sur un ancien terrain vague).

labélisé sur WAL_OCS_2018 : tant que cette zone n'est pas altérée, cette labélisation ne sera pas contestée. Ainsi l'erreur ne pourra pas être corrigée puisqu'aucun changement n'y est associé³. Cette approche n'a donc de sens que si l'on s'attend à obtenir un modèle sous-performant par rapport à SOL_2018.

La deuxième approche, Approche B, repose quant à elle sur la proposition d'un *candidat-carte OCS* sur base des orthophotos qui est ensuite comparé à WAL_OCS_2018 pour identifier les désaccords (i.e. les changements et erreurs). Ces désaccords sont ensuite segmentés par le *Modèle d'Affinement*, pour finalement être fusionné au *candidat-carte OCS* pour créer WAL_OCS_IA_2019/20. Cette approche, illustrée à la Figure 2, vise par la génération du *candidat-carte OCS* à rapidement segmenter les zones triviales pour ne laisser que les zones litigieuses et difficile à classer au « Modèle d'affinement ». En effet, ce dernier modèle étant plus demandeur en temps de calcul par son utilisation de multiples sources de données pour faire ses prédictions, il est important de filtrer, en amont, ces zones triviales.

La sélection de l'approche initiale s'est faite en concertation avec le CA lors de la réunion de lancement du 19 janvier 2021. Lors de cette réunion, AEROSPACELAB a suggéré d'investiguer l'Approche B lors de la première itération de la Phase II.

Cette approche, suggérée par AEROSPACELAB, se veut la plus proche du cahier des charges du marché tout en prenant en compte l'état de l'art dans les différents domaines concernés.

Premièrement, l'objectif premier est l'implémentation d'un modèle permettant la mise à jour de la carte d'occupation des sols. La notion de mise à jour englobe à la fois la détection des changements et la classification de ceux-ci, mais aussi la détection d'éventuelles erreurs incluses dans la version d'origine (i.e. WAL_OCS_2018) et la reclassification de ces dernières. L'Approche B répond à ces deux exigences puisque, contrairement à l'Approche A, une erreur dans WAL_OCS_2018 est théoriquement remarquée lors de l'étape de comparaison entre WAL_OCS_2018 et le « candidat-carte OCS », et ensuite traitée par le « Modèle d'Affinement »⁴.

Deuxièmement, le cahier des charges stipule que la solution proposée doit être automatisée pour éviter d'éventuelles étapes manuelles telles que la fusion de données ou l'ajustement de paramètres. Cette automatisation peut se faire au détriment de la précision de la solution puisqu'une marge d'erreurs absolue de 5% est tolérée pour autant que celle-ci permette une optimisation de l'automatisation. Ce souhait d'automatisation de la solution a poussé AEROSPACELAB à proposer une solution se reposant sur des modèles de réseaux de neurones convolutifs en apprentissage profond. Cette proposition fut acceptée par le CA lors de cette même réunion de janvier. En effet, ces modèles apprennent d'eux-mêmes à extraire les caractéristiques d'une image leur permettant d'accomplir la tâche pour laquelle ils sont entraînés. Cette faculté n'est pas propre aux anciennes approches de Computer Vision pour lesquelles cette tâche d'extraction de caractéristiques doit être faite manuellement.

De plus, dans le cadre d'un exercice de mise à jour, cette solution devait permettre de corriger les potentielles erreurs de classification de WAL_OCS_2018. Ceci conforte le choix des CNNs puisque ceux-ci sont capables de généraliser au-delà des données fournies lors de leur entraînement et se montrent robustes face aux éventuelles erreurs dans les données d'entraînement.

Allant de pair avec cette volonté d'automatisation, le SPW souhaite également avoir une solution générique qui puisse être réutilisée sur d'éventuelles données futures ou antérieures, ou dans le cadre d'autres marchés. Dès le début, le SPW a montré son intérêt pour une solution qui puisse s'appliquer à la fois sur les données de 2019 et 2020. Or les orthophotos de 2020, prises plus tôt dans l'année que celles de 2018, imposaient l'utilisation d'un modèle capable d'apprendre plusieurs représentations pour la même classe. C'est une qualité que la littérature scientifique ne cesse d'associer aux CNNs.

³ Cet exemple est illustré à la Figure 1.

⁴ Ceci est illustré à la Figure 1 2 par l'exemple de l'étang incorrectement classifié sur WAL_OCS_2018.

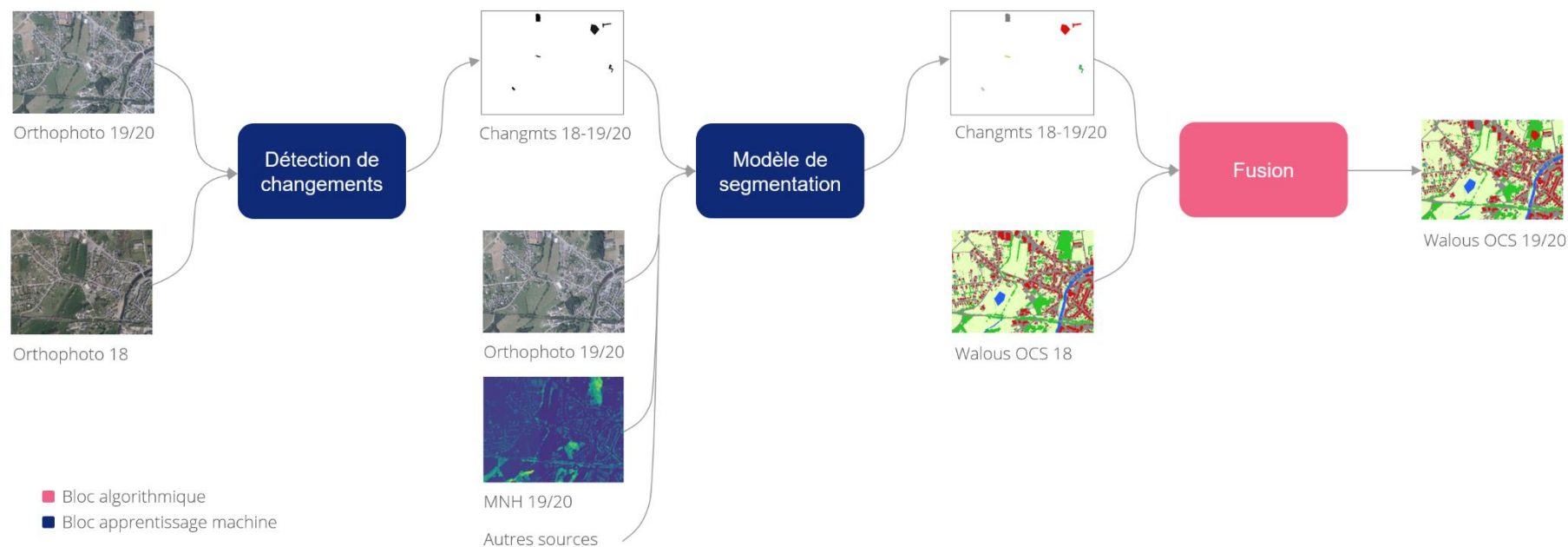


Figure 1 - Approche A : Détection, Segmentation, Fusion

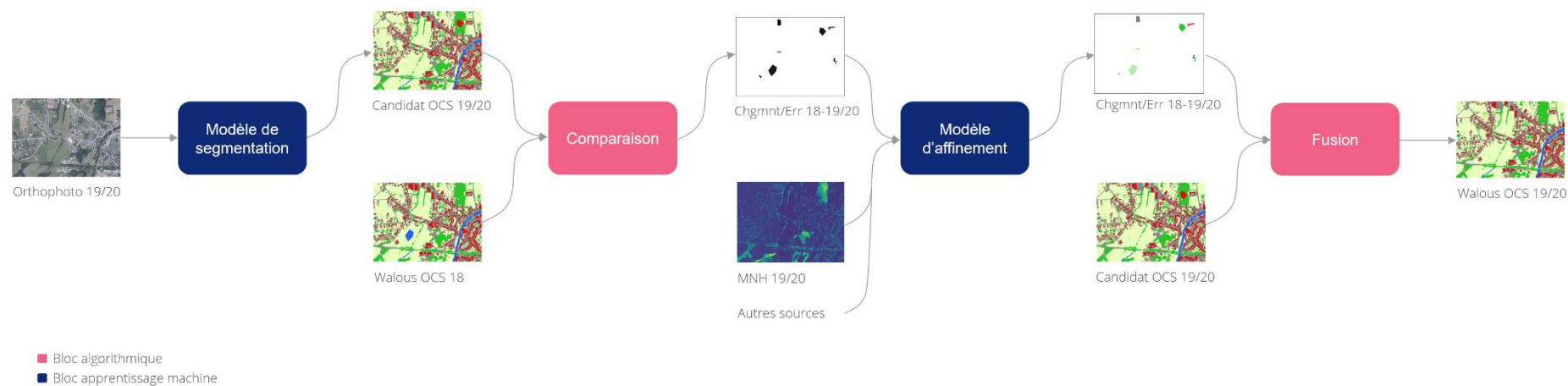


Figure 2 - Approche B : Segmentation, Comparaison, Affinement, Fusion

3. Itérations

3.1. Itération 01

3.1.1. Objectifs

L'objectif principal de cette première itération (Itération 01) fut de tester le « Modèle de segmentation » de l'Approche B. De plus, l'apport des données provenant des images multispectrales de SENTINEL-2 a également été investigué durant cette itération :

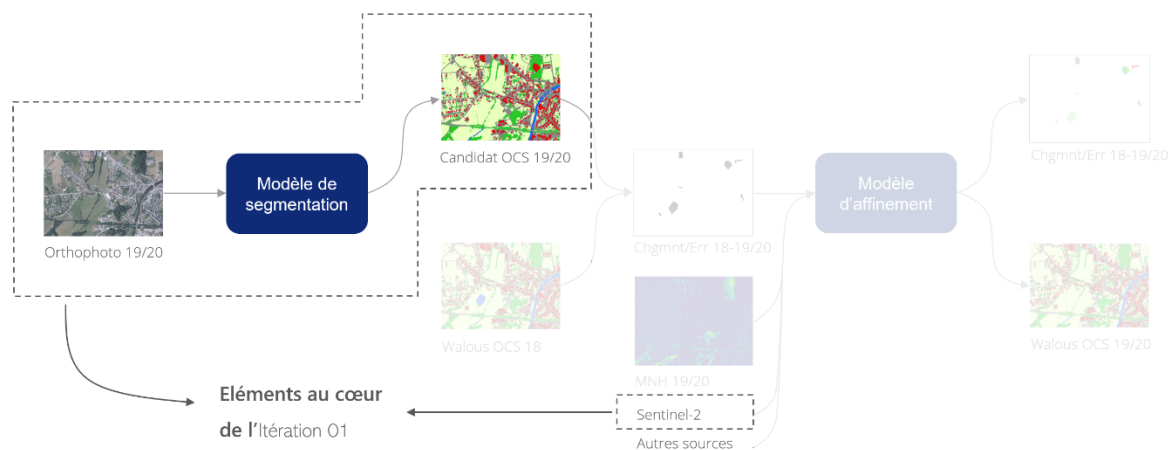


Figure 3 - Itération 01 : Objectifs

3.1.2. Implémentations

3.1.2.1. Implémentation I

Un modèle CNN de segmentation, reposant sur l'architecture DEEPLAB v3+ (Chen, et al. 2017) et acceptant en entrée des vignettes orthophotos de dimensions :

$$512 \times 512 \times 4$$

Equation 1 - Dimensions d'une vignette orthophoto

Où les deux premières dimensions font référence aux dimensions spatiales et la dernière à la dimension spectrale

fût implémenté et entraîné sur les orthophotos de 2018 en utilisant comme réalité-terrain WAL_OCS_2018. Cet entraînement s'est effectué sur 20.000 itérations⁵ de batch de 16 vignettes. Les 4 bandes spectrales ont été utilisées ; Rouge, Vert, Bleu, et Infra-Rouge proche, toutes à la résolution spatiale de 0,25m. Ce modèle fut implémenté dans le framework DETECTRON 2 de FACEBOOK RESEARCH⁶. L'architecture choisie a prouvé son efficacité, et ce pour de nombreuses et diverses applications. Son utilisation répandue apporte aussi l'avantage que des modèles pré-entraînés sont disponibles, réduisant ainsi la taille du jeu de données nécessaire et/ou la durée de l'entraînement afin d'obtenir de bonnes performances. Un autre avantage que cette architecture propose réside principalement dans l'utilisation de convolutions atrous afin d'augmenter le champ de perception des kernels. Ce champ de perception détermine le contexte dans la donnée d'entrée qui sera utilisée pour extraire une caractéristique lors de la convolution. Au plus ce contexte est étendu, au plus la

⁵ Le terme « itération » fait ici référence au passage dans le GPU d'un batch de vignettes

⁶ github.com/facebookresearch/detectron2/tree/master/detectron2

caractéristique extraite peut être pertinente puisqu'elle se base sur plus d'information. Cependant, ceci implique un temps de calcul plus important puisque la convolution s'applique sur une plus grande fenêtre. L'approche des convolutions atrous permet de considérer un contexte plus grand sans pour

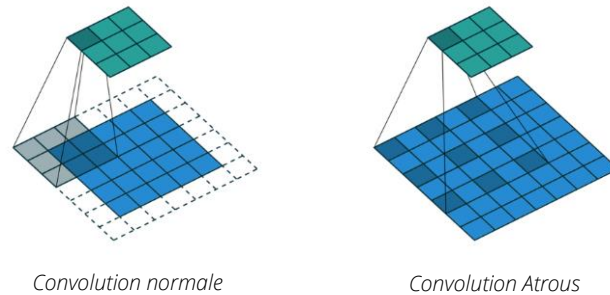


Figure 4 - Types de convolution

autant avoir un impact sur le temps de calcul (Chen, et al. 2017). Ceci est illustré par la Figure suivante : Les détails sur la configuration de ce modèle dans DETECTRON 2 sont proposés en Annexes : 6.1.1

3.1.2.2. Implémentation II

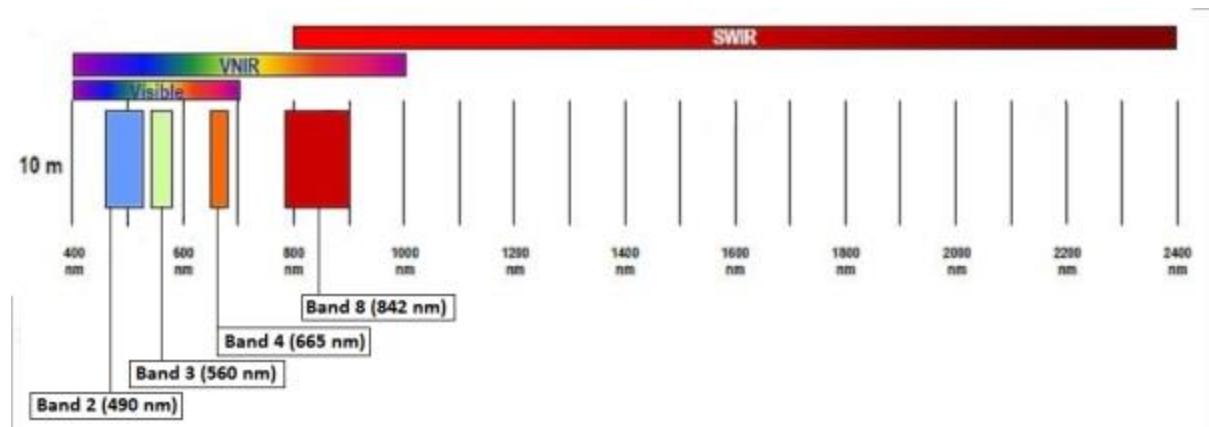
Parallèlement à ce « Modèle de segmentation », un modèle centré autour des données SENTINEL-2 fut également implémenté. Ce deuxième modèle de segmentation est, quant à lui, basé sur l'architecture CONVLSTM afin d'exploiter la temporalité des images multispectrales (Rußwurm and Korner 2018). Ainsi les données d'entrées sont composées de vignettes de dimensions

$$200 \times 200 \times 6 \times 18$$

Equation 2 - Dimension d'une vignette Sentinel-2

Où les deux premières font référence aux dimensions spatiales, la troisième à la dimension spectrale, et la dernière à la dimension temporelle

avec une résolution spatiale de 10m. Les bandes spectrales qui ont été retenues sont les suivantes : B02, B03, B04, B05, B08, et B11.



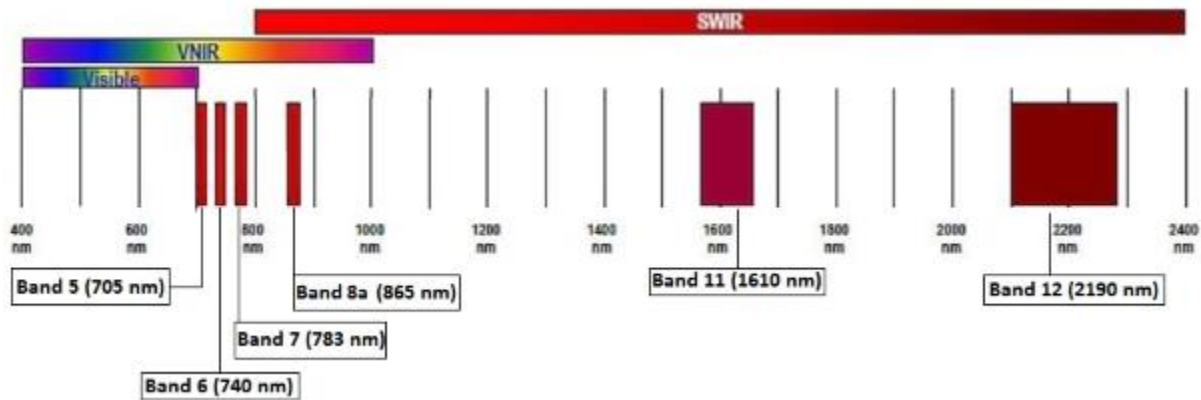


Figure 5 - Sentinel-2 bandes spectrales

Ces bandes, représentant des fenêtres de longueurs d'onde dans les domaines du visible, de l'infrarouge proche et à ondes courtes, ont été choisies puisqu'elles sont les bandes constitutives des indices spectraux relatifs à la végétation, tels que le NDVI, le NDWI, le NDMI, le ARI, et le MCARI. En effet, ce modèle est davantage conçu pour la segmentation des zones naturelles (i.e. les couverts herbacés permanents et en rotation, les feuillus, et les résineux) compte tenu de la résolution spatiale des données ne permettant pas de segmenter de manière précises les petites surfaces qui sont majoritairement artificielles.

La dimension temporelle comporte 18 points sélectionnés bimensuellement de mars à décembre. Les tuiles SENTINEL-2 ont été sélectionnées de manière à choisir, chaque mois, les 2 avec la couverture nuageuse la plus faible. Les produits 2A ont été choisis pour les corrections atmosphériques et radiométriques qu'ils apportent (ESA n.d.). Ces produits sont disponibles depuis mars 2018 ce qui est une des raisons expliquant la couverture temporelle de mars à décembre.

L'entraînement de ce modèle s'est fait sur 25 epochs.

Pour plus de détails sur cette implémentation : Annexes 6.1.2.

3.1.3. Observations et interprétations

3.1.3.1. Implémentation I

L'entraînement du « Modèle de segmentation » permet de révéler, au travers des performances observées, les limites de l'approche considérée. En effet, les 4 bandes spectrales contenues dans les orthophotos ne suffisent pas pour obtenir une première segmentation satisfaisante. Les « candidats-cartes OCS » des sols générés par ce premier modèle contenaient trop de confusions et d'imprécision pour justifier cette approche en deux étapes. Le but premier poursuivit par cette approche est, en effet, de générer des candidats suffisamment cohérents pour rapidement classifier les zones triviales, permettant ainsi au second modèle, plus complexe et exigeant sur le plan des calculs, de se concentrer uniquement sur les cas de désaccords prétendument les plus difficiles à classifier. Or, dans l'état actuel, les candidats déviaient trop par rapport à WAL_OCS_2018, si bien que le « Modèle d'affinement » aurait eu trop de zones à traiter. Les performances du modèle sont illustrées à la Figure suivante grâce aux matrices de confusion calculées sur chacun des jeux de données de validation :

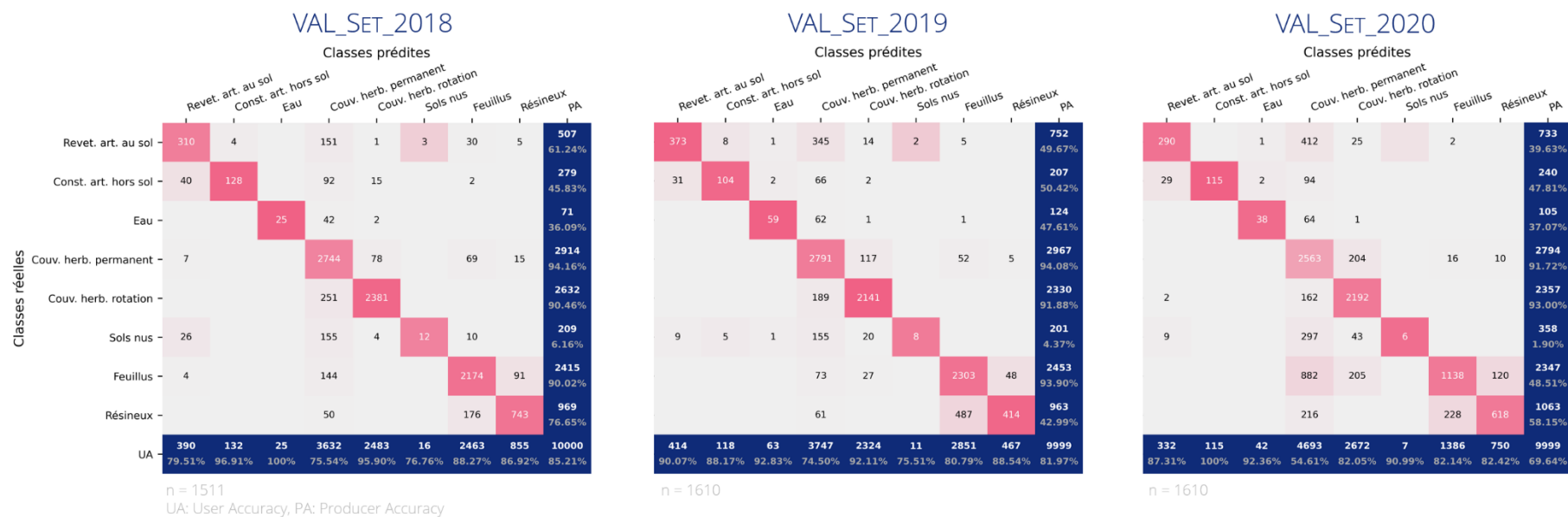


Figure 6 – Itération 01/Modèle de segmentation – Résultats.
Matrices de confusion pour chacun des VAL_Sets (première version). UA : Précision utilisateur. PA : Précision producteur.

De manière générale, on constate que l'UA de ce modèle est bien meilleure que sa PA. Un faible PA implique des taux d'erreurs d'omission élevés, alors qu'un UA élevé traduit des taux d'erreurs de commission faibles. Il peut également être observé que le modèle a tendance à se rabattre sur la prédiction de la classe « Couvert herbacé permanent ». Cela s'explique par la surreprésentation de cette classe dans le jeu de données d'entraînement. Une confusion s'observe pour les 2 classes de « Couverts herbacés », ainsi qu'entre les « Résineux » et les « Feuillus ». Cela était attendu puisque ces classes se distinguent difficilement sur base d'une unique prise de vue. Finalement une nette baisse en termes de précision se constate pour l'année 2020. Cela s'explique par le fait que le modèle est entraîné sur les orthophotos de 2018 dont les conditions saisonnières sont comparables à celles de 2019 mais pas à celles de 2020.

3.1.3.2. Implémentation II

Concernant le modèle ConvLSTM, une précision globale de 81% a été obtenue sur les classes suivantes :

- Couvert herbacé en rotation (classe WALOUS : 6)
- Couvert herbacé permanent (classe WALOUS : 7)
- Résineux (classes WALOUS : 8, 80)
- Feuillus (classes WALOUS : 9, 90)
- Autres (classes WALOUS : 1, 2, 3, 4, 5)

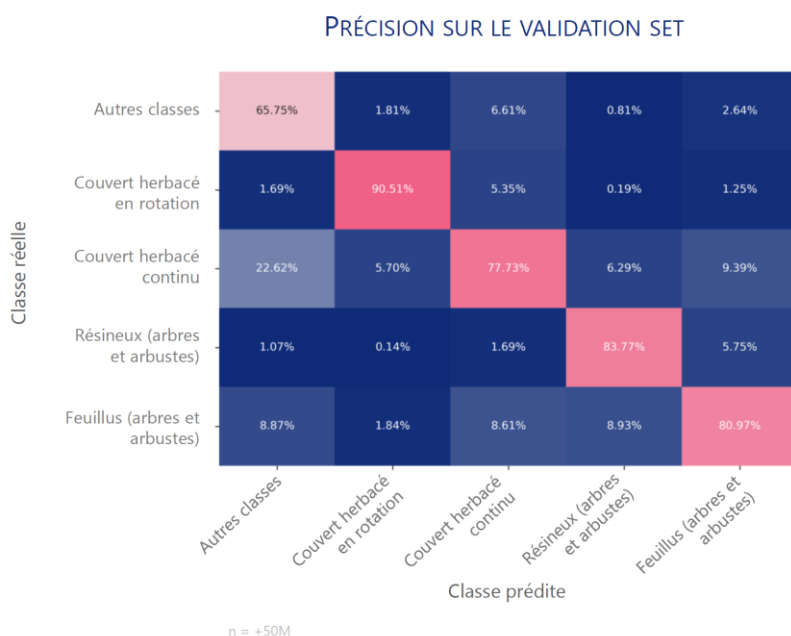


Figure 7 – Itération 01 : Performances du Modèle d'affinement.
Matrice de confusion sur un jeu de données de validation de plus de 50 millions de points propre à cette implémentation utilisant les données SENTINEL-2.

Le lecteur est redirigé vers la présentation de fin d'itération 01 pour plus de détails.

3.1.4. Suggestions

Lors de la présentation de fin d'itération du 2 avril 2021 avec le CA, les suggestions suivantes ont été faites :

1. Proposition 1A : Utilisation des orthophotos de 2019 et 2020 pour entraîner le modèle DEEPLAB V3+ afin de constituer un jeu de données d'entraînement plus représentatif et plus diversifié.

2. Proposition 1B : Concentration de l'entraînement du modèle DEEPLAB V3+ sur les classes ayant subi de fortes perturbations entre 2018 et 2019/2020 afin de diversifier le jeu de données d'entraînement et lui permettre d'apprendre les multiples représentations d'une classe (e.g. : Feuillus avec et sans feuilles).
3. Proposition 1C : Augmentation de la dimension spatiale des vignettes orthophotos utilisées pour l'entraînement du modèle DEEPLAB V3+ afin d'apporter plus de contexte spatial.
4. Proposition 1D : Utilisation des modèles numériques de surface et de terrain comme données d'entrées supplémentaires pour le modèle DEEPLAB V3+, pour faciliter la distinction entre les revêtements artificiels au sol et hors sol, ainsi que la distinction entre les arbres de petite et grande taille.
5. Proposition 1E : Utilisation des informations des relevés LIDAR pour faciliter la détection des surfaces d'eau puisque ceux-ci absorbent les échos.
6. Proposition 1F : Utilisation de squelettes vectoriels thématiques tels que le PICC et le LPIS comme données d'entrée au modèle DEEPLAB V3+ pour contextualiser davantage l'orthophoto.
7. Proposition 1G : Ajout des données de sorties du modèle CONVLSTM comme données d'entrée du modèle DEEPLAB V3+ pour aider le modèle à classer les strates naturelles.
8. Proposition 1H : Nouvelle approche, Approche C, basée sur un seul modèle de segmentation (voir Figure 8).

3.1.5. Décisions

Sachant que les problèmes de précisions aux contours des zones restent un problème récurrent pour les CNNs, l'approche initiale en deux parties ne fut pas retenue puisque les effets de bord continueraient à générer trop de zones à traiter pour le « Modèle d'affinement ». L'Approche C présentée à la Figure 8 fut sélectionnée pour être au cœur de la seconde itération. Les Propositions 1A, 1B, 1D, 1E, et 1G ont également été sélectionnées pour être étudiées lors de la prochaine itération.

Concernant la Proposition 1A, celle-ci n'est pas triviale puisqu'elle repose implicitement sur l'hypothèse que peu de changements sont survenus entre 2018 et 2019/2020 car, lors de l'entraînement, WAL_OCS_2018 restera la réalité-terrain utilisée pour calculer la fonction objectif.

La Proposition 1C n'a pas été retenue puisque le modèle DEEPLAB V3+ ne semble pas souffrir d'un manque de contexte avec ses vignettes d'environ 1,64 hectares.

Quant à la Proposition 1F, celle-ci a été rejetée puisqu'elle impliquerait l'ajout de dépendances externes au modèle impliquant que ces différents squelettes vectoriels soient mis à jour avant de pouvoir considérer la mise à jour de la carte d'occupation des sols.



Figure 8 - Approche C : Segmentation, Comparaison

3.2. Itération 02

3.2.1. Objectifs

Les objectifs de cette seconde itération sont de :

- Valider l'apport des données MNH comme données d'entrée
- Valider l'apport des données Orthophotos de 2020 comme données d'entrée
- Valider l'apport des données résultant du modèle ConvLSTM basée sur SENTINEL-2 comme données d'entrée
- Valider la nouvelle approche centrée autour d'un seul modèle de segmentation

3.2.2. Implémentations

Afin de pouvoir juger de la plus-value apportée par chacune de ces nouvelles sources de données, leur ajout s'est fait graduellement via plusieurs implémentations successives. En greffant ainsi successivement ces nouvelles sources de données à l'architecture de l'Approche C, l'apport individuel de chacune d'elle a pu être déterminé.

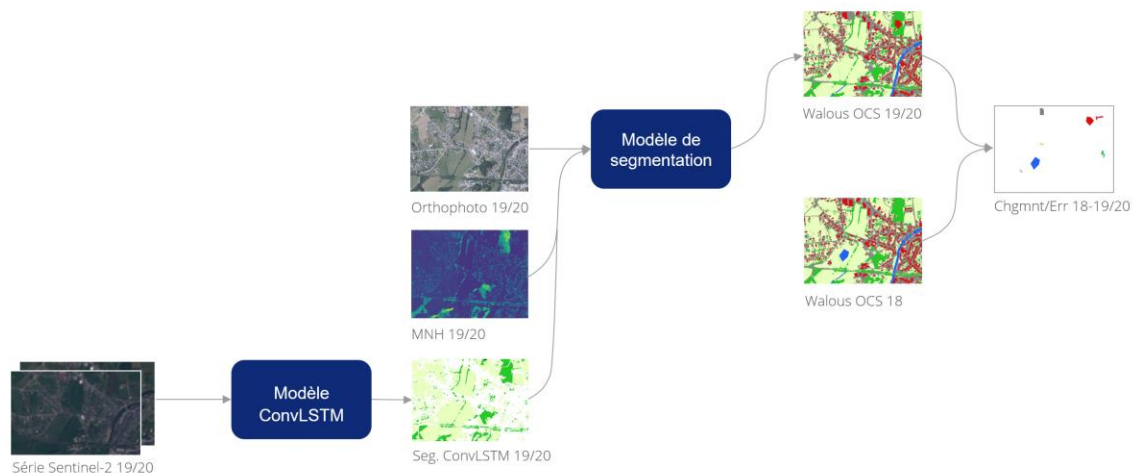


Figure 9 - Itération 02 : Approche C

Trois versions furent donc implémentées :

- Implémentation I : Dans cette version, le « Modèle de segmentation » fut entraîné sur base des orthophotos de 2018 et du MNH correspondant.
- Implémentation II : Cette version se base sur l'Implémentation I et rajoute les orthophotos de 2020 pour l'entraînement.
- Implémentation III : Cette implémentation est construite sur la précédente et rajoute en source d'entrée du « Modèle de segmentation » le résultat du modèle ConvLSTM basé sur les images multispectrales SENTINEL-2.

3.2.2.1. Implémentation I

Afin de valider l'apport des données contenues dans les MNS et MNT, une nouvelle source de données fut rajoutée aux données d'entrée au « Modèle de segmentation ».

Toujours basée sur l'architecture DEEPLAB V3+, ce modèle a été entraîné sur base de vignettes orthophotos de 2018 auxquelles sont venus se greffer les résultats du MNH borné entre -5m et 25m de hauteur. Cette décision a été prise pour plusieurs raisons :

- a) Borner les valeurs aberrantes.
- b) Maintenir des valeurs négatives peut faciliter la détection des carrières qui vont de pair avec la classe « Sols nus ».

- c) Maintenir des valeurs élevées permet de conserver la texture singulière des arbres (canopée).
- d) Garder une bonne précision pour plusieurs fenêtres de hauteurs d'intérêt. En effet, pour cette application, on souhaite une bonne précision aux alentours de zéro pour différencier la classe « Constructions artificielles hors sol » de la classe « Revêtement artificiel au sol ». Une bonne précision au niveau des 3 mètres est également souhaitée pour différencier arbustes et arbres.

Il est important de noter que le MNH pour l'année 2018 découle de la soustraction du MNT de 2013-2014 au MNS de 2018. Outre cette disparité dans les années d'acquisition, les techniques d'acquisition diffèrent également. En effet, alors que le MNT découle de données LIDAR, le MNS de 2018 est, quant à lui, déduit par photogrammétrie. De plus, puisque les données d'entrée doivent partager les mêmes dimensions spatiales, le MNH, dont la résolution spatiale native est de 1m, a été rééchantillonné pour obtenir une résolution spatiale de 0.25m, correspondante à celle des orthophotos.

L'entraînement de ce nouveau modèle poursuit celui effectué lors de l'itération 01. Cela se fait en initialisant les poids du modèle avec ceux obtenus lors de l'itération précédente, lorsque les couches sont compatibles. Cependant, puisque la première couche a été partiellement altérée par l'ajout d'une nouvelle source d'entrée, uniquement les poids relatifs à cette première couche convolutive et liée à cette nouvelle donnée entrante ont été initialisés semi-aléatoirement sur base des autres couches.

3.2.2.2. Implémentation II

La seconde implémentation, construite sur base de la précédente, s'est concentrée sur l'utilisation des orthophotos de 2020 afin d'apprendre au « Modèle de segmentation » les différentes représentations d'une classe résultant des variations de saisons (printemps – été). Ce modèle reçoit toujours en entrée le MNH, mais celui-ci est calculé sur base du MNS de 2018 ou 2020, dépendamment de l'année de référence des orthophotos. Puisque la réalité-terrain reste WAL_OCS_2018, l'utilisation des orthophotos de 2020 s'est faite avec parcimonie pour éviter de perturber l'apprentissage du modèle. Ainsi, pour chaque instance d'entraînement, l'année de référence fut sélectionnée aléatoirement parmi 2018 et 2020. AEROSPACELAB a arbitrairement attaché une probabilité de 1/3 à la sélection de l'année 2020 pour réduire les perturbations d'apprentissage du modèle. En effet, des changements survenus entre 2018 et 2020 impliquent que WAL_OCS_2018, utilisé comme réalité-terrain lors de l'entraînement, dévie de plus en plus de ce qui peut être observé sur les orthophotos de 2020. Or, il est malgré tout demandé à notre modèle de continuer à prédire ce que contient WAL_OCS_2018. La Figure suivante illustre ces dissonances :

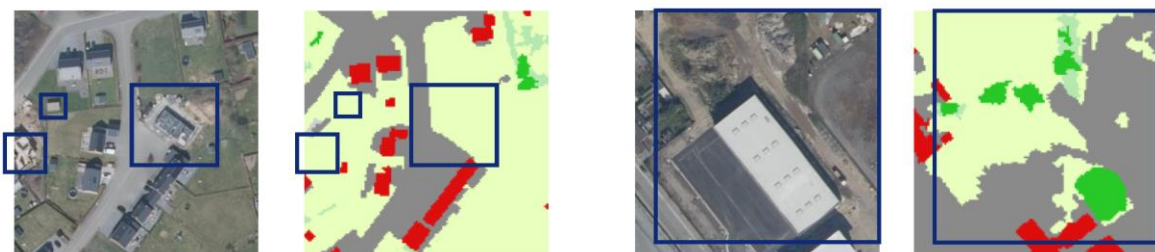


Figure 10 - Dissonances entre les orthophotos de 2020 et WAL_OCS_2018.
Deux exemples de changements survenus entre 2018 (année de référence pour WAL_OCS_2018, à droite) et 2020 (année des orthophotos, à gauche)

L'entraînement de ce modèle a été initialisé avec les poids obtenus à la fin de l'entraînement du modèle de l'implémentation I.

3.2.2.3. Implémentation III

Cette dernière implémentation de l'itération 02 est venu greffer au modèle de l'implémentation précédente le résultat du modèle ConvLSTM entraîné sur les images SENTINEL-2. Ainsi l'apport de ces nouvelles données a pu être étudié en comparant les performances de ce nouveau modèle avec le modèle obtenu à l'implémentation II.

Cette nouvelle donnée d'entrée pour le « Modèle de segmentation » se présente sous la forme d'une unique bande où la valeur du pixel renseigne sur son appartenance à l'une des 5 classes suivantes :

- Couverts herbacés permanents
- Couverts herbacés en rotation
- Résineux (sans distinction de taille)
- Feuillus (sans distinction de taille)
- Autres

Puisque la résolution spatiale des données SENTINEL-2 est nettement inférieure à celle des orthophotos, une étape d'interpolation a eu lieu avant de passer les données de sortie du modèle ConvLSTM au « Modèle de segmentation » afin d'assurer une concordance dans les résolutions spatiales des données d'entrée.

L'entraînement de ce modèle a été initialisé avec les poids obtenus à la fin de l'entraînement du modèle de l'Implémentation II pour les couches identiques, et semi-aléatoirement dans le cas échéant.

Les détails de la configuration DETECTRON 2 de chaque implémentation sont donnés en Annexes : 6.2.1, 6.2.2, et 6.2.3, respectivement.

3.2.3. Observations et interprétations

3.2.3.1. Implémentation I

Cette première implémentation témoigne de l'apport non négligeable du MNH, comme l'illustre la Figure 11.

Alors que pour le modèle de la première itération, l'UA était généralement significativement meilleure que sa PA, ceci ne peut être reproché à ce nouveau modèle. Cette amélioration des performances du PA se traduit par moins d'erreurs d'omission (i.e. un meilleur rappel). Ceci est particulièrement remarquable pour la strate « Eau » et peut être expliqué par les artefacts présents dans le MNH pour les zones appartenant à cette strate. Ceux-ci sont causés, en partie, par l'interpolation du MNT pour les plans d'eau pour lesquelles une acquisition LIDAR n'est pas rendue possible.

On observe également que le modèle commet moins de confusions entre les strates « Constructions artificielles hors sol » et « Revêtements artificiels au sol ». Cela tend à confirmer que le modèle est capable d'extraire les informations pertinentes du MNH. De plus, les erreurs d'omission ont également drastiquement diminué pour ces deux strates.

Néanmoins, on conserve toujours une légère confusion pour les 2 strates de « Couverts herbacés ». De plus, la strate « Couvert herbacé permanent » reste la prédiction par défaut du modèle. Cela était attendu car, a priori, le MNH ne contient pas d'information permettant d'adresser ces deux problèmes.

Le modèle se trouve également très prudent dans ses prédictions d'arbres de petite taille. Cela peut s'expliquer par la présence très limitée de cette classe dans le jeu de données d'entraînement.

Finalement, une nette baisse de l'OA se constate toujours pour l'année 2020. Les erreurs d'omission en hausse pour les « Feuillus » et « Résineux » sont en grande partie responsables pour cette baisse de précision. On constate notamment que les « Feuillus » sont prédits comme « Couverts herbacés permanents » à cause de la perte de leurs feuilles.

Il est à noter également qu'une partie de ces améliorations de performance est à attribuer à la poursuite de l'entraînement du modèle sur base des poids obtenus à la suite de la première itération.

3.2.3.2. Implémentation II

L'Implémentation II prouve que le « Modèle de segmentation » a pu bénéficier de ces nouvelles orthophotos puisque l'OA du modèle sur le VAL_SET_2020 a augmenté de 6% par rapport à l'implémentation précédente. Cette augmentation se traduit principalement par une diminution des erreurs d'omission pour la strate « Feuillus ».

Cette hausse des performances du modèle n'est pas à attribuer à la poursuite de l'entraînement puisqu'on ne constate pas d'amélioration significative en termes d'OA sur les autres VAL_SET (i.e. VAL_SET_2018 et VAL_SET_2019). Les matrices de confusion en Figure 12 détaillent les performances de cette nouvelle implémentation.

3.2.3.3. Implémentation III

Comme l'illustre la Figure 13, cette dernière implémentation n'a pas permis de justifier l'utilisation des données dérivées des images SENTINEL-2 puisque les performances de ce modèle sont sensiblement similaires à celles obtenues à l'Implémentation II.

Ces résultats s'expliquent par la faible résolution spatiale offerte par ces données. Même si elles s'avèrent satisfaisantes d'un point de vue macro, ces données ne permettent pas d'apporter suffisamment de détails à la résolution spatiale imposée par les orthophotos. Ceci est illustré à la Figure 14.

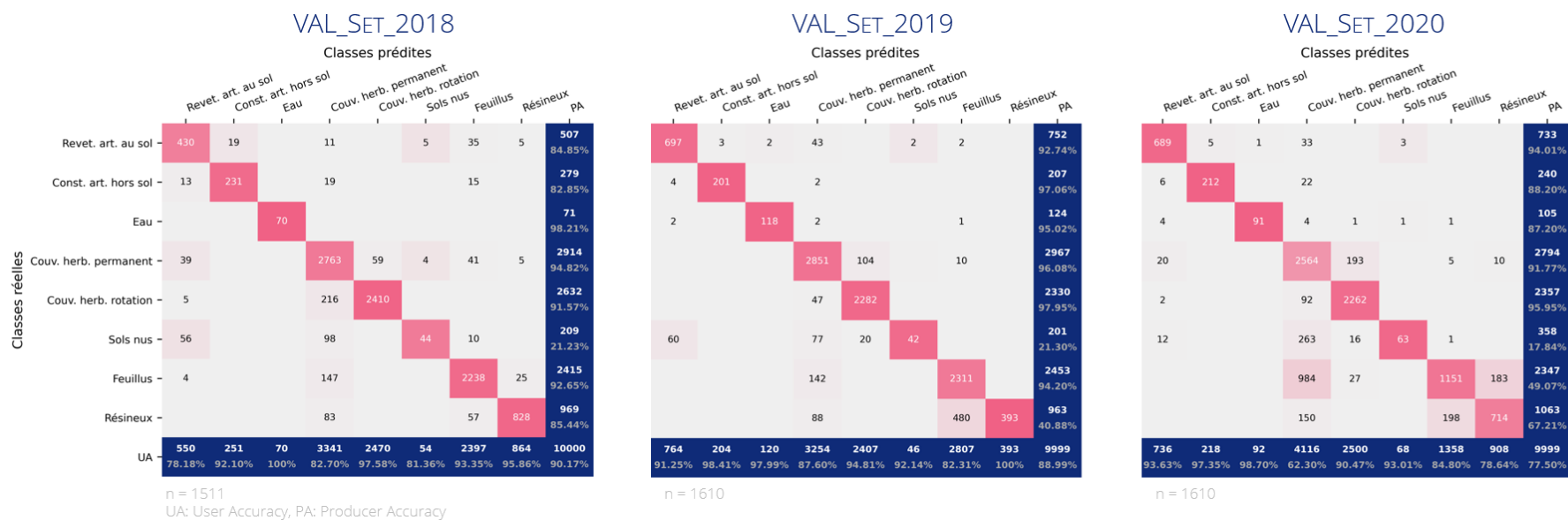


Figure 11 - Itération 02/Implémentation I – Résultats.

Matrices de confusion pour chacun des VAL_Sets (première version). UA : Précision utilisateur. PA : Précision producteur.



Figure 12 - Itération 02/Implémentation II – Résultats.

Matrices de confusion pour chacun des VAL_Sets (contre-validés). UA : Précision utilisateur. PA : Précision producteur.

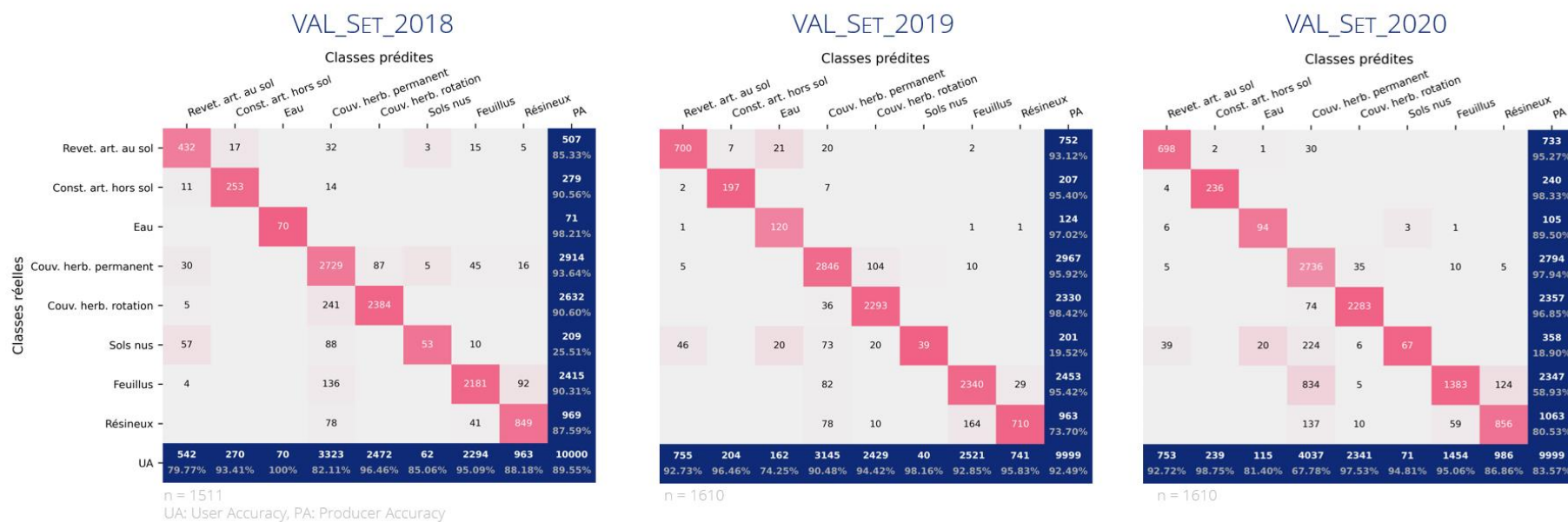


Figure 13 - Itération 02/Implémentation III – Résultats.

Matrices de confusion pour chacun des VAL_Sets (première version). UA : Précision utilisateur. PA : Précision producteur.

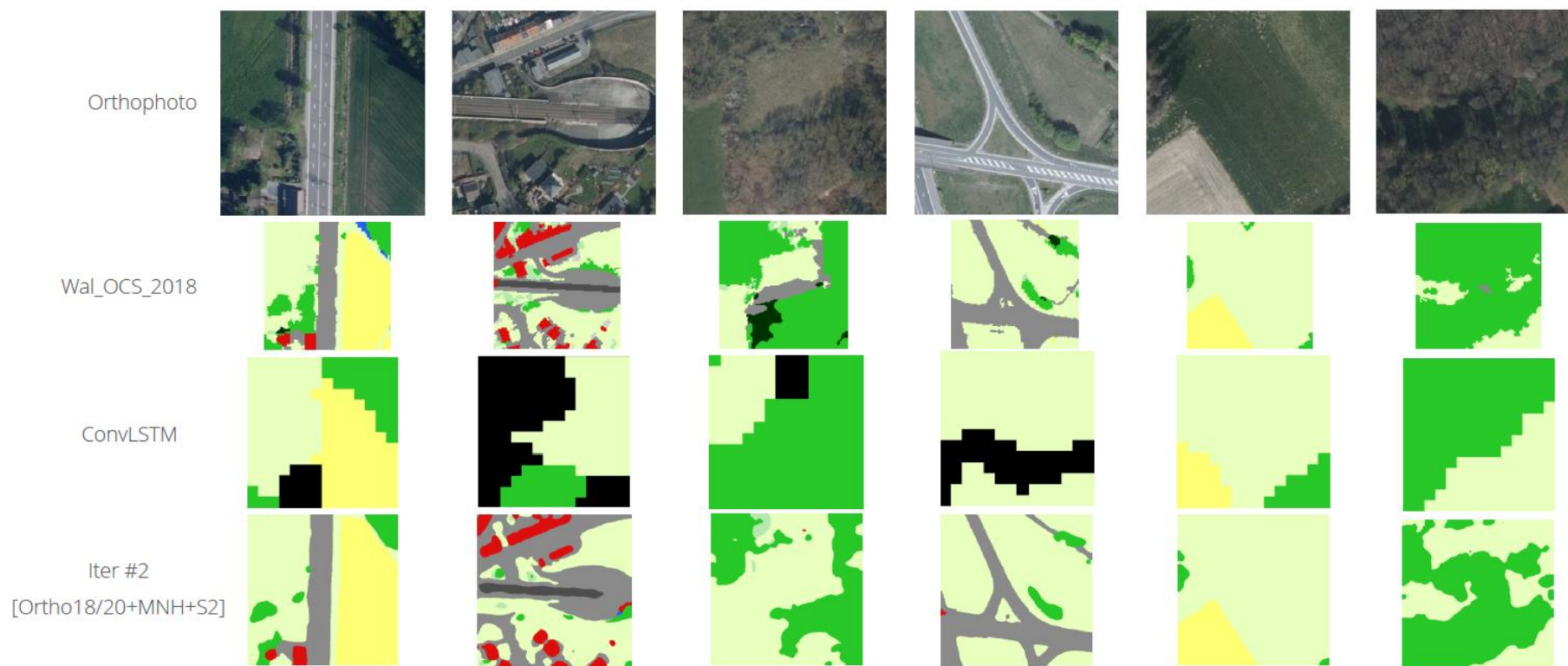


Figure 14 - Itération 02/Implémentation III - Apport du ConvLSTM.

Les données WAL_OCS_2018 sont dans leur résolution spatiale d'origine, à savoir 1m/pixel. Les données ConvLSTM ont un GSD de 10m/pixel propre aux données SENTINEL-2 d'entrée. Les prédictions « Iter #2 » font référence à l'Implémentation III de la deuxième itération et sont affichées avec une résolution spatiale de 0.25m/pixel.

Le lecteur est redirigé vers la présentation de fin d'itération 02 pour plus de détails.

3.2.4. Suggestions

Les résultats obtenus à la fin de cette Itération 02 ont été présentés le 6 avril 2021 devant le CA. Les suggestions suivantes découlent de cette présentation :

1. Poursuite de l'entraînement du « Modèle de segmentation » et échantillonnage adapté des instances d'entraînement pour :
 - 1.1. Proposition 2A : améliorer le rappel de la strate « Feuillus » sur les orthophotos de 2020.
 - 1.2. Proposition 2B : améliorer le rappel de la strate « Eau » pour les grandes étendues d'eau.
 - 1.3. Proposition 2C : améliorer la précision du modèle aux contours de zones.
2. Proposition 2D : Ajout de l'extension POINTREND à l'architecture actuelle du « Modèle de segmentation » pour améliorer la précision du modèle aux contours des zones.
3. Amélioration du rappel pour les strates « Sols nus », « Revêtements artificiels au sol », et « Couverts herbacés permanents », via :
 - 3.1. un affinement des données réalité-terrain (i.e. WAL_OCS_2018), via :
 - 3.1.1. Proposition 2E : un effort manuel
 - 3.1.2. Proposition 2F : l'utilisation du PICC
 - 3.1.3. Proposition 2G : les données vectorielles des carrières actives
 - 3.2. Proposition 2H : une intégration de l'historique comme données d'entrée
 - 3.3. Proposition 2I : une utilisation des données HSL (HARMONIZED LANDSAT SENTINEL-2) afin de déterminer l'humidité/perméabilité du sol
 - 3.4. Proposition 2J : une utilisation des données SAR de SENTINEL-1 afin de déterminer la consistance du sol
 - 3.5. Proposition 2K : une détermination des signatures spectrales pour ces strates
4. Proposition 2L : Implémentation de règles basées sur l'entropie pour affiner les prédictions faites par le modèle.
5. Proposition 2M : Utilisation du MNT 2013-2014 LIDAR sans interpolation pour améliorer le rappel de la strate « Eau ».
6. Proposition 2N : Exploitation de la densité des réponses LIDAR du MNT 2013-2014 pour faciliter la distinction entre une surface artificielle, généralement plus lisse et donc caractérisée par un écho net et précis, et un sol nu, souvent plus irrégulier et donc associé à un écho plus bruité.
7. Proposition 2O : Utilisation du LPIS pour améliorer la précision du modèle dans ses classifications de « Couverts herbacés permanents » et « Couverts herbacés en rotation ».
8. Proposition 2P : Exclusion de l'entraînement des zones où la confiance du modèle est faible.
9. Proposition 2Q : Distinction entre les coupes à blanc et les sols nus.
10. Proposition 2R : Utilisation de la couche vectorielle des 3 classes pour affiner les données réalité-terrain (i.e. WAL_OCS_2018) notamment pour les « Résineux », « Feuillus » et coupes à blanc.

3.2.5. Décisions

La décision principale prise à la suite de cette Itération 02 fut de poursuivre l'entraînement du modèle de l'implémentation II afin d'identifier les faiblesses persistantes de celui-ci. Ainsi, puisque l'apport des données SENTINEL-2 via le modèle ConvLSTM, ne s'est pas révélé concluant, la branche axée autour de cette source de données n'a pas été retenue. Dès lors, l'approche a évolué pour aboutir à celle illustrée en Figure 15.

Cette poursuite de l'entraînement du « Modèle de segmentation » adresse donc partiellement les Propositions 2A-C. Il a été décidé de n'envisager la Proposition 2D qu'après avoir identifié les faiblesses du modèle à la fin de son apprentissage. Cette décision fut prise en accord avec le souci du SPW de s'approprier la solution. En effet, si l'itération 03 confirme que le modèle a pu s'améliorer dans sa délimitation des zones, alors il convient d'éviter de complexifier le modèle par l'ajout de l'extension POINTREND.

La décision quant à la Proposition 2E a été reportée à la prochaine itération afin de mieux évaluer la charge de travail à consentir une fois que les faiblesses du modèle seront identifiées à la fin de son apprentissage.

La Proposition 2F a été rejetée puisque SOL_2018 s'est basée sur les données du PICC. De plus, cette proposition suppose que les données du PICC sont à jour. Finalement, son utilisation est loin d'être triviale puisque l'angle de vue des orthophotos implique des disparités entre son empreinte réelle au sol et celle basée sur la photo-interprétation.

La Proposition 2G a été retenue, mais fera d'abord l'objet d'une évaluation de la qualité de cette couche vectorielle.

La Proposition 2H a été rejetée par crainte d'overfitting. Il est également possible que ce type de données pousse le modèle à ne considérer que celle-ci puisqu'elle serait très proche de ce qui est attendue en sortie. Cela impliquerait que le modèle détecte moins, voire plus du tout, les changements car ceux-ci représentent en principe une infime minorité des données et seraient donc considérés comme valeurs aberrantes par le modèle. De plus, toujours suivant cette hypothèse, ce modèle ne serait pas non plus capable de corriger les erreurs car elles seraient également présentes dans l'historique. Finalement, le plus gros point d'interrogation reste l'obtention de ces données historiques et la confiance qu'on y accorde.

Les Propositions 2I-J n'ont pas été retenues à cause de la résolution spatiale qu'elles offrent.

Les Propositions 2K-L n'ont également pas été retenues car leur utilisation impliquerait l'ajout de règles qui ne sont pas forcément robustes pour l'ensemble du territoire wallon et pour diverses saisons.

De manière générale, les Propositions 2E-K ne constituent pas une réponse adéquate puisque selon Aerospacelab, les problèmes de confusion entre les strates « Sols nus », « Couverts herbacés permanents » et « Revêtements artificiels au sol » n'est pas lié à un manque d'information mais plutôt à un jeu de données d'entraînement qui n'est pas consistant. Ainsi, rajouter des sources de données supplémentaires ne règlera pas le problème car cela n'empêchera pas le modèle d'être injustement pénalisé durant son apprentissage. Les inconsistances de WAL_OCS_2018 forcent le modèle à associer deux classes différentes à la même représentation d'un objet.

Les Propositions 2M-N ont été rejetées à cause du fait que les données LIDAR ne sont disponibles que pour 2013-2014.

Concernant, la Proposition 2O n'a également pas été retenue pour deux raisons principales :

- Cette utilisation du LPIS comme source de données additionnelle rajouterait une dépendance. Ceci implique donc qu'avant de considérer la mise à jour la carte d'occupation des sols, la nouvelle version du LPIS devra être disponible.
- L'ajout de cette nouvelle source de données pourrait pousser notre modèle à ne plus tenir compte des sols nus présents dans les champs (e.g. : en bordure de champs ou à proximité des abreuvoirs).

Bien que l'utilisation de la fonction objectif soit plus adaptée que la « confiance » du modèle pour la Proposition 2P, celle-ci n'a pas été retenue puisqu'elle revient à ne pas entraîner le modèle sur les zones où il éprouve des difficultés. Et puisque le modèle devra tout de même se prononcer sur ces zones difficiles durant la phase d'inférence, ne pas l'entraîner sur celle-ci revient à reléguer le problème à plus tard.

La Proposition 2Q a été retenue. Cependant, AEROSPACELAB précise que des données labélisées doivent être fournies ou dérivables d'autres sources pour pouvoir entraîner un modèle à faire cette distinction entre coupes à blanc et « Sols nus ». En réponse, la Proposition 2R a été suggérée afin de dissocier les coupes à blanc de sa strate d'origine. De plus, cette couche vectorielle des « 3 classes » pourraient également être utilisée pour améliorer la segmentation des forêts dans WAL_OCS_2018. Cette proposition a donc été retenue pour la prochaine itération.

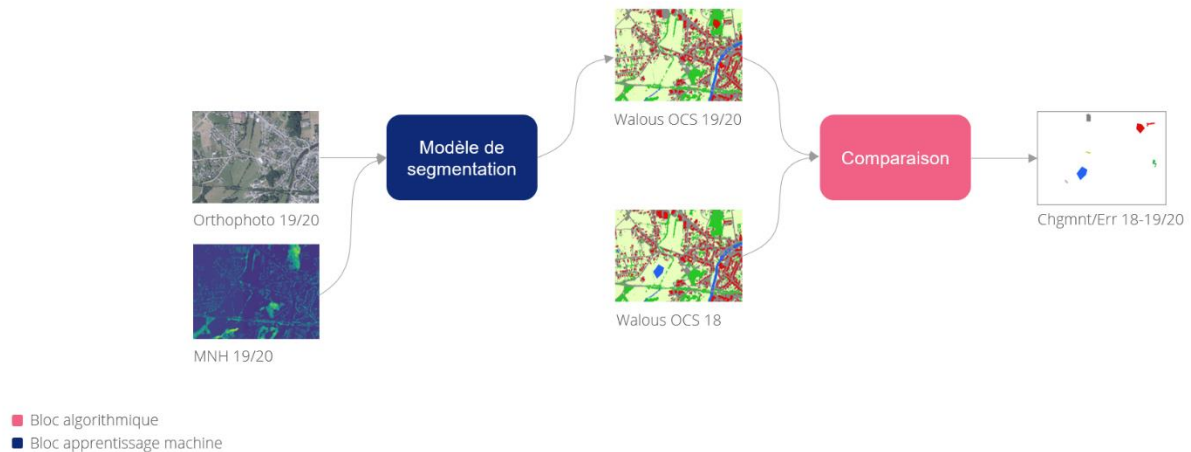


Figure 15 - Approche D

3.3. Itération 3

3.3.1. Objectifs

Les premières itérations ont servi à déterminer les sources de données nécessaires afin d'obtenir un modèle avec de bonnes performances. Ces itérations rapides ont permis de rapidement tester l'apport de chacune de ces sources. Cependant, le modèle n'a pu être entraîné très longtemps suite à ces adaptations régulières. Par conséquent, l'objectif principal de cette itération fut d'étudier les limites de l'Approche D, et plus spécifiquement du « Modèle de segmentation », basée sur les orthophotos et le MNH, en poursuivant son entraînement pendant plusieurs epochs.

3.3.2. Implémentations

Aucune nouvelle implémentation n'a été mise en place. Le modèle sélectionné utilise en données entrantes les orthophotos (2018 et 2020) et le MNH correspondant. A noter que, comme pour l'itération précédente, l'année de référence fut sélectionnée aléatoirement pour générer les instances d'entraînement en associant une probabilité d'un tiers à l'année 2020, et ce pour les mêmes raisons qu'évoquées plus haut.

L'entraînement du modèle a été poursuivi sur base des poids obtenus à la fin de l'entraînement du modèle Implémentation II de la deuxième itération. 175.000 itérations sont venues parfaire l'entraînement de ce modèle, pour un total de 200.000 itérations. Pour rappel, à chaque itération, ce sont 16 instances d'entraînement qui sont vues par le modèle.

Les détails de cette implémentation sont fournis en Annexes 6.3.1.

3.3.3. Observations et interprétations

Concernant la poursuite de l'entraînement du modèle, il demeure qu'à la fin de son apprentissage, le modèle peine toujours à précisément définir les bordures de zones artificielles sur une résolution spatiale de 0.25m/pixel et à distinguer précisément certaines strates problématiques.

Classes réelles	Classes prédites								PA
	Revet. art. au sol	Const. art. hors sol	Eau	Couv. herb. permanent	Couv. herb. rotation	Sols nus	Feuillus	Résineux	
Revet. art. au sol	448	17	19	1	15	5			507 88.45%
Const. art. hors sol	11	247	17		2				279 88.66%
Eau			70						71 98.21%
Couv. herb. permanent	35		2737	64	9	50	16		2914 93.93%
Couv. herb. rotation	5		167	2459					2632 93.42%
Sols nus	84		59		56	10			209 26.74%
Feuillus	4		136			2181	92		2415 90.31%
Résineux			78			36	854		969 88.12%
UA	590	265	70	3217	2523	67	2296	968	10000 75.90% 93.28% 100% 85.09% 97.44% 83.44% 95.00% 88.24% 90.56%

n = 1511
UA: User Accuracy, PA: Producer Accuracy

Classes réelles	Classes prédites								PA
	Revet. art. au sol	Const. art. hors sol	Eau	Couv. herb. permanent	Couv. herb. rotation	Sols nus	Feuillus	Résineux	
Revet. art. au sol	716	13	41	2	2				777 92.13%
Const. art. hors sol	2	217	7						227 95.82%
Eau			102		1	1			105 97.69%
Couv. herb. permanent			2932	118		27			3078 95.24%
Couv. herb. rotation			79	2202		27			2309 95.35%
Sols nus	25		56		14	1			97 14.71%
Feuillus			69			2305	16		2391 96.41%
Résineux			47			235	729		1012 72.03%
UA	743	231	102	3233	2320	19	2602	746	9999 96.34% 94.03% 100% 90.67% 94.88% 75.20% 88.60% 97.77% 92.21%

n = 1710

Classes réelles	Classes prédites								PA
	Revet. art. au sol	Const. art. hors sol	Eau	Couv. herb. permanent	Couv. herb. rotation	Sols nus	Feuillus	Résineux	
Revet. art. au sol	666	4	13	5					690 96.52%
Const. art. hors sol	15	217							232 93.50%
Eau	9		96				1		107 90.15%
Couv. herb. permanent	2		2831	78		5			2918 97.00%
Couv. herb. rotation			95	2228		22	15		2362 94.33%
Sols nus	11		190		44	2			249 17.65%
Feuillus			1	886		1441	78		2408 59.88%
Résineux			178			64	786		1029 76.41%
UA	705	223	98	4195	2313	44	1537	881	9999 94.43% 97.52% 98.77% 67.48% 96.33% 100% 93.79% 89.20% 83.14%

n = 1710

Figure 16 - Itération 03 – Résultats.

Matrices de confusion pour chacun des VAL_Sets (contre-validés). UA : Précision utilisateur. PA : Précision producteur.

Le manque de précision aux contours de surfaces artificielles s'explique par le fait que les données réalité-terrain découlent d'une résolution spatiale de 1m/pixel pas suffisamment précise pour obtenir une délimitation parfaite des zones artificielles sur une résolution spatiale de 0.25m/pixel.

En ce qui concerne la Proposition 2G, après une évaluation de la qualité de la couche vectorielle « Carrières actives », les problèmes suivants ont été relevés :

- Le polygone délimitant la carrière englobe non seulement la fosse d'extraction mais également les bâtiments d'exploitation, les zones inondées, etc.
- Uniquement les carrières actives sont reprises dans cette donnée.

Il découle de ces constatations que cette donnée ne peut pas être utilisée sans une intervention humaine.

La Proposition 2R visait à considérer la donnée « 3 classes » afin de dissocier les coupes à blanc de la strate « Sols nus » et de raffiner la segmentation des « Feuillus » et « Résineux ». Cependant, une étude préliminaire révèle que cette donnée n'est pas exempte de tous défauts également. On notera par exemple :

- Une délimitation entre feuillus et résineux pas toujours plus précise que WAL_OCS_2018.
- Une classe « Non ligneux (coupe récente) » trop permissive pour notre cas puisque la coupe ne doit pas avoir forcément eu lieu l'année précédente.
- Une absence de distinction entre les arbres de petite et grande taille, or cette information ne peut pas directement découler du MNH car ce dernier n'est pas toujours fiable.

Le lecteur est redirigé vers la présentation de fin d'itération 03 pour plus de détails.

3.3.4. Suggestions

Lors de la réunion de fin d'itération 03 du 9 juin 2021, AEROSPACELAB a réitéré ses deux propositions faites à l'itération précédente : Propositions 2D-E.

De même, la Proposition 2O a été réitérée.

Une nouvelle proposition, Proposition 3A, a été faite autour de la génération d'une carte regroupant les « zones complexes » sur lesquelles le modèle n'aurait pas à faire de segmentation. Ces zones pourraient être extraites sur base de la carte montrant les désaccords entre les prédictions du modèle résultant de l'itération 03 et celles de SOL_2018 (i.e. WAL_OCS_2018), et de l'avis d'un expert en photo-interprétation. L'avis de cet expert est important pour éviter d'inclure dans ces zones de simples désaccords entre « Feuillus » et « Résineux » ou encore entre les deux « Couverts herbacé ». De même, les contours doivent également être mis de côté. Une fois ces zones identifiées, le modèle ne devrait pas se prononcer sur celles-ci.

3.3.5. Décisions

Puisque le modèle éprouve encore des difficultés à définir avec précision les bordures de zones, la Proposition 2D a été choisie pour être investiguée lors de la prochaine itération. Ceci est principalement pertinent pour les classes correspondantes à des zones artificielles dont les délimitations sont plus nettes par rapport aux zones naturelles où les transitions de classes sont généralement plus graduelles.

Il a également été décidé de consacrer 2 ETP pendant un mois, à l'affinement de WAL_OCS_2018. Ainsi AEROSPACELAB et le SPW ont décidé d'allouer chacun un ETP à cette tâche dans le but de rendre plus consistant la réalité-terrain utilisée lors de l'entraînement du modèle afin de rendre son apprentissage plus uniforme. Des exemples d'inconsistances dans WAL_OCS_2018 sont illustrées à la Figure 17. Ainsi la Proposition 2E a été retenue pour être conduite lors de l'itération 04.

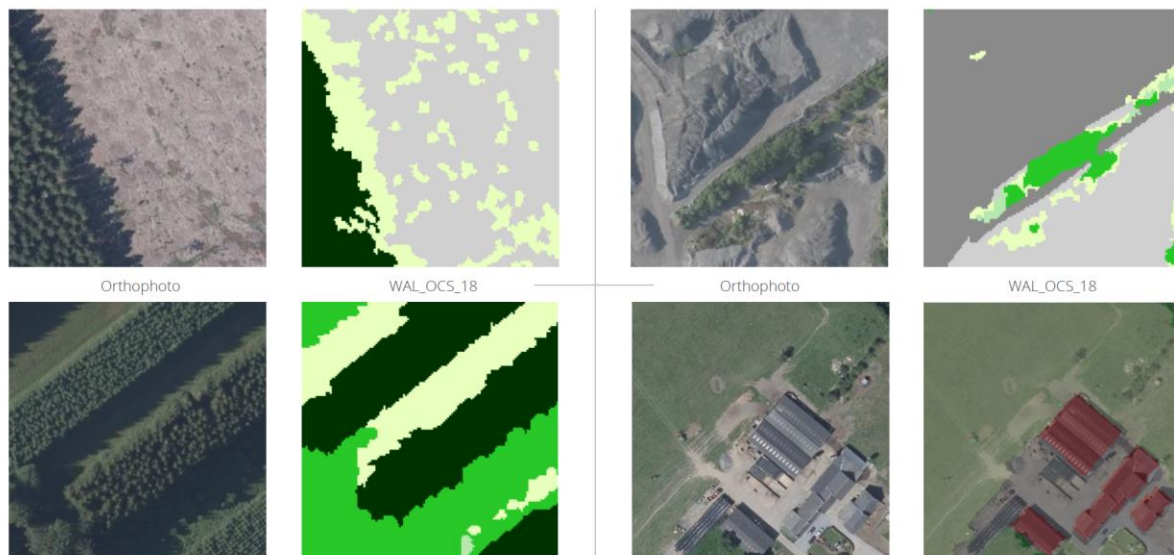


Figure 17 - Exemples d'inconsistances dans WAL_OCS_2018.

Haut à gauche : non uniformité dans la classification de surfaces homogènes. Haut à droite : non uniformité des carrières. Bas à gauche : Impacts des ombres sur la segmentation. Bas à droite : Décalage (vers le haut) de l'empreinte du bâtiment par rapport à l'orthophoto.

Pour les mêmes raisons qu'évoquées lors de la réunion de fin d'itération 02, la Proposition 20 a été, une nouvelle fois, rejetée.

Finalement, la nouvelle Proposition 3A a été écartée puisque l'expert en photo-interprétation devrait parcourir l'ensemble du territoire wallon afin de générer cette donnée de « Zones complexes ».

3.4. Itération 4

3.4.1. Objectifs

Deux objectifs encadrent cette quatrième itération :

1. Etudier l'impact de l'extension POINTREND sur la capacité du modèle à proposer des contours plus affinés pour les zones artificielles.
2. Poursuivre l'entraînement du modèle sur des mailles de WAL_OCS_2018 qui ont été affinées manuellement afin de proposer de meilleures conditions d'apprentissage pour le modèle.

3.4.2. Implémentation

L'extension POINTREND (Kirillov, et al. 2020) proposée par des chercheurs du FACEBOOK AI RESEARCH en 2020 permet d'affiner la précision d'un modèle de segmentation aux frontières de classes par une stratégie d'échantillonnage de points sur lesquels des prédictions de segmentation vont être faites (« point-based segmentation predictions »). La stratégie d'échantillonnage des points se veut itérative afin de sélectionner préférentiellement les points susceptibles d'être à la frontière entre plusieurs classes.

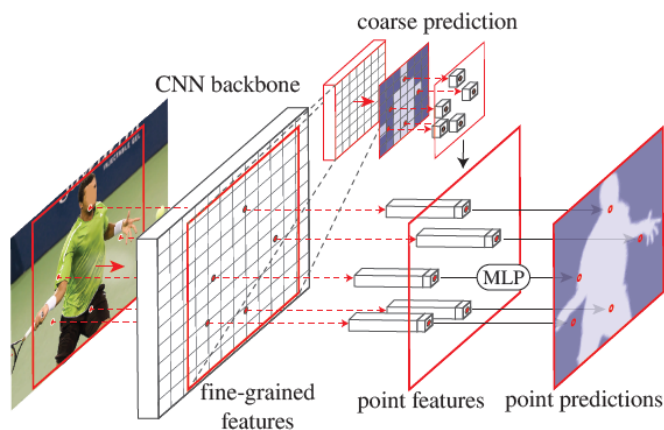


Figure 18 - Illustration de l'extension POINTREND

L'extension POINTREND étant déjà implémentée dans le framework utilisé (i.e. : DETECTRON 2), AEROSPACELAB a pu rapidement prototyper une solution intermédiaire basée sur cette extension.

Dans un premier temps, l'extension a été entraînée sur l'ensemble des mailles, en parallèle du travail d'affinement, afin de ne pas biaiser l'apprentissage du modèle. Les poids d'un modèle pré-entraîné furent néanmoins utilisés pour initialiser celui-ci mais en prenant en compte l'ajout des deux nouvelles bandes dans les données d'entrée. Dans un second temps, l'entraînement se poursuivra sur les mailles affinées et la plus-value de cette extension sera étudiée afin de constater si elle permet d'améliorer la délimitation des zones artificielles. Les configurations DETECTRON 2 sont disponibles en Annexes 6.4.1 et 6.4.2.

La tâche d'affinement visait à manuellement corriger les incohérences et erreurs de WAL_OCS_2018 en travaillant par maille. 40 mailles ont été initialement sélectionnées et l'attribution a été divisée entre le SPW et AEROSPACELAB. En tout, ce sont plus de 8 personnes qui ont contribué à cet effort et la Phase d'Affinement s'est étalée sur environ 2 mois. Cependant, le temps estimé pour affiner une maille a été sous-estimé et AEROSPACELAB a préféré mettre l'accent sur la qualité plutôt que la quantité. En effet, étant donné que le but est de « fine-tuner » les prédictions du modèle, il est préférable d'entraîner ce dernier sur un nombre restreint des données de bonne qualité que sur un plus grand nombre de données mais d'une qualité discutable. Ainsi sur l'ensemble de mailles affinées, 19 ont été jugées de bonne qualité et ont donc été retenues pour faire partie de la première vague d'entraînement. 4 de ces mailles se sont basées sur les données de 2020 et le reste sur les données de 2018. Les mailles de 2020 ont été spécifiquement sélectionnées afin de s'assurer qu'elles couvrent une zone forestière contenant des feuillus sans feuilles. En effet, puisque la solution intermédiaire présente des difficultés à segmenter correctement ce type de zones, l'accent a été mis sur l'affinement de ces dernières. La Table 1 liste les mailles ayant été retenues pour la première vague :

Mailles 2018	ORTHO_2018__00480074
	ORTHO_2018__00510089
	ORTHO_2018__00650129
	ORTHO_2018__00780084
	ORTHO_2018__00290097
	ORTHO_2018__00660127
	ORTHO_2018__00740088
	ORTHO_2018__00180122
	ORTHO_2018__00600087
	ORTHO_2018__00780023
	ORTHO_2018__00560141
	ORTHO_2018__00660121

Mailles 2020	ORTHO_2018__00720116
	ORTHO_2018__00210125
	ORTHO_2018__00350102
	ORTHO_2020__00400097
	ORTHO_2020__00250110
	ORTHO_2020__00490118
	ORTHO_2020__00410112

Table 1 - Mailles affinées / Première vague

L'effort consenti par les personnes ayant pris part à cette tâche d'affinement a été bénéfique à plus d'un sens. Outre l'amélioration du jeu de données d'entraînement, cette étape leur permet également de mieux comprendre les limites des deux modèles : celui ayant généré WAL_OCS_2018 et la solution intermédiaire de l'itération 03. En effet, afin d'aiguiller cette tâche d'affinement, un masque binaire mettant en évidence les désaccords entre les deux modèles a été fourni pour chaque maille. Les désaccords notoires repris dans ce masque ont été affinés en priorité.

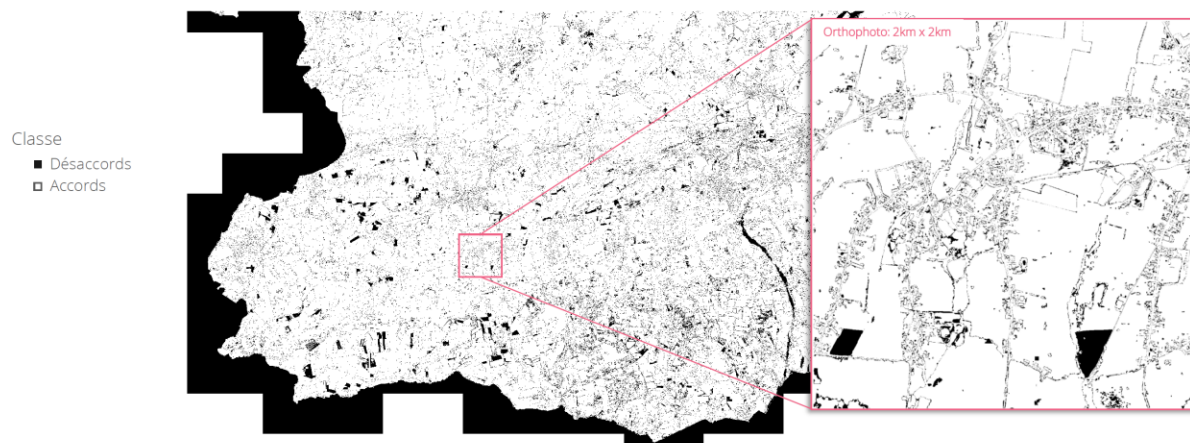


Figure 19 - Tâche d'affinement / Masque binaire de désaccords

En zones rurales et naturelles, l'affinement d'une maille peut prendre 1 jour. Cependant, cette même tâche peut prendre jusqu'à 3 jours dans les zones urbaines où les désaccords sont plus petits et diversifiés.

3.4.3. Observations et interprétations

Les résultats obtenus à la suite de cette itération montrent une nette augmentation du rappel de la strate « Sols nus ». L'écart entre le rappel de cette strate sur VAL_SET_2018 et celui sur VAL_SET_2019 s'explique par le fait que la majorité des 19 mailles affinées de la première vague ont été réalisées par les mêmes personnes qui ont photo-interprété et contre-validé VAL_SET_2019. Ainsi, puisque leur interprétation de la strate « Sols nus » diverge de celle ayant menée à labéliser VAL_SET_2018, il est normal de constater une différence significative entre les rappels sur ces deux jeux de données de validation. En effet, la réalité-terrain (i.e. : WAL_OCS_2018), correspondant aux 19 mailles, a été affinée de telle manière à ce qu'elle corresponde à l'interprétation des classes de la personne en charge de la tâche et le modèle a graduellement été entraîné à internaliser cette interprétation.

Une autre amélioration constatée est l'augmentation du rappel sur les arbustes de moins de trois mètres. Cependant, une confusion entre les deux classes est souvent observée. La petite taille de ces feuillus et résineux n'aidant pas à faire la comparaison.

Parallèlement, le rappel de la strate « Eau » a diminué par rapport aux résultats obtenus à l'itération précédente. Ainsi, il a, par exemple, été observé que des zones d'eau boueuses ont été, par endroits, confondues avec des champs labourés.

De plus, aucune amélioration notable sur la délimitation des contours des zones artificielles n'a été constatée. Et la performance du modèle sur le VAL_SET_2020 n'a pas été améliorée.

Aucune amélioration notable n'a été constatée concernant la précision du modèle dans sa délimitation des classes artificielles lorsque la segmentation est analysée avec une résolution spatiale de 0.25m. Cela peut s'expliquer par le fait que, lors de la première phase d'entraînement, le modèle a été entraîné à reproduire la segmentation de WAL_OCS_2018 à une résolution spatiale de 0.25m. Or le rééchantillonnage de cette donnée réalité-terrain n'offre pas une base d'apprentissage suffisamment précise pour l'extension POINTREND. En effet, le rééchantillonnage ne permet pas d'améliorer la précision de WAL_OCS_2018 aux frontières de classes. De plus, les décalages observés sur cette donnée entre l'empreinte cadastrale des bâtiments (dérivée des données du PICC) et celle sur l'orthophoto peut entraîner des confusions lors de l'entraînement du modèle (voir Figure 21). Finalement, la deuxième partie de l'entraînement du modèle s'est faite sur un jeu de données avec une réalité-terrain affinée mais de taille restreinte.

La performance du modèle à l'issue de cette itération est renseignée dans les matrices de confusion de la Figure 20.

Le lecteur est redirigé vers la présentation de fin d'itération 04 pour plus de détails.

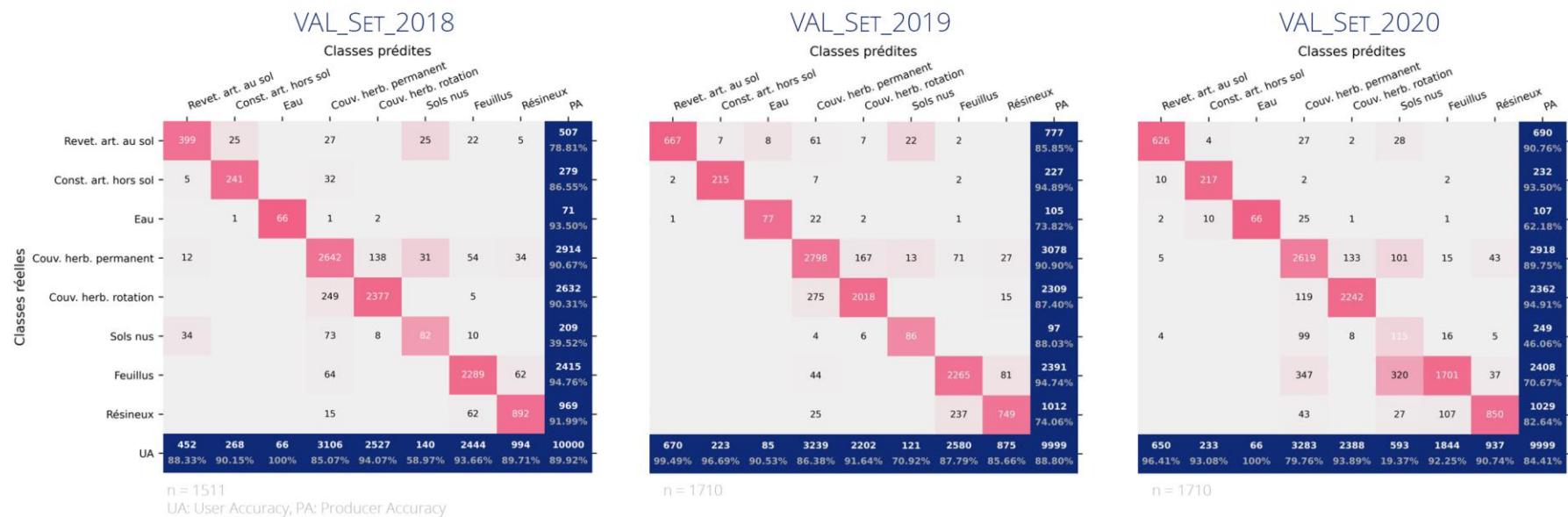


Figure 20 - Itération 04 – Résultats.

Matrices de confusion pour chacun des VAL_Sets (contre-validés). UA : Précision utilisateur. PA : Précision producteur.

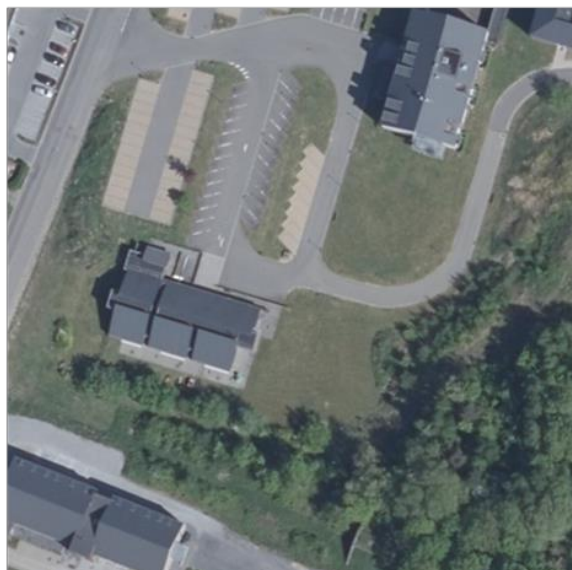
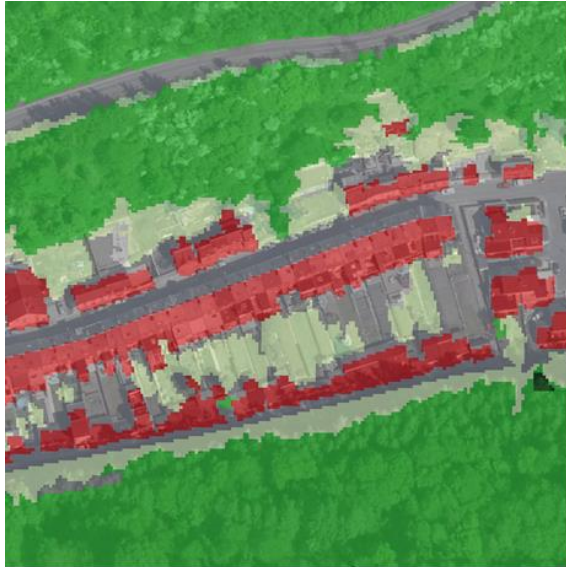


Figure 21 - Exemples de décalage sur WAL_OCS_2018 pour les "Constructions artificielles hors sol".

A gauche : WAL_OCS_2018 en transparence sur orthophoto de 2018. A droite : orthophoto de 2018 associée. Pour les trois exemples, un décalage vers le bas est constaté pour les bâtiments (couleur rouge dans la symbologie WALOUS).

3.4.4. Suggestions

Etant donné la disparité des mailles affinées en termes de qualité, uniquement les mailles jugées de bonne qualité ont servi à entraîner le modèle de cette itération. Ainsi sur les 40 mailles initialement sélectionnées pour être affinées, 38 ont été affinées et la moitié de celles-ci ont été retenues pour la première vague. L'autre moitié se rapporte à des mailles dont la qualité est moindre et qui doivent être revisitées pour être considérées pour une deuxième vague d'entraînement. Ceci fait l'objet de la Proposition 4A : consolider les 19 mailles restantes pour poursuivre l'entraînement du modèle.

La Proposition 4B vise à entraîner un modèle spécifiquement sur les données de 2020 afin de voir si un modèle uniquement destiné à faire des prédictions sur les données relatives à cette année est capable d'avoir des performances satisfaisantes. Cette proposition permettra de limiter les différentes représentations que le modèle doit apprendre car ce dernier sera spécialisé sur la saison relative aux données de 2020.

Lors de la réunion de présentation des résultats du 7 octobre 2021, la Proposition 4C a été faite. Celle-ci propose d'utiliser des règles d'entropie afin de filtrer les prédictions du modèle qui iraient à leur rencontre. Ces règles pourraient permettre d'éviter de voir des péniches segmentées sur les canaux fluviaux. Une approche moins drastique pourrait utiliser ces règles basées sur l'entropie non pas pour altérer les prédictions du modèle mais plutôt pour signaler les zones où des erreurs de segmentation peuvent avoir été commises.

3.4.5. Décisions

Le CA a décidé avec l'accord d'AERSOPACELAB de poursuivre l'intégration des mailles affinées n'ayant pas été jugées de bonne qualité. Cette nouvelle phase, appelée Phase de Consolidation, répond donc à la Proposition 4A.

La Proposition 4B a également été retenue. AERSOPACELAB a donc décidé d'entraîner un modèle DEEPLAB v3+ en utilisant exclusivement des données d'entrée de 2020.

Cependant, la Proposition 4C visant à utiliser des règles basées sur l'entropie n'a pas été retenue car le CA a convenu qu'il était très difficile de considérer toutes les exceptions lors de la définition d'une telle règle. Par exemple, on pourrait être tenté de définir une règle pour enlever la segmentation des péniches dans les canaux fluviaux. Pour cela, une simple règle indiquant qu'une zone classée comme « Revêtement artificiel au sol » (ou « Construction artificielle hors sol ») complètement entourée d'eau doit être reclassifiée en « Eau ». Cependant, cette règle pourrait également occasionner la suppression d'écluses, d'îles artificielles, ou encore de ponts mal connectés aux rives. Afin de prendre en compte ces exceptions, la règle pourrait être altérée pour considérer la taille de la zone, ainsi que sa forme. Cependant, l'ajout de ces différentes variables complexifie la règle et n'assure pas que toutes les exceptions soient prises en charge.

Puisque la plus-value de l'extension POINTREND n'a pas pu être démontrée visuellement, celle-ci ne sera pas considérée pour la suite des itérations.

3.5. Itération 5

3.5.1. Objectifs

Cette cinquième et dernière itération a eu pour objectif :

1. D'étudier les performances et limites d'un modèle exclusivement entraîné sur des données de 2020.
2. De consolider les mailles affinées restantes afin de les inclure dans la deuxième vague d'entraînement.
3. D'optimiser certains hyperparamètres du modèle

3.5.2. Implémentations

3.5.2.1. Implémentation I

La première implémentation consiste en un modèle DEEPLAB v3+ d'architecture identique à ceux utilisés dans les itérations précédentes. Les poids de l'Implémentation II de l'itération 03 ont été utilisés pour initialiser le modèle.

Puisque l'objectif est de connaître les limites du modèle lorsqu'il n'est entraîné que sur les données de 2020, les données d'entrée sont les Orthophotos et le MNH de 2020. Cependant, à défaut d'avoir une autre réalité-terrain, WAL_OCS_2018 fut utilisée.

L'entraînement s'est fait sur 35000 itérations, où chaque itération contenait un batch de 16 vignettes.

La configuration de ce modèle et de son entraînement est disponible en Annexes 6.5.1.

3.5.2.2. Implémentation II

La Phase de Consolidation a été assignée à 2 personnes : une chez AEROSPACELAB, ainsi qu'une personne du SPW. Finalement, ce sont 8 mailles supplémentaires qui ont pu être consolidées. La Table 2 liste les mailles ayant été rajoutée pour la deuxième vague d'entraînement.

ORTHO_2018__00690045
ORTHO_2018__00440085
ORTHO_2018__00470107
ORTHO_2018__00500111
ORTHO_2018__00570094
ORTHO_2018__00700042
ORTHO_2018__00430120
ORTHO_2018__00380101

Table 2 - Mailles affinées et consolidées / Ajout pour la deuxième vague

Cette tâche de consolidation ne s'est concentrée que sur les données de l'année 2018. En effet, l'itération précédente avait mis en lumière le fait qu'aucune amélioration des performances du modèle sur le VAL_SET_2020 n'avait été constatée, malgré l'affinement de WAL_OCS_2018 sur base de données de 2020 pour quelques mailles. De plus, il incombe à l'Implémentation I de vérifier la capacité d'un modèle DEEPLAB v3+ à segmenter de manière satisfaisante les données de 2020.

Comme lors de la Phase d'Affinement, l'effort s'est concentré sur la labélisation des contours de zones artificielles, l'uniformisation des zones, la correction d'erreurs, la segmentation des arbustes, et la distinction entre « Feuillus » et « Résineux ».

Ainsi le modèle a été ré-entraîné sur l'ensemble des mailles affinées de 2018, soit 23 mailles de 4km². Les détails sur la configuration de ce modèle sont fournis en Annexes 6.5.2.

3.5.2.3. Implémentation III

En amont de l'entraînement d'un modèle, certains paramètres de ce dernier doivent être sélectionnés. Ceux-ci, appelés hyperparamètres, resteront inchangés pour la durée de l'entraînement. Ils déterminent, entre autres, le comportement du modèle et son architecture.

Puisqu'il est impossible d'optimiser totalement l'ensemble des hyperparamètres du modèle, à cause des domaines continus de certains d'entre eux, AEROSPACELAB s'est uniquement restreint à certains hyperparamètres qui, selon son expérience, peuvent avoir un impact non négligeable. La Table suivante présente les 8 hyperparamètres sélectionnées :

Hyperparamètre	Description	Domaine de recherche	Nom dans le fichier de configuration
Profondeur du modèle	Au plus le modèle est profond, au plus il sera capable d'apprendre des représentations complexes. Cependant, le risque d'overfitting et les ressources nécessaires pour l'entraînement et l'inférence du modèle augmentent également	101, 50	Resnet Depth
Normalisation du modèle	La normalisation interne du modèle (aux activations) permet une convergence plus rapide du modèle et une insensibilité plus importante à l'initialisation des poids. La normalisation peut se faire sur différentes dimensions avec certains avantages et inconvénients	Synchronized Batch Normalization, Group Normalization	Resnet Norm
Taux d'apprentissage	Facteur utilisé pour déterminer l'importance de la mise à jour des poids du modèle lors de son entraînement. Un taux d'apprentissage faible entraîne une convergence plus lente alors qu'un taux d'apprentissage élevé peut rendre instable l'apprentissage	[1E-3 : 1E-4]	Learning Rate
Gamma	Facteur définissant la diminution du taux d'apprentissage au cours de l'entraînement. Cette diminution permet de transitionner d'une phase exploratoire vers une phase d'exploitation	0.05, 0.1, 0.2	Gamma
Itérations d'échauffement	Permettent de graduellement augmenter le taux d'apprentissage de 0 à sa valeur de base	10, 500, 1000	Warmup Iters
Momentum	Facteur utilisé lors du calcul du gradient déterminant l'influence de l'itération précédente (inertie)	[0.01 : 0.9]	Momentum
Paramètre de la régularisation L2	Permet la régularisation du modèle en le pénalisant en fonction de la norme L2 de ses poids. Un facteur élevé contraint davantage le modèle mais peut engendrer un modèle qui n'est pas assez flexible pour modéliser la distribution des données	[1E-5 : 1E-3]	Weight Decay
Nesterov	Approche permettant d'influencer le calcul du momentum dans le gradient	0, 1	Nesterov

Table 3 - Hyperparamètres sélectionnés pour l'optimisation

Une approche de recherche aléatoire a été sélectionnée pour observer l'impact des certaines configuration d'hyperparamètres sur les performances du modèle. Cette approche, facile à mettre en place, peut être parallélisée et présente un avantage majeur par rapport à l'approche traditionnelle de recherche en grille puisqu'elle est moins impactée par l'ajout d'hyperparamètres qui n'ont peu ou pas d'impact significatif sur les performances du modèle (Bergstra and Bengio 2012).

Ainsi, pour chacune des 18 expériences, la valeur de chaque hyperparamètre fut prise aléatoirement dans son domaine prédéfini et le modèle obtenu fut entraîné pendant 10000 itérations avec une taille de batch de 8 vignettes.

3.5.3. Observations et interprétations

3.5.3.1. Implémentation I

Les résultats de cette implémentation tendent à montrer que dans cette configuration d'entraînement, le modèle éprouve toujours des difficultés à correctement distinguer les « Couverts herbacés permanent » des « Feuillus » sans feuilles puisque les performances sur VAL_Set_2020 sont similaires à celles observées à l'itération précédente. Les rappels et précisions de cette implémentation sont donnés, pour chacune des strates, dans la matrice de confusion à la Figure 22.

	Classes prédites								
	Revet. art. au sol	Const. art. hors sol	Eau	Couv. herb. permanent	Couv. herb. rotation	Sols nus	Feuillus	Résineux	
Revet. art. au sol	701	2		22		3	2		733 95.71%
Const. art. hors sol	6	234							240 97.47%
Eau	3		91			7	1	1	105 87.20%
Couv. herb. permanent	13			2736	21	10	5	5	2794 97.94%
Couv. herb. rotation	2			186	2157		10		2357 91.52%
Sols nus	53			192	6	99	6		358 27.86%
Feuillus				893	5		1368	79	2347 58.30%
Résineux				168			80	814	1063 76.54%
UA	781	237	91	4200	2191	121	1475	900	9999
	89.81%	98.77%	100%	65.16%	98.47%	81.97%	92.74%	90.41%	82.05%

Figure 22 - Itération 05/Implémentation I – Résultats.

Matrice de confusion pour le VAL_Set_2020 (contre-validé). UA : Précision utilisateur. PA : Précision producteur.

Cette confusion avec les « Couverts herbacés permanent » est également observée pour d'autres strates, telles que « Sols nus », « Couverts herbacés en rotation », « Résineux », mais dans une moindre mesure.

Dans cette configuration, et malgré un entraînement spécifique sur les données de 2020, le modèle ne parvient toujours pas à proposer une segmentation qui réponde au cahier des charges pour l'année 2020. Les confusions observées sont causées par des signatures spectrales similaires pour certains points de ces strates. En effet, les orthophotos ayant été prises plus tôt dans l'année que celles de 2018 et 2019, le manque de végétation réduit l'information contenue dans ces orthophotos entraînant une baisse de performance du modèle. Les Figures suivantes permettent de visualiser cette perte d'information sur les orthophotos de 2020 par rapport à 2018 ou 2019.



Figure 23 - Comparaison des orthophotos de 2018, 2019 et 2020

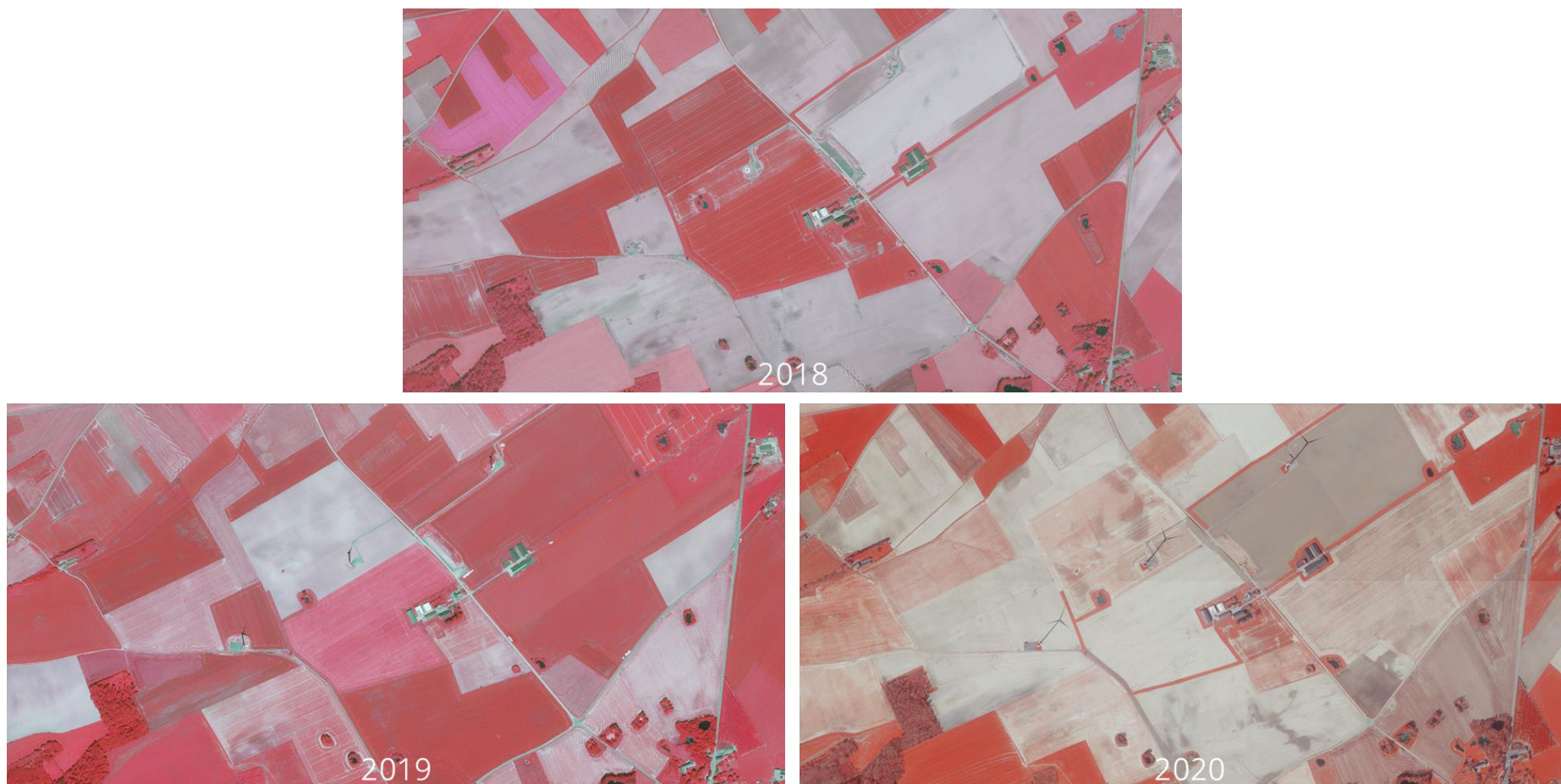


Figure 24 - Comparaison des orthophotos infra-rouge de 2018, 2019 et 2020

Plusieurs perspectives pour améliorer le modèle sur les données de 2020 ont été proposées par AEROSPACELAB dans le document « Rapport Solution ». Par manque de ressources, celles-ci n'ont pu être investiguées lors de ce marché.

3.5.3.2. Implémentation II

En concentrant l'entraînement de ce modèle uniquement sur les mailles affinées, celui-ci a pu apprendre les nuances entre certaines strates qui ont été apportées lors de la labélisation manuelle de ces mailles. En effet, le modèle s'est nettement amélioré sur son rappel des « Sols nus » et des arbres des moins de 3m, comme cela est illustré ci-dessous :

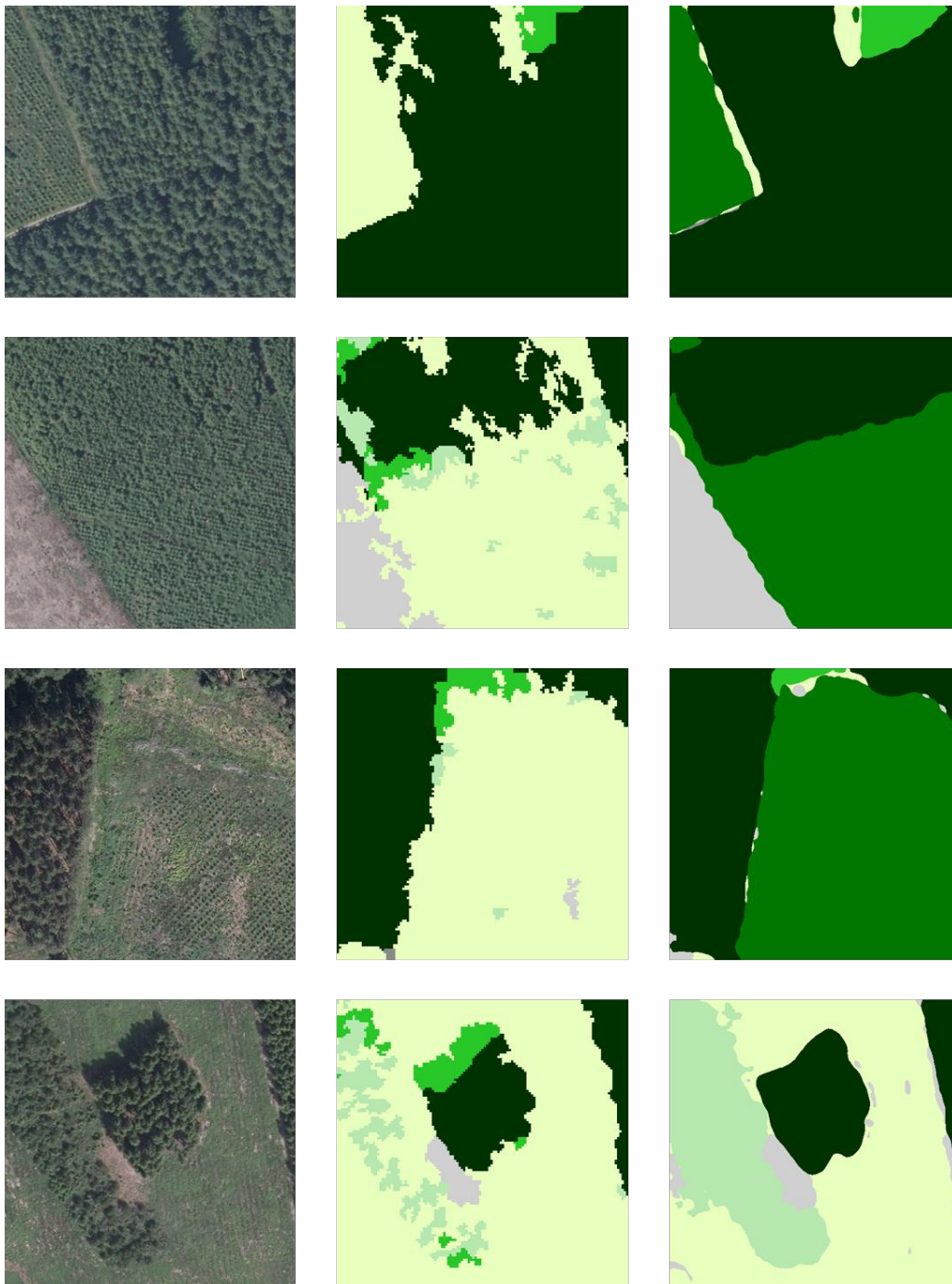


Figure 25 - Itération 05/Implémentation II – Amélioration du rappel des arbustes.
A gauche : Orthophoto 2018. Au centre : WAL_OCS_2018 (avec une résolution spatiale de 1m). A droite : Prédiction de l'Implémentation II de l'itération 5 (avec une résolution spatiale de 0.25m)

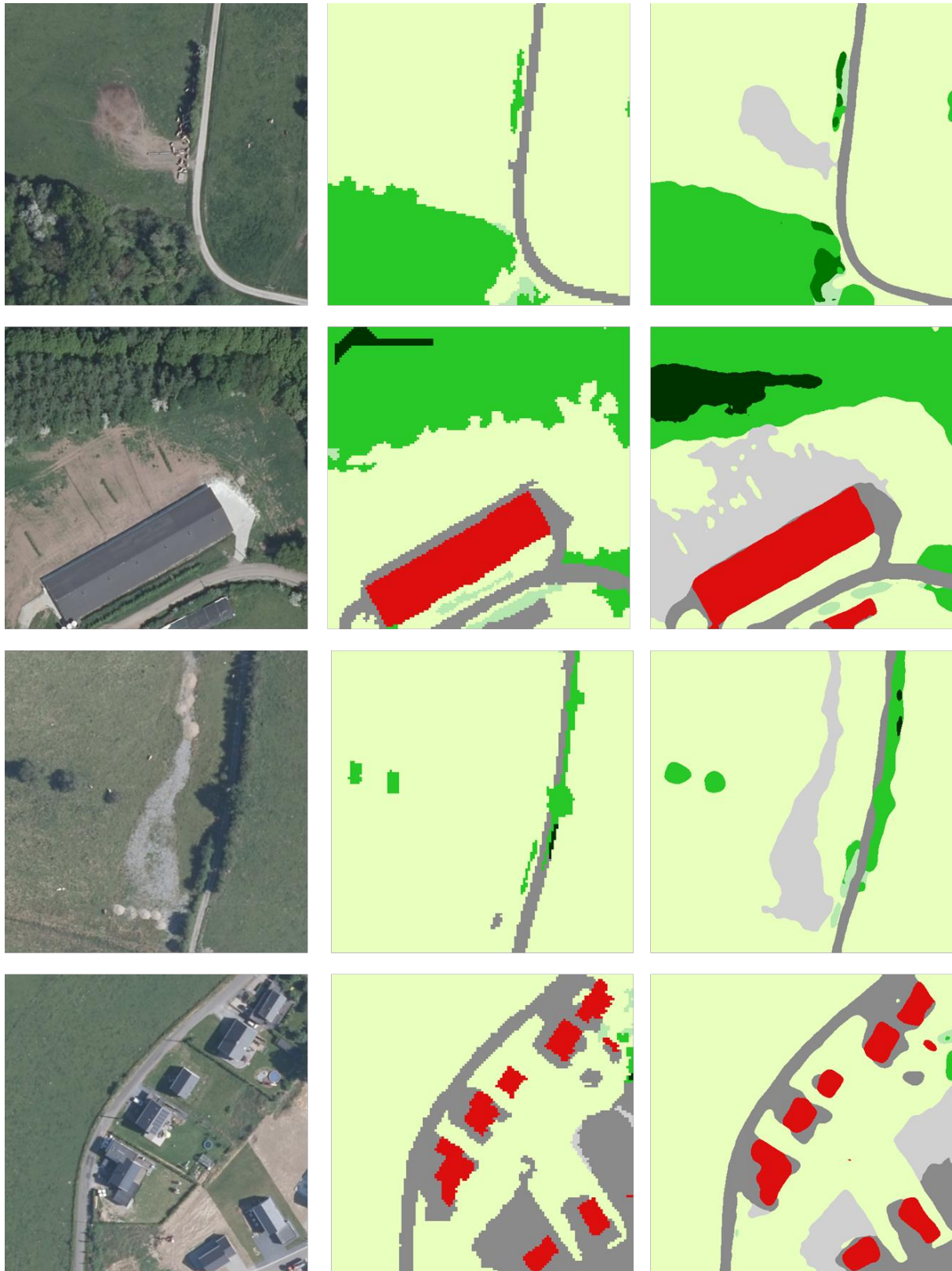


Figure 26 - Itération 05/Implémentation II – Amélioration du rappel des "Sols nus".
A gauche : Orthophoto 2018. Au centre : WAL_OCS_2018 (avec une résolution spatiale de 1m). A droite :
Prédictions de l'Implémentation II de l'Itération 5 (avec une résolution spatiale de 0.25m)

L'affinement de certaines mailles de la réalité-terrain a également permis une meilleure délimitation des zones artificielles. En effet, il peut être, par exemple, constaté que les intersections routières sont parfois mieux segmentées que sur WAL_OCS_2018. En effet, les données vectorielles du PICC ont été

utilisées par SOL_2018 pour dériver les réseaux routiers. Plusieurs adaptations à ces données vectorielles ont dû être apportées afin de les convertir en raster. Par exemple, des espaces « tampons » de chaque côté des axes routiers ont dû être appliqués afin de représenter la largeur des routes. D'autres adaptations ont été faites et celles-ci sont détaillées en partie dans les documents WALOUS_D2_squelettevectoriel (Service Public de Wallonie (SPW) - Walous 2018) et WALOUS_D6_consolidation (Service Public de Wallonie (SPW) - Walous 2020).

Dans certains cas, il peut être observé que les « tampons » utilisés ne correspondent pas à la largeur réelle de la route et que les prédictions du modèle issue de l'itération 05 sont plus adaptées. Cela est illustré à la Figure 27.

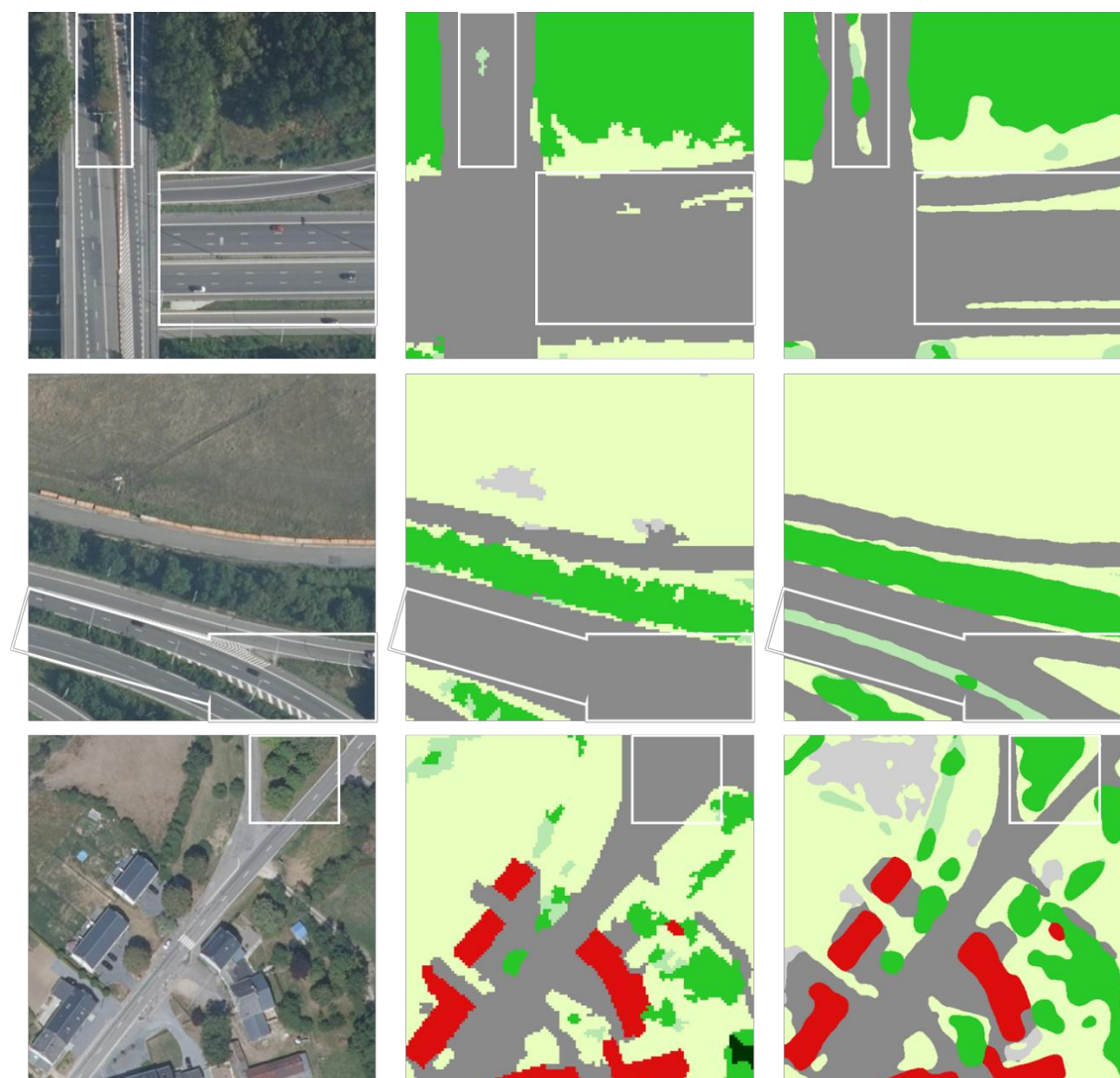


Figure 27 - Améliorations sur la segmentation des grands axes routiers.

A gauche : Orthophoto 2018. Au centre : WAL_OCS_2018 (avec une résolution spatiale de 1m). A droite : Prédictions de SOL_2019/20 (avec une résolution spatiale de 0.25m). Les rectangles aux bords blancs mettent en évidence les différences entre WAL_OCS_2018 et les prédictions du nouveau modèle.

Cependant, les rappels de certaines autres strates, telles que l'« Eau » et les « Couverts herbacés en rotation », ont été impactés négativement causant des artéfacts visuels rendant cette implémentation difficilement viable pour la production des cartes d'occupation du sol. De manière systématique, l'« Eau » est confondue par endroit en « Sols nus » et certaines parcelles agricoles en « Résineux < 3m ». Même si la confusion entre « Couverts herbacés en rotation » et « Résineux < 3m » peut être en partie expliquée par le fait que certaines cultures peuvent avoir l'apparence de jeunes sapinières, la première confusion est plus difficilement interprétable. Cependant, ces confusions sont également et en partie le

résultat d'éventuelles erreurs dans les étapes d'affinement et/ou de consolidation, voire d'un manque de diversité dans le jeu de données d'entraînement. La Figure 28 permet d'observer ces confusions.

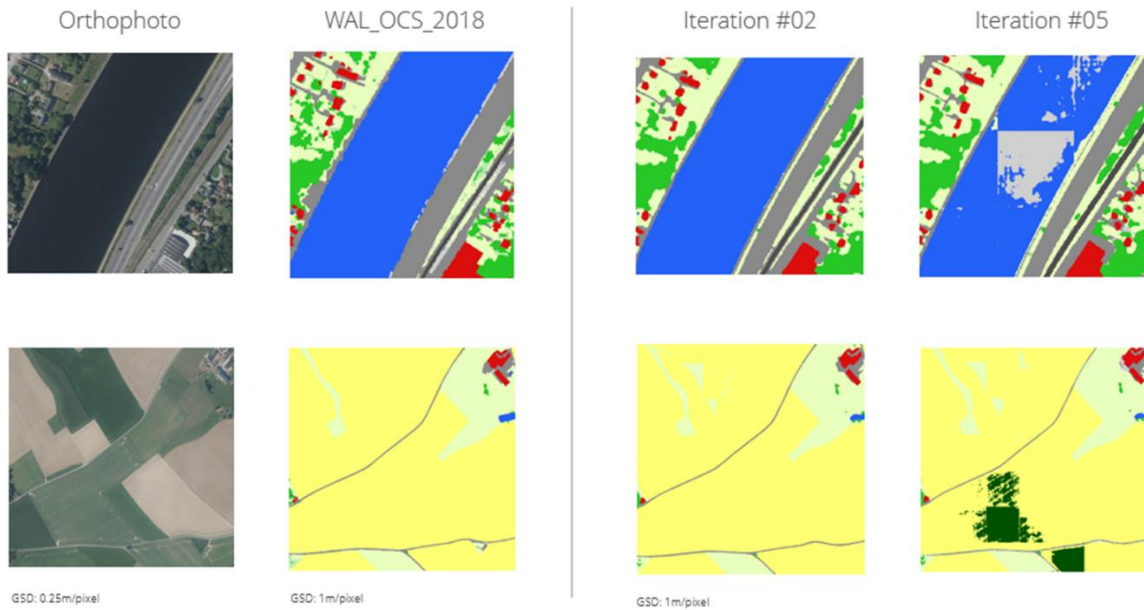


Figure 28 - Itération 05/Implémentation II - Confusions entre certaines strates notamment "Eau" / "Sols nus" et "Couverts herbacés en rotation" / "Résineux < 3m"

La matrice de confusion dérivée du VAL_Set_2019 ne laissait néanmoins pas présager ces confusions importantes. Les matrices de confusion pour chacun des VAL_Sets sont reprises à la Figure 29.

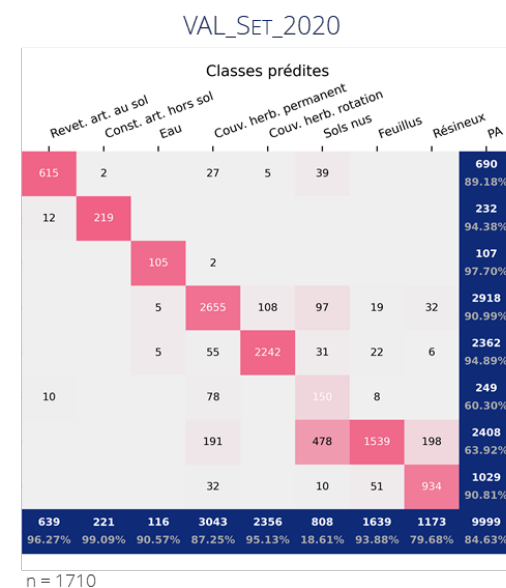
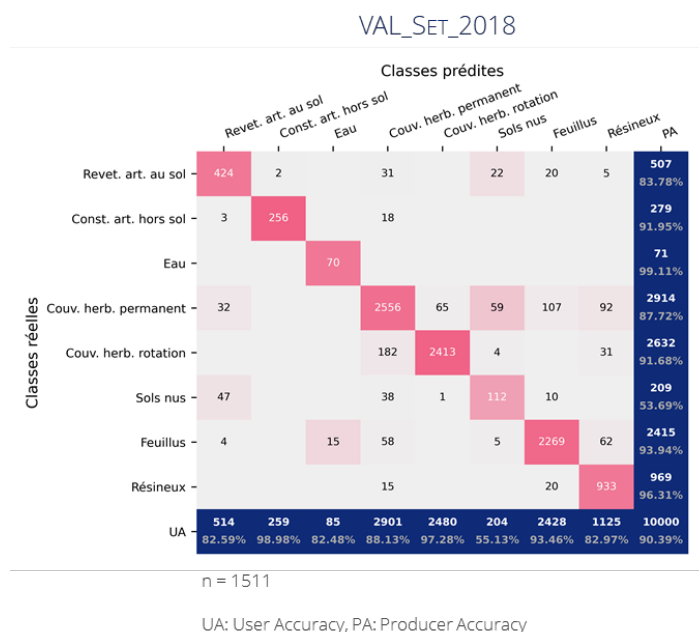


Figure 29 - Itération 05/Implémentation II – Résultats.

Matrices de confusion pour chacun des VAL_Sets (contre-validés). UA : Précision utilisateur. PA : Précision producteur.

3.5.3.3. Implémentation III

Il ressort de l'analyse de ces 8 hyperparamètres, que seul deux ont un impact non négligeable sur les performances du modèle, plus spécifiquement sa précision globale sur VAL_Set_2018. En effet, les résultats de cette analyse (voir Figure 30) tendent à montrer que le paramètre de normalisation du modèle (Resnet Norm), ainsi que, et dans une moindre mesure, celui lié à sa profondeur (Resnet Depth), peuvent influencer les performances du modèle :

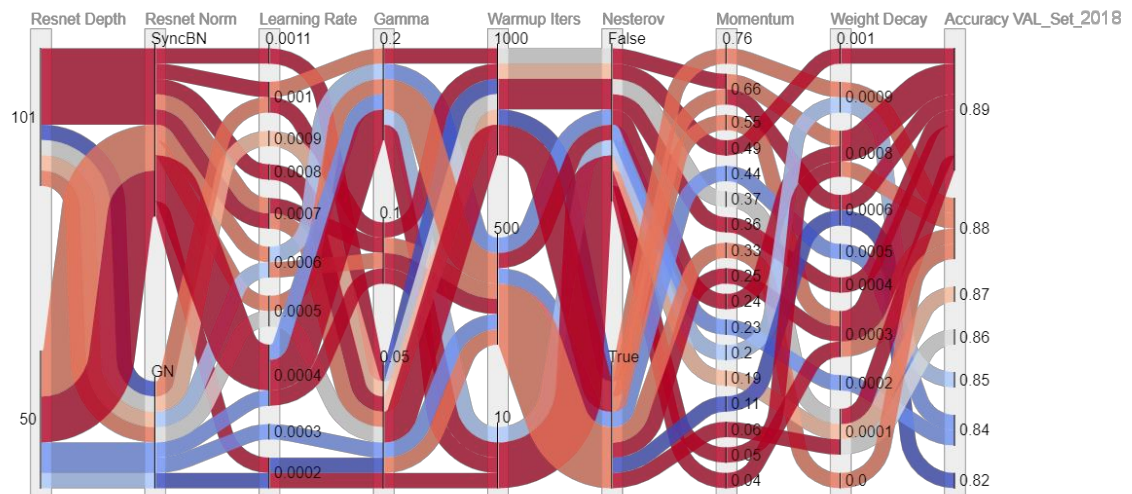


Figure 30 – Résultats de l'analyse d'un set d'hyperparamètres sur le VAL_Set_2018

Les résultats tendent à suggérer que l'utilisation d'une « Batch Norm » synchronisée (SyncBN) permettrait d'obtenir un meilleur modèle pour cette tâche que la normalisation par groupe (GN).

La normalisation par batch (BN) s'applique à plusieurs endroits au sein du modèle : sur les données d'entrée mais également sur les résultats d'activation d'une couche. Cette normalisation standardise donc les données en ne tenant compte uniquement de celles présentes dans le batch. De manière générale, la « Batch Norm » permet de stabiliser l'apprentissage du modèle réduisant donc le nombre d'epochs requis pour le faire converger (Santurkar, et al. 2019) (Bengio, Goodfellow and Courville 2016). La version synchronisée de cette normalisation permet de synchroniser la normalisation sur l'ensemble du batch lorsque plusieurs GPUs sont utilisées lors de l'entraînement.

La GN normalise les données en regroupant d'abord les bandes (i.e. Rouge, Vert, Bleu, Infra-rouge, MNH) en sous-groupe. Cette normalisation est ainsi indépendante de la taille du batch, mais nécessite l'ajout d'un nouveau hyperparamètre définissant le nombre de groupes. Par défaut, celui-ci a été laissé à 32 comme suggéré dans par les auteurs (Wu and He 2018).

La profondeur du modèle semble avoir également un impact sur la stabilité de l'apprentissage. Dans sa configuration à 101 couches, la variance des performances du modèle est plus petite que celle obtenue en entraînant un Resnet50. Cependant, sur base des résultats obtenus, l'impact de cet hyperparamètre reste relativement moins important que celui de la normalisation des données.

Il est important de noter qu'une analyse plus approfondie, incluant également davantage d'hyperparamètres, devrait être réalisée afin d'en dériver des conclusions plus importantes et solides.

Le lecteur est redirigé vers la présentation de fin d'itération 05 pour plus de détails.

3.5.4. Suggestions

Etant donné que cette itération fut la dernière de la Phase II, celle-ci n'a pas débouchée sur des propositions qui furent considérées pour une implémentation lors de ce marché. Cependant, un rapport détaillant les perspectives d'amélioration de la solution actuelle sur base des faiblesses observées a été rédigé par AEROSPACELAB : « Rapport - Solution ».

3.5.5. Décisions

Concernant les hyperparamètres, la décision a été prise de conserver ceux utilisés jusqu'à présent dans les diverses itérations puisque seul l'hyperparamètre de normalisation semble pouvoir jouer un rôle significatif sur la performance du modèle, or la meilleure option, la normalisation par batch, fut déjà sélectionnée lors des précédentes itérations.

Puisque cette itération est la dernière, le choix de l'itération a été laissé au CA et les votes ont indiqués la préférence des membres du CA pour les Itérations 05, 02, et 04, ordonnées suivant le nombre de votes reçus :

1. Quelle est l'itération qui vous semble la plus pertinente?

[Plus de détails](#)

- itération 1 (basée uniquement... 0
- itération 2 (Ortho (2018 et 202... 0
- itération 3 (Ortho (2018 et 202... 0
- itération 4 (MNH+ortho+Affin... 1
- itération 5 (MNH+ortho+Affin... 5



2. Quel serait votre second choix ?

[Plus de détails](#)

- itération 1 (basée uniquement... 0
- itération 2 (Ortho (2018 et 202... 3
- itération 3 (Ortho (2018 et 202... 2
- itération 4 (MNH+ortho+Affin... 1
- itération 5 (MNH+ortho+Affin... 0



Figure 31 - Sélection de l'itération

Cependant, suite aux augmentations des taux d'omissions pour les classes "Eau" et "Couverts herbacés en rotation", l'itération 05 n'a pas été retenue et l'itération 02 lui a été préférée.

4. Récapitulatif des décisions influençant la solution retenue

Date	Sujet	Décision	Raison
19/01/2021	Approche A centrée sur la détection des changements sur base des orthophotos puis de la segmentation de ceux-ci.	Rejeté	<ul style="list-style-type: none"> • Une approche supervisée n'est pas envisageable puisqu'il manque la réalité-terrain (i.e. carte des changements survenus entre 2018 et 2019/2020). Les approches non-supervisées telles que CCVA ne sont pas robustes pour faire abstraction des nombreuses perturbations accentuées ce niveau de résolution spatiale • Cette approche ne permet pas de directement corriger les erreurs éventuellement présentes de WAL_OCS_2018 puisqu'uniquement les changements seront reconsidérés
19/01/2021	Approche B composée d'un premier modèle en charge de la génération d'un candidat-carte OCS basé sur les orthophotos et affiné par un second modèle se concentrant sur la segmentation des désaccords entre le candidat-carte OCS et WAL_OCS_2018 via l'utilisation de diverses sources telles que Sentinel-2, MNH, ...	Approuvé	<ul style="list-style-type: none"> • Modèle robuste • Approche générique • Automatisation
02/04/2021	Abandon de l'Approche B	Approuvé	Le modèle de segmentation ne permet pas, sur base des

			orthophotos uniquement, de générer des candidats-cartes OCS suffisamment fiables pour justifier une approche en deux étapes ayant pour but une segmentation rapide des zones triviales
02/04/2021	Proposition 1A : Utilisation des orthophotos de 2019 et 2020 pour entraîner le modèle DEEPLAB v3 afin de constituer un jeu de données d'entraînement plus représentatif et plus diversifié.	Approuvé	Ceci permettrait au modèle d'associer plusieurs représentations différentes liées aux diverses dates de prise de vue pour les mêmes objets (e.g. : feuillus sans feuilles, couvert herbacé permanent plus sec, ...)
02/04/2021	Proposition 1B : Concentration de l'entraînement du modèle DEEPLAB v3 sur les classes ayant subi de fortes perturbations entre 2018 et 2019/2020 afin de diversifier le jeu de données d'entraînement et lui permettre d'apprendre les multiples représentations d'une classe (e.g. : Feuillus avec et sans feuilles).	Approuvé	Dû au déséquilibre des classes dans le jeu de données d'entraînement, favoriser l'échantillonnage des classes sous représentées permettrait au modèle d'affiner son apprentissage sur ces dernières
02/04/2021	Proposition 1C : Augmentation de la dimension spatiale des vignettes orthophotos utilisées pour l'entraînement du modèle DEEPLAB v3 afin d'apporter plus de contexte spatial.	Rejeté	A ce stade, ce modèle de segmentation ne semble pas souffrir de problèmes liés à un manque de contexte avec ces vignettes de 128m de côté
02/04/2021	Proposition 1D : Utilisation des modèles numériques de surface et de terrain comme données d'entrées supplémentaires pour le modèle DEEPLAB v3, pour faciliter la distinction entre les revêtements artificiels au sol et hors sol, ainsi que la distinction entre les arbres de petite et grande taille.	Approuvé	Ces données seront indirectement utilisées afin d'en dériver un modèle numérique de hauteur qui devrait permettre au modèle de plus facilement faire la distinction entre les arbres de petites et grandes tailles ainsi que les revêtements artificiels au et hors sol.

02/04/2021	Proposition 1E : Utilisation des informations des relevés LIDAR pour faciliter la détection des surfaces d'eau puisque ceux-ci absorbent les échos.	Approuvé	Ceci se fait au travers de l'utilisation du modèle numérique de terrain de 2013/2014
02/04/2021	Proposition 1F : Utilisation de squelettes vectoriels thématiques tels que le PICC et le LPIS comme données d'entrée au modèle DEEPLAB V3 pour contextualiser davantage l'orthophoto.	Rejeté	Ceci rajouterait une nouvelle dépendance qui ne semble pas être nécessaire à ce stade. En effet, les apports du modèle numérique de hauteur et des données de sortie du modèle ConvLSTM peuvent suffire pour apporter le contexte manquant au modèle. De plus, rajouter une nouvelle dépendance externe de ce type impliquerait que pour pouvoir mettre à jour la carte d'occupation du sol, cette donnée externe devra d'abord être mis à jour elle-même.
02/04/2021	Proposition 1G : Ajout des données de sorties du modèle CONV LSTM comme données d'entrée du modèle DEEPLAB V3 pour aider le modèle à classifier les strates naturelles.	Approuvé	Afin de tester l'apport de ces données dérivées d'images SENTINEL-2, ces premières doivent être proposées en données d'entrée du modèle de segmentation principal.
02/04/2021	Proposition 1H : Nouvelle approche, Approche C, basée sur un seul modèle de segmentation.	Approuvé	
06/05/2021	Proposition 2A : Poursuite de l'entraînement du « Modèle de segmentation » et échantillonnage adapté des instances d'entraînement pour améliorer le rappel de la strate « Feuillus » sur les orthophotos de 2020.	Approuvé	Le modèle ayant été entraîné que sur une demie epoch, la poursuite de son entraînement pourra lui permettre d'affiner son apprentissage
06/05/2021	Proposition 2B : Poursuite de l'entraînement du « Modèle de	Approuvé	

	segmentation » et échantillonnage adapté des instances d'entraînement pour améliorer le rappel de la strate « Eau » pour les grandes étendues d'eau.		
06/05/2021	Proposition 2C : Poursuite de l'entraînement du « Modèle de segmentation » et échantillonnage adapté des instances d'entraînement pour améliorer la précision du modèle aux contours de zones.	Approuvé	
06/05/2021	Proposition 2D : Ajout de l'extension POINTREND à l'architecture actuelle du « Modèle de segmentation » pour améliorer la précision du modèle aux contours des zones.	Rejeté	Cette extension pourrait venir complexifier inutilement la solution s'il s'avère que la Proposition 2C porte ses fruits
06/05/2021	Proposition 2E : Amélioration du rappel pour les strates « Sols nus », « Revêtements artificiels au sol », et « Couverts herbacés permanents », via un affinement des données réalité-terrain (i.e. WAL_OCS_2018), grâce à un effort manuel.	Rejeté	La charge de travail pourra être mieux quantifiée lorsque les faiblesses du modèle seront clairement identifiées suite à l'entraînement prolongé du modèle
06/05/2021	Proposition 2F : Amélioration du rappel pour les strates « Sols nus », « Revêtements artificiels au sol », et « Couverts herbacés permanents », via un affinement des données réalité-terrain (i.e. WAL_OCS_2018), en utilisant les données du PICC.	Rejeté	<ul style="list-style-type: none"> • Les données du PICC sont déjà indirectement prises en compte puisque celles-ci ont été utilisées pour générer WAL_OCS_2018 • L'angle de vue des orthophotos entraînera des disparités entre l'empreinte cadastrale/au sol d'un bâtiment et la surface qu'il occupe sur l'orthophoto • Cette proposition présuppose que les données du PICC soient à jour

06/05/2021	Proposition 2G : Amélioration du rappel pour les strates « Sols nus », « Revêtements artificiels au sol », et « Couverts herbacés permanents », via un affinement des données réalité-terrain (i.e. WAL_OCS_2018), en utilisant les données vectorielles des carrières actives.	Approuvé partiellement	Cette couche vectorielle de données est à jour mais une étude préliminaire de la qualité des données a été demandée à AEROSPACELAB
06/05/2021	Proposition 2H : Amélioration du rappel pour les strates « Sols nus », « Revêtements artificiels au sol », et « Couverts herbacés permanents », via une intégration de l'historique comme données d'entrée.	Rejeté	Le risque d'overfitting est trop important puisque ces données historiques de la carte d'occupation du sol seraient trop proches du résultat attendu en sortie
06/05/2021	Proposition 2I : Amélioration du rappel pour les strates « Sols nus », « Revêtements artificiels au sol », et « Couverts herbacés permanents », via une utilisation des données HSL (Harmonized Landsat Sentinel-2) afin de déterminer l'humidité/perméabilité du sol.	Rejeté	La résolution spatiale de ces données ne pourra pas, selon AEROSPACELAB, apporter suffisamment d'information pour être bénéfique au modèle
06/05/2021	Proposition 2J : Amélioration du rappel pour les strates « Sols nus », « Revêtements artificiels au sol », et « Couverts herbacés permanents », via une utilisation des données SAR de SENTINEL-1 afin de déterminer la consistance du sol.	Rejeté	
06/05/2021	Proposition 2K : Amélioration du rappel pour les strates « Sols nus », « Revêtements artificiels au sol », et « Couverts herbacés permanents », via une détermination des signatures spectrales pour ces strates	Rejeté	Ces règles définies manuellement sont généralement difficile à rendre robustes sur l'ensemble du territoire wallon et peuvent ne pas s'appliquer pour les différentes périodes d'acquisition des données d'entrée.
06/05/2021	Proposition 2L : Implémentation de règles basées sur l'entropie pour	Rejeté	

	affiner les prédictions faites par le modèle.		
06/05/2021	Proposition 2M : Utilisation du MNT 2013-2014 LIDAR sans interpolation pour améliorer le rappel de la strate « Eau ».	Rejeté	Les données LIDAR ne sont disponibles que pour les années 2013-2014 et aucune perspective de mise à jour n'est pour le moment envisagée au SPW
06/05/2021	Proposition 2N : Exploitation de la densité des réponses LIDAR du MNT 2013-2014 pour faciliter la distinction entre une surface artificielle, généralement plus lisse et donc caractérisée par un écho net et précis, et un sol nu, souvent plus irrégulier et donc associé à un écho plus bruité.	Rejeté	
06/05/2021	Proposition 2O : Utilisation du LPIS pour améliorer la précision du modèle dans ses classifications de « Couverts herbacés permanents » et « Couverts herbacés en rotation ».	Rejeté	<ul style="list-style-type: none"> • Ajout d'une dépendance • Comment traiter les conflits : e.g. : « Construction artificielle hors sol » dans un champ ?
06/05/2021	Proposition 2P : Exclusion de l'entraînement des zones où la confiance du modèle est faible.	Rejeté	Ne pas entrainer un modèle sur les cas où il se trompe ne permettra pas de régler le problème
06/05/2021	Proposition 2Q : Distinction entre les coupes à blanc et les sols nus.	Accepté partiellement	Pour autant que des données labélisées puissent être obtenues, le modèle pourrait être capable de faire cette distinction
06/05/2021	Proposition 2R : Utilisation de la couche vectorielle des « 3 classes » pour affiner les données réalité-terrain (i.e. WAL_OCS_2018) notamment pour les « Résineux », « Feuillus » et coupes à blanc.	Accepté partiellement	Avant l'utilisation de ces données, une étude préliminaire de leur qualité a été demandée à AEROSPACELAB
09/06/2021	Proposition 2D : (reproposée) Ajout de l'extension POINTREND à l'architecture actuelle du « Modèle	Accepté	Les délimitations des zones artificielles ne sont toujours pas précises lorsqu'on observe les prédictions à une résolution spatiale

	de segmentation » pour améliorer la précision du modèle aux contour des zones.		de 0.25m malgré le prolongement de l'entraînement
09/06/2021	Proposition 2E : (reproposée) Amélioration du rappel pour les strates « Sols nus », « Revêtements artificiels au sol », et « Couverts herbacés permanents », via un affinement des données réalité-terrain (i.e. WAL_OCS_2018), grâce à un effort manuel.	Accepté	Selon AEROSPACELAB une amélioration significative des performances ne peut, à ce stade, qu'être obtenue par l'affinement des données réalité-terrain
09/06/2021	Proposition 2O : (reproposée) Utilisation du LPIS pour améliorer la précision du modèle dans ses classifications de « Couverts herbacés permanents » et « Couverts herbacés en rotation ».	Rejeté	Idem que pour la proposition initiale
09/06/2021	Proposition 3A : Génération d'une carte regroupant les « zones complexes » sur lesquelles le modèle n'aurait pas à faire de segmentation.	Rejeté	La mise en pratique de cette proposition est rendue ardue par le fait qu'un expert en photo-interprétation devrait passer en revue toutes les zones de désaccords sur l'ensemble du territoire wallon pour distinguer les « zones complexes » des simples changements
07/10/2021	Proposition 4A : Consolidation des 19 mailles restantes pour poursuivre l'entraînement du modèle.	Accepté	Afin de ne pas exploiter le travail d'affinement partiellement effectué sur les 19 mailles non reprises dans la première vague. De plus, ces nouvelles mailles peuvent apporter de la diversité dans les données d'entraînement
07/10/2021	Proposition 4B : Entraînement d'un modèle spécifiquement sur les données de 2020 afin de voir si un modèle uniquement destiné à faire des prédictions sur les données	Accepté	Afin de comprendre les limites du modèle en simplifiant son apprentissage par la réduction du nombre de représentations de chaque classe à apprendre

	relatives à cette année est capable d'avoir des performances satisfaisantes.		
07/10/2021	Proposition 4C : Utilisation des règles d'entropie afin de filtrer les prédictions du modèle qui iraient à leur rencontre. Ceci pourrait permettre de corriger les prédictions aberrantes du modèle qui peuvent être capturées par une simple règle d'entropie.	Rejeté	Ces règles définies manuellement sont en générale difficile à rendre robustes sur l'ensemble du territoire wallon. De plus, la liste de toutes les exceptions à la règle doit être exhaustive pour éviter de rajouter malencontreusement des erreurs
07/10/2021	Abandon de l'extension POINTREND.	Approuvé	L'extension n'a pas eu de plus-value notoire. Pour éviter de complexifier inutilement la solution, l'extension a donc été enlevée
19/01/2022	Sélection de l'itération à retenir pour la génération de la carte d'occupation du sol.	Itération 02	Vote du CA sur base des performances du modèle sur le VAL_Set

5. Références

- Bengio, Yoshua, Ian Goodfellow, et Aaron Courville. 2016. *Deep Learning*. 1st. Cambridge: The MIT Press. <http://www.deeplearningbook.org>.
- Bergstra, James, et Yoshua Bengio. 2012. «Random Search for Hyper-Parameter Optimization.» Édité par Leon Bottou. *Journal of Machine Learning Research* (13): 281-305. <http://jmlr.org/papers/v13/bergstra12a.html>.
- Card, Don H. 1982. «Using known map category marginal frequencies to improve estimates of thematic map accuracy.» *Photogrammetric Engineering and Remote Sensing* (NASA Ames Research Center) 48: 431-439. https://www.asprs.org/wp-content/uploads/pers/1982journal/mar/1982_mar_431-439.pdf.
- Chen, Liang-Chieh, George Papandreou, Florian Schroff, et Hartwig Adam. 2017. «Rethinking Atrous Convolution for Semantic Image Segmentation.» *Cornell Univeristy / Computer Vision and Pattern Recognition*, 5 December. doi:1706.05587.
- ESA. s.d. *Level-2A Algorithm Overview*. Accès le November 2021. <https://sentinels.copernicus.eu/web/sentinel/technical-guides/sentinel-2-msi/level-2a/algorithm>.
- Fernandez-Moral, Eduardo, Renato Wolf, Denis Martins, et Patrick Rives. 2018. «A new metric for evaluating semantic segmentation: leveraging global and contour accuracy.» *2018 IEEE Intelligent Vehicles Symposium (IV)*. IEEE. 1051-1056. doi:10.1109/IVS.2018.8500497.
- Kirillov, Alexander, Yuxin Wu, Kaiming He, et Ross Girshick. 2020. «PointRend: Image Segmentation as Rendering.» *CoRR*, 16 February. doi:1912.08193.
- Rußwurm, M., et M. Korner. 2018. «Multi-Temporal Land Cover Classification with Sequential Recurrent Encoders.» *International Journal of Geo-Information*, 21 March. doi:1802.02080.
- Santurkar, Shibani, Dimitris Tsipras, Andrew Ilyas, et Aleksander Madry. 2019. «How Does Batch Normalization Help Optimization?» Édité par S Bengio, H Wallach, H Larochelle, K Grauman, N Cesa-Bianchi et R Garnett. *Advances in Neural Information Processing Systems*. Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2018/file/905056c1ac1dad141560467e0a99e1cf-Paper.pdf>.
- Service Public de Wallonie (SPW) - Walous. 2019. «Guide méthodologique finalisé et recommandations d'usages pour la fusion des données de la carte d'occupation du sol de Wallonie 2018.» 29. https://geoportail.wallonie.be/files/PDF/annexes_walous/WALOUS_WP5_D5.1_Fusion.pdf.
- Service Public de Wallonie (SPW) - Walous. 2020. «Consolidation et enrichissement de la carte d'occupation du sol de Wallonie 2018.» 15. https://geoportail.wallonie.be/files/PDF/annexes_walous/WALOUS_WP6_Consolidation.pdf.
- Service Public de Wallonie (SPW) - Walous. 2018. «Guide méthodologique de production du squelette vectoriel.» 30. https://geoportail.wallonie.be/files/PDF/annexes_walous/WALOUS_WP2_squelettevectoriel.pdf.
- Service Public de Wallonie (SPW) - Walous. 2020. «Validation de la carte d'occupation du sol de Wallonie 2018.» 7. https://geoportail.wallonie.be/files/PDF/annexes_walous/WALOUS_WP6_validation.pdf.
- Wu, Yuxin, et Kaiming He. 2018. «Group Normalization.» *CoRR*. doi:1803.08494.

6. Annexes

6.1. Itération 01

6.1.1. Configuration – Implementation I

```
CUDNN_BENCHMARK: false
DATALOADER:
  ASPECT_RATIO_GROUPING: true
  FILTER_EMPTY_ANNOTATIONS: true
  NUM_WORKERS: 10
  REPEAT_THRESHOLD: 0.0
  SAMPLER_TRAIN: TrainingSampler
DATASETS:
  PRECOMPUTED_PROPOSAL_TOPK_TEST: 1000
  PRECOMPUTED_PROPOSAL_TOPK_TRAIN: 2000
  PROPOSAL_FILES_TEST: []
  PROPOSAL_FILES_TRAIN: []
TEST:
  - walous_sem_seg_val
TRAIN:
  - walous_sem_seg_train
GLOBAL:
  HACK: 1.0
INPUT:
  CROP:
    ENABLED: false
    SINGLE_CATEGORY_MAX_AREA: 1.0
    SIZE:
      - 512
      - 1024
    TYPE: absolute
  FORMAT: RGB
  MASK_FORMAT: polygon
  MAX_SIZE_TEST: 512
  MAX_SIZE_TRAIN: 512
  MIN_SIZE_TEST: 512
  MIN_SIZE_TRAIN:
    - 512
  MIN_SIZE_TRAIN_SAMPLING: choice
  RANDOM_FLIP: horizontal
MODEL:
  ANCHOR_GENERATOR:
    ANGLES:
      - -90
      - 0
      - 90
    ASPECT RATIOS:
      - 0.5
```

```
- 1.0
- 2.0
NAME: DefaultAnchorGenerator
OFFSET: 0.0
SIZES:
- - 32
- - 64
- - 128
- - 256
- - 512
BACKBONE:
  FREEZE_AT: 0
  NAME: build_resnet_deeplab_backbone
DEVICE: cuda
FPN:
  FUSE_TYPE: sum
  IN_FEATURES: []
  NORM: ''
  OUT_CHANNELS: 256
KEYPOINT_ON: false
LOAD_PROPOSALS: false
MASK_ON: false
META_ARCHITECTURE: SemanticSegmentor
PANOPTIC_FPN:
  COMBINE:
    ENABLED: true
    INSTANCES_CONFIDENCE_THRESH: 0.5
    OVERLAP_THRESH: 0.5
    STUFF_AREA_LIMIT: 4096
    INSTANCE_LOSS_WEIGHT: 1.0
PIXEL_MEAN:
- 27108.4928
- 29066.0223
- 26921.2527
- 39606.5559
PIXEL_STD:
- 7118.1259
- 6383.3102
- 6026.0482
- 6578.2162
PROPOSAL_GENERATOR:
  MIN_SIZE: 0
  NAME: RPN
RESNETS:
  DEFORM_MODULATED: false
  DEFORM_NUM_GROUPS: 1
  DEFORM_ON_PER_STAGE:
- false
- false
```

```
- false
- false
DEPTH: 101
NORM: SyncBN
NUM_GROUPS: 1
OUT_FEATURES:
- res2
- res5
RES2_OUT_CHANNELS: 256
RES4_DILATION: 1
RES5_DILATION: 2
RES5_MULTI_GRID:
- 1
- 2
- 4
STEM_OUT_CHANNELS: 128
STEM_TYPE: deeplab
STRIDE_IN_1X1: false
WIDTH_PER_GROUP: 64
RETINANET:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_WEIGHTS: &id001
  - 1.0
  - 1.0
  - 1.0
  - 1.0
  FOCAL_LOSS_ALPHA: 0.25
  FOCAL_LOSS_GAMMA: 2.0
  IN_FEATURES:
  - p3
  - p4
  - p5
  - p6
  - p7
  IOU_LABELS:
  - 0
  - -1
  - 1
  IOU_THRESHOLDS:
  - 0.4
  - 0.5
  NMS_THRESH_TEST: 0.5
  NORM: ''
  NUM_CLASSES: 80
  NUM_CONVS: 4
  PRIOR_PROB: 0.01
  SCORE_THRESH_TEST: 0.05
  SMOOTH_L1_LOSS_BETA: 0.1
  TOPK_CANDIDATES_TEST: 1000
```

```
ROI_BOX_CASCADE_HEAD:
  BBOX_REG_WEIGHTS:
    - 10.0
    - 10.0
    - 5.0
    - 5.0
    - 20.0
    - 20.0
    - 10.0
    - 10.0
    - 30.0
    - 30.0
    - 15.0
    - 15.0
  IOUS:
    - 0.5
    - 0.6
    - 0.7
ROI_BOX_HEAD:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS:
    - 10.0
    - 10.0
    - 5.0
    - 5.0
  CLS_AGNOSTIC_BBOX_REG: false
  CONV_DIM: 256
  FC_DIM: 1024
  NAME: FastRCNNConvFCHead
  NORM: ''
  NUM_CONV: 0
  NUM_FC: 2
  POOLER_RESOLUTION: 7
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAlignV2
  SMOOTH_L1_BETA: 0.0
  TRAIN_ON_PRED_BOXES: false
ROI_HEADS:
  BATCH_SIZE_PER_IMAGE: 512
  IN_FEATURES:
    - res5
  IOU_LABELS:
    - 0
    - 1
  IOU_THRESHOLDS:
    - 0.5
  NAME: StandardROIHeads
  NMS_THRESH_TEST: 0.5
```



```
NUM_CLASSES: 80
POSITIVE_FRACTION: 0.25
PROPOSAL_APPEND_GT: true
SCORE_THRESH_TEST: 0.05
ROI_KEYPOINT_HEAD:
  CONV_DIMS:
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
  LOSS_WEIGHT: 1.0
  MIN_KEYPOINTS_PER_IMAGE: 1
  NAME: KRCNNConvDeconvUpsampleHead
  NORMALIZE_LOSS_BY_VISIBLE_KEYPOINTS: true
  NUM_KEYPOINTS: 17
  POOLER_RESOLUTION: 14
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAlignV2
ROI_MASK_HEAD:
  CLS_AGNOSTIC_MASK: false
  CONV_DIM: 256
  NAME: MaskRCNNConvUpsampleHead
  NORM: ''
  NUM_CONV: 4
  POOLER_RESOLUTION: 14
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAlignV2
RPN:
  BATCH_SIZE_PER_IMAGE: 256
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS: *id001
  BOUNDARY_THRESH: -1
  HEAD_NAME: StandardRPNHead
  IN_FEATURES:
    - res5
  IOU_LABELS:
    - 0
    - -1
    - 1
  IOU_THRESHOLDS:
    - 0.3
    - 0.7
  LOSS_WEIGHT: 1.0
  NMS_THRESH: 0.7
```

```
POSITIVE_FRACTION: 0.5
POST_NMS_TOPK_TEST: 1000
POST_NMS_TOPK_TRAIN: 2000
PRE_NMS_TOPK_TEST: 6000
PRE_NMS_TOPK_TRAIN: 12000
SMOOTH_L1_BETA: 0.0
SEM_SEG_HEAD:
  ASPP_CHANNELS: 256
  ASPP_DILATIONS:
    - 6
    - 12
    - 18
  ASPP_DROPOUT: 0.1
  COMMON_STRIDE: 4
  CONVS_DIM: 256
  IGNORE_VALUE: 255
  IN_FEATURES:
    - res2
    - res5
  LOSS_TYPE: hard_pixel_mining
  LOSS_WEIGHT: 1.0
  NAME: DeepLabV3PlusHead
  NORM: SyncBN
  NUM_CLASSES: 11
  PROJECT_CHANNELS:
    - 48
  PROJECT_FEATURES:
    - res2
  USE_DEPTHWISE_SEPARABLE_CONV: false
  WEIGHTS: detectron2://DeepLab/R-103.pkl
OUTPUT_DIR: output_walious_iter-1
SEED: -1
SOLVER:
  AMP:
    ENABLED: false
  BASE_LR: 0.001
  BIAS_LR_FACTOR: 1.0
  CHECKPOINT_PERIOD: 5000
  CLIP_GRADIENTS:
    CLIP_TYPE: value
    CLIP_VALUE: 1.0
    ENABLED: false
    NORM_TYPE: 2.0
  GAMMA: 0.1
  IMS_PER_BATCH: 16
  LR_SCHEDULER_NAME: WarmupPolyLR
  MAX_ITER: 35000
  MOMENTUM: 0.9
  NESTEROV: false
```

```

POLY_LR_CONSTANT_ENDING: 0.0
POLY_LR_POWER: 0.9
REFERENCE_WORLD_SIZE: 0
STEPS:
- 60000
- 80000
WARMUP_FACTOR: 0.001
WARMUP_ITERS: 1000
WARMUP_METHOD: linear
WEIGHT_DECAY: 0.0001
WEIGHT_DECAY_BIAS: 0.0001
WEIGHT_DECAY_NORM: 0.0
TEST:
AUG:
  ENABLED: false
  FLIP: true
  MAX_SIZE: 4000
  MIN_SIZES:
    - 400
    - 500
    - 600
    - 700
    - 800
    - 900
    - 1000
    - 1100
    - 1200
  DETECTIONS_PER_IMAGE: 100
  EVAL_PERIOD: 1500
  EXPECTED_RESULTS: []
  KEYPOINT_OKS_SIGMAS: []
  PRECISE_BN:
    ENABLED: false
    NUM_ITER: 200
VERSION: 2
VIS_PERIOD: 0

```

6.1.2. Configuration – Implementation II

```

BATCH_SIZE: 1
INPUT:
  BANDS:
    - B02
    - B03
    - B04
    - B05
    - B08

```

```
- B11
INDICES:
- NDVI
- NDWI
- NDMI
- ARI
SEQUENCE_LENGTH: 18
SIZE:
- 200
- 200
MODEL:
  BIDIRECTIONAL: true
  HIDDEN_SIZE: 64
  LOSS_TYPE: CrossEntropyLoss
  LOSS_WEIGHT:
    - 0.9388 # Rest
    - 0.7543 # 6
    - 0.6553 # 7
    - 0.8927 # 8
    - 0.7586 # 9
  NUM_CLASSES: 5
  NUM_LAYERS: 2
  PIXEL_MEAN:
    - 102.06
    - 115.68
    - 92.60
    - 124.67
    - 134.88
    - 134.31
  PIXEL_STD:
    - 44.93
    - 44.63
    - 49.26
    - 45.59
    - 46.33
    - 46.41
  SOLVER:
    BETAS:
      - 0.9
      - 0.999
    LR: 0.001
    MAX_ITER: 25
    NAME: Adam
    WEIGHT_DECAY: 0.0
```

6.2. Itération 02

6.2.1. Configuration – Implementation I

```
CUDNN_BENCHMARK: false
DATALOADER:
  ASPECT_RATIO_GROUPING: true
  FILTER_EMPTY_ANNOTATIONS: true
  NUM_WORKERS: 10
  REPEAT_THRESHOLD: 0.0
  SAMPLER_TRAIN: TrainingSampler
DATASETS:
  PRECOMPUTED_PROPOSAL_TOPK_TEST: 1000
  PRECOMPUTED_PROPOSAL_TOPK_TRAIN: 2000
  PROPOSAL_FILES_TEST: []
  PROPOSAL_FILES_TRAIN: []
  TEST:
    - walous_sem_seg_val
  TRAIN:
    - walous_sem_seg_train
GLOBAL:
  HACK: 1.0
INPUT:
  CROP:
    ENABLED: false
    SINGLE_CATEGORY_MAX_AREA: 1.0
    SIZE:
      - 512
      - 1024
    TYPE: absolute
  FORMAT: RGB
  MASK_FORMAT: polygon
  MAX_SIZE_TEST: 512
  MAX_SIZE_TRAIN: 512
  MIN_SIZE_TEST: 512
  MIN_SIZE_TRAIN:
    - 512
  MIN_SIZE_TRAIN_SAMPLING: choice
  RANDOM_FLIP: horizontal
MODEL:
  ANCHOR_GENERATOR:
    ANGLES:
      - -90
      - 0
      - 90
    ASPECT RATIOS:
      - 0.5
      - 1.0
      - 2.0
    NAME: DefaultAnchorGenerator
    OFFSET: 0.0
```

```
SIZES:
- - 32
- - 64
- - 128
- - 256
- - 512
BACKBONE:
  FREEZE_AT: 0
  NAME: build_resnet_deeplab_backbone
DEVICE: cuda
FPN:
  FUSE_TYPE: sum
  IN_FEATURES: []
  NORM: ''
  OUT_CHANNELS: 256
KEYPOINT_ON: false
LOAD_PROPOSALS: false
MASK_ON: false
META_ARCHITECTURE: SemanticSegmentor
PANOPTIC_FPN:
  COMBINE:
    ENABLED: true
    INSTANCES_CONFIDENCE_THRESH: 0.5
    OVERLAP_THRESH: 0.5
    STUFF_AREA_LIMIT: 4096
  INSTANCE_LOSS_WEIGHT: 1.0
PIXEL_MEAN:
- 27108.4928
- 29066.0223
- 26921.2527
- 39606.5559
- 6.0366
PIXEL_STD:
- 7118.1259
- 6383.3102
- 6026.0482
- 6578.2162
- 6.1443
PROPOSAL_GENERATOR:
  MIN_SIZE: 0
  NAME: RPN
RESNETS:
  DEFORM_MODULATED: false
  DEFORM_NUM_GROUPS: 1
  DEFORM_ON_PER_STAGE:
- false
- false
- false
- false
```



```
DEPTH: 101
NORM: SyncBN
NUM_GROUPS: 1
OUT_FEATURES:
- res2
- res5
RES2_OUT_CHANNELS: 256
RES4_DILATION: 1
RES5_DILATION: 2
RES5_MULTI_GRID:
- 1
- 2
- 4
STEM_OUT_CHANNELS: 128
STEM_TYPE: deeplab
STRIDE_IN_1X1: false
WIDTH_PER_GROUP: 64
RETINANET:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_WEIGHTS: &id001
  - 1.0
  - 1.0
  - 1.0
  - 1.0
  FOCAL_LOSS_ALPHA: 0.25
  FOCAL_LOSS_GAMMA: 2.0
  IN_FEATURES:
  - p3
  - p4
  - p5
  - p6
  - p7
  IOU_LABELS:
  - 0
  - -1
  - 1
  IOU_THRESHOLDS:
  - 0.4
  - 0.5
  NMS_THRESH_TEST: 0.5
  NORM: ''
  NUM_CLASSES: 80
  NUM_CONVS: 4
  PRIOR_PROB: 0.01
  SCORE_THRESH_TEST: 0.05
  SMOOTH_L1_LOSS_BETA: 0.1
  TOPK_CANDIDATES_TEST: 1000
ROI_BOX_CASCADE_HEAD:
  BBOX_REG_WEIGHTS:
```

```

- - 10.0
- 10.0
- 5.0
- 5.0
- - 20.0
- 20.0
- 10.0
- 10.0
- - 30.0
- 30.0
- 15.0
- 15.0

IOUS:
- 0.5
- 0.6
- 0.7

ROI_BOX_HEAD:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS:
    - 10.0
    - 10.0
    - 5.0
    - 5.0
  CLS_AGNOSTIC_BBOX_REG: false
  CONV_DIM: 256
  FC_DIM: 1024
  NAME: FastRCNNConvFCHead
  NORM: ''
  NUM_CONV: 0
  NUM_FC: 2
  POOLER_RESOLUTION: 7
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAlignV2
  SMOOTH_L1_BETA: 0.0
  TRAIN_ON_PRED_BOXES: false

ROI_HEADS:
  BATCH_SIZE_PER_IMAGE: 512
  IN_FEATURES:
    - res5
  IOU_LABELS:
    - 0
    - 1
  IOU_THRESHOLDS:
    - 0.5
  NAME: StandardROIHeads
  NMS_THRESH_TEST: 0.5
  NUM_CLASSES: 80
  POSITIVE_FRACTION: 0.25

```

```
PROPOSAL_APPEND_GT: true
SCORE_THRESH_TEST: 0.05
ROI_KEYPOINT_HEAD:
  CONV_DIMS:
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
  LOSS_WEIGHT: 1.0
  MIN_KEYPOINTS_PER_IMAGE: 1
  NAME: KRCNNConvDeconvUpsampleHead
  NORMALIZE_LOSS_BY_VISIBLE_KEYPOINTS: true
  NUM_KEYPOINTS: 17
  POOLER_RESOLUTION: 14
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAlignV2
ROI_MASK_HEAD:
  CLS_AGNOSTIC_MASK: false
  CONV_DIM: 256
  NAME: MaskRCNNConvUpsampleHead
  NORM: ''
  NUM_CONV: 4
  POOLER_RESOLUTION: 14
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAlignV2
RPN:
  BATCH_SIZE_PER_IMAGE: 256
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS: *id001
  BOUNDARY_THRESH: -1
  HEAD_NAME: StandardRPNHead
  IN_FEATURES:
    - res5
  IOU_LABELS:
    - 0
    - -1
    - 1
  IOU_THRESHOLDS:
    - 0.3
    - 0.7
  LOSS_WEIGHT: 1.0
  NMS_THRESH: 0.7
  POSITIVE_FRACTION: 0.5
  POST_NMS_TOPK_TEST: 1000
```

```
POST_NMS_TOPK_TRAIN: 2000
PRE_NMS_TOPK_TEST: 6000
PRE_NMS_TOPK_TRAIN: 12000
SMOOTH_L1_BETA: 0.0
SEM_SEG_HEAD:
  ASPP_CHANNELS: 256
  ASPP_DILATIONS:
    - 6
    - 12
    - 18
  ASPP_DROPOUT: 0.1
  COMMON_STRIDE: 4
  CONVS_DIM: 256
  IGNORE_VALUE: 255
  IN_FEATURES:
    - res2
    - res5
  LOSS_TYPE: hard_pixel_mining
  LOSS_WEIGHT: 1.0
  NAME: DeepLabV3PlusHead
  NORM: SyncBN
  NUM_CLASSES: 11
  PROJECT_CHANNELS:
    - 48
  PROJECT_FEATURES:
    - res2
  USE_DEPTHWISE_SEPARABLE_CONV: false
  WEIGHTS: model_walous_iter-1_extraconv.pth
OUTPUT_DIR: output_walous_iter-2.RGBIH
SEED: -1
SOLVER:
  AMP:
    ENABLED: false
  BASE_LR: 0.001
  BIAS_LR_FACTOR: 1.0
  CHECKPOINT_PERIOD: 5000
  CLIP_GRADIENTS:
    CLIP_TYPE: value
    CLIP_VALUE: 1.0
    ENABLED: false
    NORM_TYPE: 2.0
  GAMMA: 0.1
  IMS_PER_BATCH: 16
  LR_SCHEDULER_NAME: WarmupPolyLR
  MAX_ITER: 25000
  MOMENTUM: 0.9
  NESTEROV: false
  POLY_LR_CONSTANT_ENDING: 0.0
  POLY_LR_POWER: 0.9
```

```
REFERENCE_WORLD_SIZE: 0
STEPS:
- 60000
- 80000
WARMUP_FACTOR: 0.001
WARMUP_ITERS: 1000
WARMUP_METHOD: linear
WEIGHT_DECAY: 0.0001
WEIGHT_DECAY_BIAS: 0.0001
WEIGHT_DECAY_NORM: 0.0
TEST:
AUG:
  ENABLED: false
  FLIP: true
  MAX_SIZE: 4000
  MIN_SIZES:
  - 400
  - 500
  - 600
  - 700
  - 800
  - 900
  - 1000
  - 1100
  - 1200
DETECTIONS_PER_IMAGE: 100
EVAL_PERIOD: 1000
EXPECTED_RESULTS: []
KEYPOINT_OKS_SIGMAS: []
PRECISE_BN:
  ENABLED: false
  NUM_ITER: 200
VERSION: 2
VIS_PERIOD: 0
```

6.2.2. Configuration – Implementation II

```
CUDNN_BENCHMARK: false
DATALOADER:
  ASPECT_RATIO_GROUPING: true
  FILTER_EMPTY_ANNOTATIONS: true
  NUM_WORKERS: 10
  REPEAT_THRESHOLD: 0.0
  SAMPLER_TRAIN: TrainingSampler
DATASETS:
  PRECOMPUTED_PROPOSAL_TOPK_TEST: 1000
  PRECOMPUTED_PROPOSAL_TOPK_TRAIN: 2000
```

```
PROPOSAL_FILES_TEST: []
PROPOSAL_FILES_TRAIN: []
TEST:
- walous_sem_seg_val
TRAIN:
- walous_sem_seg_train
GLOBAL:
  HACK: 1.0
INPUT:
  CROP:
    ENABLED: false
    SINGLE_CATEGORY_MAX_AREA: 1.0
    SIZE:
      - 512
      - 1024
    TYPE: absolute
  FORMAT: RGB
  MASK_FORMAT: polygon
  MAX_SIZE_TEST: 512
  MAX_SIZE_TRAIN: 512
  MIN_SIZE_TEST: 512
  MIN_SIZE_TRAIN:
    - 512
  MIN_SIZE_TRAIN_SAMPLING: choice
  RANDOM_FLIP: horizontal
MODEL:
  ANCHOR_GENERATOR:
    ANGLES:
      - -90
      - 0
      - 90
    ASPECT RATIOS:
      - 0.5
      - 1.0
      - 2.0
    NAME: DefaultAnchorGenerator
    OFFSET: 0.0
    SIZES:
      - 32
      - 64
      - 128
      - 256
      - 512
  BACKBONE:
    FREEZE_AT: 0
    NAME: build_resnet_deeplab_backbone
  DEVICE: cuda
  FPN:
    FUSE_TYPE: sum
```



```
IN_FEATURES: []
NORM: ''
OUT_CHANNELS: 256
KEYPOINT_ON: false
LOAD_PROPOSALS: false
MASK_ON: false
META_ARCHITECTURE: SemanticSegmentor
PANOPTIC_FPN:
  COMBINE:
    ENABLED: true
    INSTANCES_CONFIDENCE_THRESH: 0.5
    OVERLAP_THRESH: 0.5
    STUFF_AREA_LIMIT: 4096
    INSTANCE_LOSS_WEIGHT: 1.0
PIXEL_MEAN:
- 27108.4928
- 29066.0223
- 26921.2527
- 39606.5559
- 6.0366
PIXEL_STD:
- 7118.1259
- 6383.3102
- 6026.0482
- 6578.2162
- 6.1443
PROPOSAL_GENERATOR:
  MIN_SIZE: 0
  NAME: RPN
RESNETS:
  DEFORM_MODULATED: false
  DEFORM_NUM_GROUPS: 1
  DEFORM_ON_PER_STAGE:
    - false
    - false
    - false
    - false
  DEPTH: 101
  NORM: SyncBN
  NUM_GROUPS: 1
  OUT_FEATURES:
    - res2
    - res5
  RES2_OUT_CHANNELS: 256
  RES4_DILATION: 1
  RES5_DILATION: 2
  RES5_MULTI_GRID:
    - 1
    - 2
```

```
- 4
STEM_OUT_CHANNELS: 128
STEM_TYPE: deeplab
STRIDE_IN_1X1: false
WIDTH_PER_GROUP: 64
RETINANET:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_WEIGHTS: &id001
  - 1.0
  - 1.0
  - 1.0
  - 1.0
  FOCAL_LOSS_ALPHA: 0.25
  FOCAL_LOSS_GAMMA: 2.0
  IN_FEATURES:
    - p3
    - p4
    - p5
    - p6
    - p7
  IOU_LABELS:
    - 0
    - -1
    - 1
  IOU_THRESHOLDS:
    - 0.4
    - 0.5
  NMS_THRESH_TEST: 0.5
  NORM: ''
  NUM_CLASSES: 80
  NUM_CONVS: 4
  PRIOR_PROB: 0.01
  SCORE_THRESH_TEST: 0.05
  SMOOTH_L1_LOSS_BETA: 0.1
  TOPK_CANDIDATES_TEST: 1000
ROI_BOX_CASCADE_HEAD:
  BBOX_REG_WEIGHTS:
    - - 10.0
    - 10.0
    - 5.0
    - 5.0
    - - 20.0
    - 20.0
    - 10.0
    - 10.0
    - - 30.0
    - 30.0
    - 15.0
    - 15.0
```

```

IOUS:
- 0.5
- 0.6
- 0.7
ROI_BOX_HEAD:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS:
    - 10.0
    - 10.0
    - 5.0
    - 5.0
  CLS_AGNOSTIC_BBOX_REG: false
  CONV_DIM: 256
  FC_DIM: 1024
  NAME: FastRCNNConvFCHead
  NORM: ''
  NUM_CONV: 0
  NUM_FC: 2
  POOLER_RESOLUTION: 7
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAlignV2
  SMOOTH_L1_BETA: 0.0
  TRAIN_ON_PRED_BOXES: false
ROI_HEADS:
  BATCH_SIZE_PER_IMAGE: 512
  IN_FEATURES:
    - res5
  IOU_LABELS:
    - 0
    - 1
  IOU_THRESHOLDS:
    - 0.5
  NAME: StandardROIHeads
  NMS_THRESH_TEST: 0.5
  NUM_CLASSES: 80
  POSITIVE_FRACTION: 0.25
  PROPOSAL_APPEND_GT: true
  SCORE_THRESH_TEST: 0.05
ROI_KEYPOINT_HEAD:
  CONV_DIMS:
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512

```

```
LOSS_WEIGHT: 1.0
MIN_KEYPOINTS_PER_IMAGE: 1
NAME: KRCNNConvDeconvUpsampleHead
NORMALIZE_LOSS_BY_VISIBLE_KEYPOINTS: true
NUM_KEYPOINTS: 17
POOLER_RESOLUTION: 14
POOLER_SAMPLING_RATIO: 0
POOLER_TYPE: ROIALignV2
ROI_MASK_HEAD:
  CLS_AGNOSTIC_MASK: false
  CONV_DIM: 256
  NAME: MaskRCNNConvUpsampleHead
  NORM: ''
  NUM_CONV: 4
  POOLER_RESOLUTION: 14
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIALignV2
RPN:
  BATCH_SIZE_PER_IMAGE: 256
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS: *id001
  BOUNDARY_THRESH: -1
  HEAD_NAME: StandardRPNHead
  IN_FEATURES:
    - res5
  IOU_LABELS:
    - 0
    - -1
    - 1
  IOU_THRESHOLDS:
    - 0.3
    - 0.7
  LOSS_WEIGHT: 1.0
  NMS_THRESH: 0.7
  POSITIVE_FRACTION: 0.5
  POST_NMS_TOPK_TEST: 1000
  POST_NMS_TOPK_TRAIN: 2000
  PRE_NMS_TOPK_TEST: 6000
  PRE_NMS_TOPK_TRAIN: 12000
  SMOOTH_L1_BETA: 0.0
SEM_SEG_HEAD:
  ASPP_CHANNELS: 256
  ASPP_DILATIONS:
    - 6
    - 12
    - 18
  ASPP_DROPOUT: 0.1
  COMMON_STRIDE: 4
```

```
CONVS_DIM: 256
IGNORE_VALUE: 255
IN_FEATURES:
- res2
- res5
LOSS_TYPE: hard_pixel_mining
LOSS_WEIGHT: 1.0
NAME: DeepLabV3PlusHead
NORM: SyncBN
NUM_CLASSES: 11
PROJECT_CHANNELS:
- 48
PROJECT_FEATURES:
- res2
USE_DEPTHWISE_SEPARABLE_CONV: false
WEIGHTS: model_walous_iter-2.RGBIH.pth
OUTPUT_DIR: output_walous_iter-2.RGBIH20
SEED: -1
SOLVER:
  AMP:
    ENABLED: false
  BASE_LR: 0.001
  BIAS_LR_FACTOR: 1.0
  CHECKPOINT_PERIOD: 5000
  CLIP_GRADIENTS:
    CLIP_TYPE: value
    CLIP_VALUE: 1.0
    ENABLED: false
    NORM_TYPE: 2.0
  GAMMA: 0.1
  IMS_PER_BATCH: 16
  LR_SCHEDULER_NAME: WarmupPolyLR
  MAX_ITER: 35000
  MOMENTUM: 0.9
  NESTEROV: false
  POLY_LR_CONSTANT_ENDING: 0.0
  POLY_LR_POWER: 0.9
  REFERENCE_WORLD_SIZE: 0
  STEPS:
    - 60000
    - 80000
  WARMUP_FACTOR: 0.001
  WARMUP_ITERS: 1000
  WARMUP_METHOD: linear
  WEIGHT_DECAY: 0.0001
  WEIGHT_DECAY_BIAS: 0.0001
  WEIGHT_DECAY_NORM: 0.0
TEST:
  AUG:
```

```
ENABLED: false
FLIP: true
MAX_SIZE: 4000
MIN_SIZES:
- 400
- 500
- 600
- 700
- 800
- 900
- 1000
- 1100
- 1200
DETECTIONS_PER_IMAGE: 100
EVAL_PERIOD: 1000
EXPECTED_RESULTS: []
KEYPOINT_OKS_SIGMAS: []
PRECISE_BN:
  ENABLED: false
  NUM_ITER: 200
VERSION: 2
VIS_PERIOD: 0
```

6.2.3. Configuration – Implementation III

```
CUDNN_BENCHMARK: false
DATALOADER:
  ASPECT_RATIO_GROUPING: true
  FILTER_EMPTY_ANNOTATIONS: true
  NUM_WORKERS: 10
  REPEAT_THRESHOLD: 0.0
  SAMPLER_TRAIN: TrainingSampler
DATASETS:
  PRECOMPUTED_PROPOSAL_TOPK_TEST: 1000
  PRECOMPUTED_PROPOSAL_TOPK_TRAIN: 2000
  PROPOSAL_FILES_TEST: []
  PROPOSAL_FILES_TRAIN: []
  TEST:
    - walous_sem_seg_val
  TRAIN:
    - walous_sem_seg_train
GLOBAL:
  HACK: 1.0
INPUT:
  CROP:
    ENABLED: false
  SINGLE_CATEGORY_MAX_AREA: 1.0
```

```
SIZE:
- 512
- 1024
TYPE: absolute
FORMAT: RGB
MASK_FORMAT: polygon
MAX_SIZE_TEST: 512
MAX_SIZE_TRAIN: 512
MIN_SIZE_TEST: 512
MIN_SIZE_TRAIN:
- 512
MIN_SIZE_TRAIN_SAMPLING: choice
RANDOM_FLIP: horizontal
MODEL:
ANCHOR_GENERATOR:
  ANGLES:
  - -90
  - 0
  - 90
  ASPECT RATIOS:
  - 0.5
  - 1.0
  - 2.0
  NAME: DefaultAnchorGenerator
  OFFSET: 0.0
  SIZES:
  - 32
  - 64
  - 128
  - 256
  - 512
BACKBONE:
  FREEZE_AT: 0
  NAME: build_resnet_deeplab_backbone
DEVICE: cuda
FPN:
  FUSE_TYPE: sum
  IN_FEATURES: []
  NORM: ''
  OUT_CHANNELS: 256
KEYPOINT_ON: false
LOAD_PROPOSALS: false
MASK_ON: false
META_ARCHITECTURE: SemanticSegmentor
PANOPTIC_FPN:
  COMBINE:
    ENABLED: true
    INSTANCES_CONFIDENCE_THRESH: 0.5
    OVERLAP_THRESH: 0.5
```



```
STUFF_AREA_LIMIT: 4096
INSTANCE_LOSS_WEIGHT: 1.0
PIXEL_MEAN:
- 27108.4928
- 29066.0223
- 26921.2527
- 39606.5559
- 6.0366
- 0
PIXEL_STD:
- 7118.1259
- 6383.3102
- 6026.0482
- 6578.2162
- 6.1443
- 1
PROPOSAL_GENERATOR:
MIN_SIZE: 0
NAME: RPN
RESNETS:
DEFORM_MODULATED: false
DEFORM_NUM_GROUPS: 1
DEFORM_ON_PER_STAGE:
- false
- false
- false
- false
DEPTH: 101
NORM: SyncBN
NUM_GROUPS: 1
OUT_FEATURES:
- res2
- res5
RES2_OUT_CHANNELS: 256
RES4_DILATION: 1
RES5_DILATION: 2
RES5_MULTI_GRID:
- 1
- 2
- 4
STEM_OUT_CHANNELS: 128
STEM_TYPE: deeplab
STRIDE_IN_1X1: false
WIDTH_PER_GROUP: 64
RETINANET:
BBBOX_REG_LOSS_TYPE: smooth_l1
BBBOX_REG_WEIGHTS: &id001
- 1.0
- 1.0
```

```
- 1.0
- 1.0
FOCAL_LOSS_ALPHA: 0.25
FOCAL_LOSS_GAMMA: 2.0
IN_FEATURES:
- p3
- p4
- p5
- p6
- p7
IOU_LABELS:
- 0
- -1
- 1
IOU_THRESHOLDS:
- 0.4
- 0.5
NMS_THRESH_TEST: 0.5
NORM: ''
NUM_CLASSES: 80
NUM_CONVS: 4
PRIOR_PROB: 0.01
SCORE_THRESH_TEST: 0.05
SMOOTH_L1_LOSS_BETA: 0.1
TOPK_CANDIDATES_TEST: 1000
ROI_BOX_CASCADE_HEAD:
  BBOX_REG_WEIGHTS:
    - - 10.0
      - 10.0
      - 5.0
      - 5.0
    - - 20.0
      - 20.0
      - 10.0
      - 10.0
    - - 30.0
      - 30.0
      - 15.0
      - 15.0
  IOUS:
    - 0.5
    - 0.6
    - 0.7
ROI_BOX_HEAD:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS:
    - 10.0
    - 10.0
```

```

- 5.0
- 5.0
CLS_AGNOSTIC_BBOX_REG: false
CONV_DIM: 256
FC_DIM: 1024
NAME: FastRCNNConvFCHead
NORM: ''
NUM_CONV: 0
NUM_FC: 2
POOLER_RESOLUTION: 7
POOLER_SAMPLING_RATIO: 0
POOLER_TYPE: ROIAlignV2
SMOOTH_L1_BETA: 0.0
TRAIN_ON_PRED_BOXES: false
ROI_HEADS:
  BATCH_SIZE_PER_IMAGE: 512
  IN_FEATURES:
    - res5
  IOU_LABELS:
    - 0
    - 1
  IOU_THRESHOLDS:
    - 0.5
  NAME: StandardROIHeads
  NMS_THRESH_TEST: 0.5
  NUM_CLASSES: 80
  POSITIVE_FRACTION: 0.25
  PROPOSAL_APPEND_GT: true
  SCORE_THRESH_TEST: 0.05
ROI_KEYPOINT_HEAD:
  CONV_DIMS:
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
  LOSS_WEIGHT: 1.0
  MIN_KEYPOINTS_PER_IMAGE: 1
  NAME: KRCNNConvDeconvUpsampleHead
  NORMALIZE_LOSS_BY_VISIBLE_KEYPOINTS: true
  NUM_KEYPOINTS: 17
  POOLER_RESOLUTION: 14
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAlignV2
ROI_MASK_HEAD:
  CLS_AGNOSTIC_MASK: false

```

```
CONV_DIM: 256
NAME: MaskRCNNConvUpsampleHead
NORM: ''
NUM_CONV: 4
POOLER_RESOLUTION: 14
POOLER_SAMPLING_RATIO: 0
POOLER_TYPE: ROIAlignV2
RPN:
  BATCH_SIZE_PER_IMAGE: 256
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS: *id001
  BOUNDARY_THRESH: -1
  HEAD_NAME: StandardRPNHead
  IN_FEATURES:
    - res5
  IOU_LABELS:
    - 0
    - -1
    - 1
  IOU_THRESHOLDS:
    - 0.3
    - 0.7
  LOSS_WEIGHT: 1.0
  NMS_THRESH: 0.7
  POSITIVE_FRACTION: 0.5
  POST_NMS_TOPK_TEST: 1000
  POST_NMS_TOPK_TRAIN: 2000
  PRE_NMS_TOPK_TEST: 6000
  PRE_NMS_TOPK_TRAIN: 12000
  SMOOTH_L1_BETA: 0.0
SEM_SEG_HEAD:
  ASPP_CHANNELS: 256
  ASPP_DILATIONS:
    - 6
    - 12
    - 18
  ASPP_DROPOUT: 0.1
  COMMON_STRIDE: 4
  CONVS_DIM: 256
  IGNORE_VALUE: 255
  IN_FEATURES:
    - res2
    - res5
  LOSS_TYPE: hard_pixel_mining
  LOSS_WEIGHT: 1.0
  NAME: DeepLabV3PlusHead
  NORM: SyncBN
  NUM_CLASSES: 11
```

```
PROJECT_CHANNELS:
- 48
PROJECT_FEATURES:
- res2
USE_DEPTHWISE_SEPARABLE_CONV: false
WEIGHTS: model_walous_iter-2.RGBIH20S2.pth
OUTPUT_DIR: output_walous_iter-2.RGBIH20S2
SEED: -1
SOLVER:
  AMP:
    ENABLED: false
  BASE_LR: 0.001
  BIAS_LR_FACTOR: 1.0
  CHECKPOINT_PERIOD: 5000
  CLIP_GRADIENTS:
    CLIP_TYPE: value
    CLIP_VALUE: 1.0
    ENABLED: false
    NORM_TYPE: 2.0
  GAMMA: 0.1
  IMS_PER_BATCH: 16
  LR_SCHEDULER_NAME: WarmupPolyLR
  MAX_ITER: 35000
  MOMENTUM: 0.9
  NESTEROV: false
  POLY_LR_CONSTANT_ENDING: 0.0
  POLY_LR_POWER: 0.9
  REFERENCE_WORLD_SIZE: 0
  STEPS:
    - 60000
    - 80000
  WARMUP_FACTOR: 0.001
  WARMUP_ITERS: 1000
  WARMUP_METHOD: linear
  WEIGHT_DECAY: 0.0001
  WEIGHT_DECAY_BIAS: 0.0001
  WEIGHT_DECAY_NORM: 0.0
TEST:
  AUG:
    ENABLED: false
    FLIP: true
    MAX_SIZE: 4000
    MIN_SIZES:
      - 400
      - 500
      - 600
      - 700
      - 800
      - 900
```

```
- 1000
- 1100
- 1200
DETECTIONS_PER_IMAGE: 100
EVAL_PERIOD: 1000
EXPECTED_RESULTS: []
KEYPOINT_OKS_SIGMAS: []
PRECISE_BN:
  ENABLED: false
  NUM_ITER: 200
VERSION: 2
VIS_PERIOD: 0
```

6.3. Itération 03

6.3.1. Configuration

```
CUDNN_BENCHMARK: false
DATALOADER:
  ASPECT_RATIO_GROUPING: true
  FILTER_EMPTY_ANNOTATIONS: true
  NUM_WORKERS: 10
  REPEAT_THRESHOLD: 0.0
  SAMPLER_TRAIN: TrainingSampler
DATASETS:
  PRECOMPUTED_PROPOSAL_TOPK_TEST: 1000
  PRECOMPUTED_PROPOSAL_TOPK_TRAIN: 2000
  PROPOSAL_FILES_TEST: []
  PROPOSAL_FILES_TRAIN: []
TEST:
  - walous_sem_seg_val
TRAIN:
  - walous_sem_seg_train
GLOBAL:
  HACK: 1.0
INPUT:
  CROP:
    ENABLED: false
    SINGLE_CATEGORY_MAX_AREA: 1.0
  SIZE:
    - 512
    - 1024
  TYPE: absolute
FORMAT: RGB
MASK_FORMAT: polygon
MAX_SIZE_TEST: 512
MAX_SIZE_TRAIN: 512
```



```
MIN_SIZE_TEST: 512
MIN_SIZE_TRAIN:
- 512
MIN_SIZE_TRAIN_SAMPLING: choice
RANDOM_FLIP: horizontal
MODEL:
  ANCHOR_GENERATOR:
    ANGLES:
      - -90
      - 0
      - 90
    ASPECT RATIOS:
      - 0.5
      - 1.0
      - 2.0
    NAME: DefaultAnchorGenerator
    OFFSET: 0.0
    SIZES:
      - 32
      - 64
      - 128
      - 256
      - 512
  BACKBONE:
    FREEZE_AT: 0
    NAME: build_resnet_deeplab_backbone
  DEVICE: cuda
  FPN:
    FUSE_TYPE: sum
    IN_FEATURES: []
    NORM: ''
    OUT_CHANNELS: 256
  KEYPOINT_ON: false
  LOAD_PROPOSALS: false
  MASK_ON: false
  META_ARCHITECTURE: SemanticSegmentor
  PANOPTIC_FPN:
    COMBINE:
      ENABLED: true
      INSTANCES_CONFIDENCE_THRESH: 0.5
      OVERLAP_THRESH: 0.5
      STUFF_AREA_LIMIT: 4096
    INSTANCE_LOSS_WEIGHT: 1.0
  PIXEL_MEAN:
    - 27108.4928
    - 29066.0223
    - 26921.2527
    - 39606.5559
    - 6.0366
```

```
PIXEL_STD:
- 7118.1259
- 6383.3102
- 6026.0482
- 6578.2162
- 6.1443
PROPOSAL_GENERATOR:
  MIN_SIZE: 0
  NAME: RPN
RESNETS:
  DEFORM_MODULATED: false
  DEFORM_NUM_GROUPS: 1
  DEFORM_ON_PER_STAGE:
    - false
    - false
    - false
    - false
  DEPTH: 101
  NORM: SyncBN
  NUM_GROUPS: 1
  OUT_FEATURES:
    - res2
    - res5
  RES2_OUT_CHANNELS: 256
  RES4_DILATION: 1
  RES5_DILATION: 2
  RES5_MULTI_GRID:
    - 1
    - 2
    - 4
  STEM_OUT_CHANNELS: 128
  STEM_TYPE: deeplab
  STRIDE_IN_1X1: false
  WIDTH_PER_GROUP: 64
RETINANET:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_WEIGHTS: &id001
    - 1.0
    - 1.0
    - 1.0
    - 1.0
  FOCAL_LOSS_ALPHA: 0.25
  FOCAL_LOSS_GAMMA: 2.0
  IN_FEATURES:
    - p3
    - p4
    - p5
    - p6
    - p7
```

```
IOU_LABELS:
- 0
- -1
- 1
IOU_THRESHOLDS:
- 0.4
- 0.5
NMS_THRESH_TEST: 0.5
NORM: ''
NUM_CLASSES: 80
NUM_CONVS: 4
PRIOR_PROB: 0.01
SCORE_THRESH_TEST: 0.05
SMOOTH_L1_LOSS_BETA: 0.1
TOPK_CANDIDATES_TEST: 1000
ROI_BOX_CASCADE_HEAD:
  BBOX_REG_WEIGHTS:
    - - 10.0
    - 10.0
    - 5.0
    - 5.0
    - - 20.0
    - 20.0
    - 10.0
    - 10.0
    - - 30.0
    - 30.0
    - 15.0
    - 15.0
  IOUS:
    - 0.5
    - 0.6
    - 0.7
ROI_BOX_HEAD:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS:
    - 10.0
    - 10.0
    - 5.0
    - 5.0
  CLS_AGNOSTIC_BBOX_REG: false
  CONV_DIM: 256
  FC_DIM: 1024
  NAME: FastRCNNConvFCHead
  NORM: ''
  NUM_CONV: 0
  NUM_FC: 2
  POOLER_RESOLUTION: 7
```

```
POOLER_SAMPLING_RATIO: 0
POOLER_TYPE: ROIAlignV2
SMOOTH_L1_BETA: 0.0
TRAIN_ON_PRED_BOXES: false
ROI_HEADS:
  BATCH_SIZE_PER_IMAGE: 512
  IN_FEATURES:
    - res5
  IOU_LABELS:
    - 0
    - 1
  IOU_THRESHOLDS:
    - 0.5
  NAME: StandardROIHeads
  NMS_THRESH_TEST: 0.5
  NUM_CLASSES: 80
  POSITIVE_FRACTION: 0.25
  PROPOSAL_APPEND_GT: true
  SCORE_THRESH_TEST: 0.05
ROI_KEYPOINT_HEAD:
  CONV_DIMS:
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
  LOSS_WEIGHT: 1.0
  MIN_KEYPOINTS_PER_IMAGE: 1
  NAME: KRCNNConvDeconvUpsampleHead
  NORMALIZE_LOSS_BY_VISIBLE_KEYPOINTS: true
  NUM_KEYPOINTS: 17
  POOLER_RESOLUTION: 14
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAlignV2
ROI_MASK_HEAD:
  CLS_AGNOSTIC_MASK: false
  CONV_DIM: 256
  NAME: MaskRCNNConvUpsampleHead
  NORM: ''
  NUM_CONV: 4
  POOLER_RESOLUTION: 14
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAlignV2
RPN:
  BATCH_SIZE_PER_IMAGE: 256
  BBOX_REG_LOSS_TYPE: smooth_l1
```

```
BBOX_REG_LOSS_WEIGHT: 1.0
BBOX_REG_WEIGHTS: *id001
BOUNDARY_THRESH: -1
HEAD_NAME: StandardRPNHead
IN_FEATURES:
- res5
IOU_LABELS:
- 0
- -1
- 1
IOU_THRESHOLDS:
- 0.3
- 0.7
LOSS_WEIGHT: 1.0
NMS_THRESH: 0.7
POSITIVE_FRACTION: 0.5
POST_NMS_TOPK_TEST: 1000
POST_NMS_TOPK_TRAIN: 2000
PRE_NMS_TOPK_TEST: 6000
PRE_NMS_TOPK_TRAIN: 12000
SMOOTH_L1_BETA: 0.0
SEM_SEG_HEAD:
  ASPP_CHANNELS: 256
  ASPP_DILATIONS:
  - 6
  - 12
  - 18
  ASPP_DROPOUT: 0.1
  COMMON_STRIDE: 4
  CONVS_DIM: 256
  IGNORE_VALUE: 255
  IN_FEATURES:
  - res2
  - res5
  LOSS_TYPE: hard_pixel_mining
  LOSS_WEIGHT: 1.0
  NAME: DeepLabV3PlusHead
  NORM: SyncBN
  NUM_CLASSES: 11
  PROJECT_CHANNELS:
  - 48
  PROJECT_FEATURES:
  - res2
  USE_DEPTHWISE_SEPARABLE_CONV: false
WEIGHTS: model_walious_iter-2.RGBIH20.pth
OUTPUT_DIR: output_walious_iter-3
SEED: -1
SOLVER:
  AMP:
```

```
    ENABLED: false
BASE_LR: 0.001
BIAS_LR_FACTOR: 1.0
CHECKPOINT_PERIOD: 5000
CLIP_GRADIENTS:
  CLIP_TYPE: value
  CLIP_VALUE: 1.0
  ENABLED: false
  NORM_TYPE: 2.0
GAMMA: 0.1
IMS_PER_BATCH: 16
LR_SCHEDULER_NAME: WarmupPolyLR
MAX_ITER: 175000
MOMENTUM: 0.9
NESTEROV: false
POLY_LR_CONSTANT_ENDING: 0.0
POLY_LR_POWER: 0.9
REFERENCE_WORLD_SIZE: 0
STEPS:
- 60000
- 80000
WARMUP_FACTOR: 0.001
WARMUP_ITERS: 1000
WARMUP_METHOD: linear
WEIGHT_DECAY: 0.0001
WEIGHT_DECAY_BIAS: 0.0001
WEIGHT_DECAY_NORM: 0.0
TEST:
AUG:
  ENABLED: false
  FLIP: true
  MAX_SIZE: 4000
  MIN_SIZES:
  - 400
  - 500
  - 600
  - 700
  - 800
  - 900
  - 1000
  - 1100
  - 1200
DETECTIONS_PER_IMAGE: 100
EVAL_PERIOD: 1000
EXPECTED_RESULTS: []
KEYPOINT_OKS_SIGMAS: []
PRECISE_BN:
  ENABLED: false
NUM_ITER: 200
```



```
VERSION: 2  
VIS_PERIOD: 0
```

6.4. Itération 04

6.4.1. Configuration – Pre-entrainement

```
CUDNN_BENCHMARK: false  
DATALOADER:  
  ASPECT_RATIO_GROUPING: true  
  FILTER_EMPTY_ANNOTATIONS: true  
  NUM_WORKERS: 4  
  REPEAT_THRESHOLD: 0.0  
  SAMPLER_TRAIN: TrainingSampler  
DATASETS:  
  PRECOMPUTED_PROPOSAL_TOPK_TEST: 1000  
  PRECOMPUTED_PROPOSAL_TOPK_TRAIN: 2000  
  PROPOSAL_FILES_TEST: []  
  PROPOSAL_FILES_TRAIN: []  
  TEST:  
    - walous_sem_seg_val  
  TRAIN:  
    - walous_sem_seg_train  
GLOBAL:  
  HACK: 1.0  
INPUT:  
  COLOR_AUG_SSD: false  
  CROP:  
    ENABLED: false  
    SINGLE_CATEGORY_MAX_AREA: 1.0  
    SIZE:  
      - 0.9  
      - 0.9  
    TYPE: relative_range  
  FORMAT: BGR  
  MASK_FORMAT: polygon  
  MAX_SIZE_TEST: 512  
  MAX_SIZE_TRAIN: 512  
  MIN_SIZE_TEST: 512  
  MIN_SIZE_TRAIN:  
    - 512  
  MIN_SIZE_TRAIN_SAMPLING: choice  
  RANDOM_FLIP: horizontal  
MODEL:  
  ANCHOR_GENERATOR:  
    ANGLES:  
      - -90
```

```

- 0
- 90
ASPECT_RATIOS:
- - 0.5
- 1.0
- 2.0
NAME: DefaultAnchorGenerator
OFFSET: 0.0
SIZES:
- - 32
- - 64
- - 128
- - 256
- - 512
BACKBONE:
FREEZE_AT: 0
NAME: build_resnet_fpn_backbone
DEVICE: cuda
FPN:
FUUSE_TYPE: sum
IN_FEATURES:
- res2
- res3
- res4
- res5
NORM: ''
OUT_CHANNELS: 256
KEYPOINT_ON: false
LOAD_PROPOSALS: false
MASK_ON: false
META_ARCHITECTURE: SemanticSegmentor
PANOPTIC_FPN:
COMBINE:
ENABLED: true
INSTANCES_CONFIDENCE_THRESH: 0.5
OVERLAP_THRESH: 0.5
STUFF_AREA_LIMIT: 4096
INSTANCE_LOSS_WEIGHT: 1.0
PIXEL_MEAN:
- 27108.4928
- 29066.0223
- 26921.2527
- 39606.5559
- 6.0366
PIXEL_STD:
- 7118.1259
- 6383.3102
- 6026.0482
- 6578.2162

```

```
- 6.1443
POINT_HEAD:
  CLS_AGNOSTIC_MASK: false
  COARSE_PRED_EACH_LAYER: false
  COARSE_SEM_SEG_HEAD_NAME: SemSegFPNHead
  FC_DIM: 256
  IMPORTANCE_SAMPLE_RATIO: 0.75
  IN_FEATURES:
    - p2
  NAME: StandardPointHead
  NUM_CLASSES: 11
  NUM_FC: 3
  OVERSAMPLE_RATIO: 3
  SUBDIVISION_NUM_POINTS: 8192
  SUBDIVISION_STEPS: 2
  TRAIN_NUM_POINTS: 2048
PROPOSAL_GENERATOR:
  MIN_SIZE: 0
  NAME: RPN
RESNETS:
  DEFORM_MODULATED: false
  DEFORM_NUM_GROUPS: 1
  DEFORM_ON_PER_STAGE:
    - false
    - false
    - false
    - false
  DEPTH: 50
  NORM: SyncBN
  NUM_GROUPS: 1
  OUT_FEATURES:
    - res2
    - res3
    - res4
    - res5
  RES2_OUT_CHANNELS: 256
  RES5_DILATION: 1
  STEM_OUT_CHANNELS: 64
  STRIDE_IN_1X1: true
  WIDTH_PER_GROUP: 64
RETINANET:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_WEIGHTS: &id001
    - 1.0
    - 1.0
    - 1.0
    - 1.0
  FOCAL_LOSS_ALPHA: 0.25
  FOCAL_LOSS_GAMMA: 2.0
```

```
IN_FEATURES:
- p3
- p4
- p5
- p6
- p7
IOU_LABELS:
- 0
- -1
- 1
IOU_THRESHOLDS:
- 0.4
- 0.5
NMS_THRESH_TEST: 0.5
NORM: ''
NUM_CLASSES: 80
NUM_CONVS: 4
PRIOR_PROB: 0.01
SCORE_THRESH_TEST: 0.05
SMOOTH_L1_LOSS_BETA: 0.1
TOPK_CANDIDATES_TEST: 1000
ROI_BOX_CASCADE_HEAD:
  BBOX_REG_WEIGHTS:
    - - 10.0
    - 10.0
    - 5.0
    - 5.0
    - - 20.0
    - 20.0
    - 10.0
    - 10.0
    - - 30.0
    - 30.0
    - 15.0
    - 15.0
  IOUS:
    - 0.5
    - 0.6
    - 0.7
ROI_BOX_HEAD:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS:
    - 10.0
    - 10.0
    - 5.0
    - 5.0
  CLS_AGNOSTIC_BBOX_REG: false
  CONV_DIM: 256
```

```
FC_DIM: 1024
NAME: FastRCNNConvFCHead
NORM: ''
NUM_CONV: 0
NUM_FC: 2
POOLER_RESOLUTION: 7
POOLER_SAMPLING_RATIO: 0
POOLER_TYPE: ROIALignV2
SMOOTH_L1_BETA: 0.0
TRAIN_ON_PRED_BOXES: false
ROI_HEADS:
  BATCH_SIZE_PER_IMAGE: 512
  IN_FEATURES:
    - p2
    - p3
    - p4
    - p5
  IOU_LABELS:
    - 0
    - 1
  IOU_THRESHOLDS:
    - 0.5
  NAME: StandardROIHeads
  NMS_THRESH_TEST: 0.5
  NUM_CLASSES: 80
  POSITIVE_FRACTION: 0.25
  PROPOSAL_APPEND_GT: true
  SCORE_THRESH_TEST: 0.05
ROI_KEYPOINT_HEAD:
  CONV_DIMS:
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
  LOSS_WEIGHT: 1.0
  MIN_KEYPOINTS_PER_IMAGE: 1
  NAME: KRCNNConvDeconvUpsampleHead
  NORMALIZE_LOSS_BY_VISIBLE_KEYPOINTS: true
  NUM_KEYPOINTS: 17
  POOLER_RESOLUTION: 14
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIALignV2
ROI_MASK_HEAD:
  CLS_AGNOSTIC_MASK: false
  CONV_DIM: 256
```

```
FC_DIM: 1024
IN_FEATURES:
- p2
NAME: MaskRCNNConvUpsampleHead
NORM: ''
NUM_CONV: 4
NUM_FC: 2
OUTPUT_SIDE_RESOLUTION: 7
POINT_HEAD_ON: false
POOLER_RESOLUTION: 14
POOLER_SAMPLING_RATIO: 0
POOLER_TYPE: ROIALignV2
RPN:
  BATCH_SIZE_PER_IMAGE: 256
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS: *id001
  BOUNDARY_THRESH: -1
  HEAD_NAME: StandardRPNHead
  IN_FEATURES:
    - p2
    - p3
    - p4
    - p5
    - p6
  IOU_LABELS:
    - 0
    - -1
    - 1
  IOU_THRESHOLDS:
    - 0.3
    - 0.7
  LOSS_WEIGHT: 1.0
  NMS_THRESH: 0.7
  POSITIVE_FRACTION: 0.5
  POST_NMS_TOPK_TEST: 1000
  POST_NMS_TOPK_TRAIN: 1000
  PRE_NMS_TOPK_TEST: 1000
  PRE_NMS_TOPK_TRAIN: 2000
  SMOOTH_L1_BETA: 0.0
SEM_SEG_HEAD:
  COMMON_STRIDE: 4
  CONVS_DIM: 128
  IGNORE_VALUE: 255
  IN_FEATURES:
    - p2
    - p3
    - p4
    - p5
```

```
LOSS_WEIGHT: 1.0
NAME: PointRendSemSegHead
NORM: GN
NUM_CLASSES: 11
WEIGHTS:
detectron2://PointRend/SemanticSegmentation/pointrend_semantic_R_101_FPN_1x_cit
yscapes/202576688/model_final_cf6ac1.pkl

OUTPUT_DIR: output_walous_iter-4.PRE
SEED: -1
SOLVER:
  AMP:
    ENABLED: false
  BASE_LR: 0.01
  BIAS_LR_FACTOR: 1.0
  CHECKPOINT_PERIOD: 1000
  CLIP_GRADIENTS:
    CLIP_TYPE: value
    CLIP_VALUE: 1.0
    ENABLED: false
    NORM_TYPE: 2.0
  GAMMA: 0.1
  IMS_PER_BATCH: 8
  LR_SCHEDULER_NAME: WarmupMultiStepLR
  MAX_ITER: 35000
  MOMENTUM: 0.9
  NESTEROV: false
  REFERENCE_WORLD_SIZE: 0
  STEPS:
    - 60000
    - 80000
  WARMUP_FACTOR: 0.001
  WARMUP_ITERS: 1000
  WARMUP_METHOD: linear
  WEIGHT_DECAY: 0.0001
  WEIGHT_DECAY_BIAS: 0.0001
  WEIGHT_DECAY_NORM: 0.0
TEST:
  AUG:
    ENABLED: false
    FLIP: true
    MAX_SIZE: 4000
    MIN_SIZES:
      - 400
      - 500
      - 600
      - 700
      - 800
      - 900
```



```
- 1000
- 1100
- 1200
DETECTIONS_PER_IMAGE: 100
EVAL_PERIOD: 1000
EXPECTED_RESULTS: []
KEYPOINT_OKS_SIGMAS: []
PRECISE_BN:
  ENABLED: false
  NUM_ITER: 200
VERSION: 2
VIS_PERIOD: 0
```

6.4.2. Configuration – Affinement

```
CUDNN_BENCHMARK: false
DATALOADER:
  ASPECT_RATIO_GROUPING: true
  FILTER_EMPTY_ANNOTATIONS: true
  NUM_WORKERS: 4
  REPEAT_THRESHOLD: 0.0
  SAMPLER_TRAIN: TrainingSampler
DATASETS:
  PRECOMPUTED_PROPOSAL_TOPK_TEST: 1000
  PRECOMPUTED_PROPOSAL_TOPK_TRAIN: 2000
  PROPOSAL_FILES_TEST: []
  PROPOSAL_FILES_TRAIN: []
  TEST:
    - walous_sem_seg_val
  TRAIN:
    - walous_sem_seg_train
GLOBAL:
  HACK: 1.0
INPUT:
  COLOR_AUG_SSD: false
  CROP:
    ENABLED: false
    SINGLE_CATEGORY_MAX_AREA: 1.0
    SIZE:
      - 0.9
      - 0.9
    TYPE: relative_range
  FORMAT: BGR
  MASK_FORMAT: polygon
  MAX_SIZE_TEST: 512
  MAX_SIZE_TRAIN: 512
  MIN_SIZE_TEST: 512
  MIN_SIZE_TRAIN:
```

```

- 512
MIN_SIZE_TRAIN_SAMPLING: choice
RANDOM_FLIP: horizontal
MODEL:
  ANCHOR_GENERATOR:
    ANGLES:
      - -90
      - 0
      - 90
    ASPECT RATIOS:
      - 0.5
      - 1.0
      - 2.0
    NAME: DefaultAnchorGenerator
    OFFSET: 0.0
    SIZES:
      - 32
      - 64
      - 128
      - 256
      - 512
  BACKBONE:
    FREEZE_AT: 0
    NAME: build_resnet_fpn_backbone
  DEVICE: cuda
  FPN:
    FUSE_TYPE: sum
    IN_FEATURES:
      - res2
      - res3
      - res4
      - res5
    NORM: ''
    OUT_CHANNELS: 256
  KEYPOINT_ON: false
  LOAD_PROPOSALS: false
  MASK_ON: false
  META_ARCHITECTURE: SemanticSegmentor
  PANOPTIC_FPN:
    COMBINE:
      ENABLED: true
      INSTANCES_CONFIDENCE_THRESH: 0.5
      OVERLAP_THRESH: 0.5
      STUFF_AREA_LIMIT: 4096
    INSTANCE_LOSS_WEIGHT: 1.0
  PIXEL_MEAN:
    - 27108.4928
    - 29066.0223
    - 26921.2527

```

```
- 39606.5559
- 6.0366
PIXEL_STD:
- 7118.1259
- 6383.3102
- 6026.0482
- 6578.2162
- 6.1443
POINT_HEAD:
  CLS_AGNOSTIC_MASK: false
  COARSE_PRED_EACH_LAYER: false
  COARSE_SEM_SEG_HEAD_NAME: SemSegFPNHead
  FC_DIM: 256
  IMPORTANCE_SAMPLE_RATIO: 0.75
  IN_FEATURES:
    - p2
  NAME: StandardPointHead
  NUM_CLASSES: 11
  NUM_FC: 3
  OVERSAMPLE_RATIO: 3
  SUBDIVISION_NUM_POINTS: 8192
  SUBDIVISION_STEPS: 2
  TRAIN_NUM_POINTS: 2048
PROPOSAL_GENERATOR:
  MIN_SIZE: 0
  NAME: RPN
RESNETS:
  DEFORM_MODULATED: false
  DEFORM_NUM_GROUPS: 1
  DEFORM_ON_PER_STAGE:
    - false
    - false
    - false
    - false
  DEPTH: 50
  NORM: SyncBN
  NUM_GROUPS: 1
  OUT_FEATURES:
    - res2
    - res3
    - res4
    - res5
  RES2_OUT_CHANNELS: 256
  RES5_DILATION: 1
  STEM_OUT_CHANNELS: 64
  STRIDE_IN_1X1: true
  WIDTH_PER_GROUP: 64
RETINANET:
  BBOX_REG_LOSS_TYPE: smooth_l1
```

```
BBOX_REG_WEIGHTS: &id001
- 1.0
- 1.0
- 1.0
- 1.0
FOCAL_LOSS_ALPHA: 0.25
FOCAL_LOSS_GAMMA: 2.0
IN_FEATURES:
- p3
- p4
- p5
- p6
- p7
IOU_LABELS:
- 0
- -1
- 1
IOU_THRESHOLDS:
- 0.4
- 0.5
NMS_THRESH_TEST: 0.5
NORM: ''
NUM_CLASSES: 80
NUM_CONVS: 4
PRIOR_PROB: 0.01
SCORE_THRESH_TEST: 0.05
SMOOTH_L1_LOSS_BETA: 0.1
TOPK_CANDIDATES_TEST: 1000
ROI_BOX_CASCADE_HEAD:
  BBOX_REG_WEIGHTS:
    - - 10.0
      - 10.0
      - 5.0
      - 5.0
    - - 20.0
      - 20.0
      - 10.0
      - 10.0
    - - 30.0
      - 30.0
      - 15.0
      - 15.0
  IOUS:
    - 0.5
    - 0.6
    - 0.7
ROI_BOX_HEAD:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
```

```
BBOX_REG_WEIGHTS:
- 10.0
- 10.0
- 5.0
- 5.0
CLS_AGNOSTIC_BBOX_REG: false
CONV_DIM: 256
FC_DIM: 1024
NAME: FastRCNNConvFCHead
NORM: ''
NUM_CONV: 0
NUM_FC: 2
POOLER_RESOLUTION: 7
POOLER_SAMPLING_RATIO: 0
POOLER_TYPE: ROIAlignV2
SMOOTH_L1_BETA: 0.0
TRAIN_ON_PRED_BOXES: false
ROI_HEADS:
  BATCH_SIZE_PER_IMAGE: 512
  IN_FEATURES:
    - p2
    - p3
    - p4
    - p5
  IOU_LABELS:
    - 0
    - 1
  IOU_THRESHOLDS:
    - 0.5
  NAME: StandardROIHeads
  NMS_THRESH_TEST: 0.5
  NUM_CLASSES: 80
  POSITIVE_FRACTION: 0.25
  PROPOSAL_APPEND_GT: true
  SCORE_THRESH_TEST: 0.05
ROI_KEYPOINT_HEAD:
  CONV_DIMS:
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
  LOSS_WEIGHT: 1.0
  MIN_KEYPOINTS_PER_IMAGE: 1
  NAME: KRCNNConvDeconvUpsampleHead
  NORMALIZE_LOSS_BY_VISIBLE_KEYPOINTS: true
```

```
NUM_KEYPOINTS: 17
POOLER_RESOLUTION: 14
POOLER_SAMPLING_RATIO: 0
POOLER_TYPE: ROIAlignV2
ROI_MASK_HEAD:
  CLS_AGNOSTIC_MASK: false
  CONV_DIM: 256
  FC_DIM: 1024
  IN_FEATURES:
    - p2
  NAME: MaskRCNNConvUpsampleHead
  NORM: ''
  NUM_CONV: 4
  NUM_FC: 2
  OUTPUT_SIDE_RESOLUTION: 7
  POINT_HEAD_ON: false
  POOLER_RESOLUTION: 14
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAlignV2
RPN:
  BATCH_SIZE_PER_IMAGE: 256
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS: *id001
  BOUNDARY_THRESH: -1
  HEAD_NAME: StandardRPNHead
  IN_FEATURES:
    - p2
    - p3
    - p4
    - p5
    - p6
  IOU_LABELS:
    - 0
    - -1
    - 1
  IOU_THRESHOLDS:
    - 0.3
    - 0.7
  LOSS_WEIGHT: 1.0
  NMS_THRESH: 0.7
  POSITIVE_FRACTION: 0.5
  POST_NMS_TOPK_TEST: 1000
  POST_NMS_TOPK_TRAIN: 1000
  PRE_NMS_TOPK_TEST: 1000
  PRE_NMS_TOPK_TRAIN: 2000
  SMOOTH_L1_BETA: 0.0
SEM_SEG_HEAD:
  COMMON_STRIDE: 4
```

```
CONVS_DIM: 128
IGNORE_VALUE: 255
IN_FEATURES:
- p2
- p3
- p4
- p5
LOSS_WEIGHT: 1.0
NAME: PointRendSemSegHead
NORM: GN
NUM_CLASSES: 11
WEIGHTS: model_walous_iter-4-PR.PRE.pth
OUTPUT_DIR: output_walous_iter-4.AFF
SEED: -1
SOLVER:
  AMP:
    ENABLED: false
  BASE_LR: 0.01
  BIAS_LR_FACTOR: 1.0
  CHECKPOINT_PERIOD: 1000
  CLIP_GRADIENTS:
    CLIP_TYPE: value
    CLIP_VALUE: 1.0
    ENABLED: false
    NORM_TYPE: 2.0
  GAMMA: 0.1
  IMS_PER_BATCH: 8
  LR_SCHEDULER_NAME: WarmupMultiStepLR
  MAX_ITER: 5000
  MOMENTUM: 0.9
  NESTEROV: false
  REFERENCE_WORLD_SIZE: 0
  STEPS:
    - 60000
    - 80000
  WARMUP_FACTOR: 0.001
  WARMUP_ITERS: 1000
  WARMUP_METHOD: linear
  WEIGHT_DECAY: 0.0001
  WEIGHT_DECAY_BIAS: 0.0001
  WEIGHT_DECAY_NORM: 0.0
TEST:
  AUG:
    ENABLED: false
    FLIP: true
    MAX_SIZE: 4000
    MIN_SIZES:
      - 400
      - 500
```



```
- 600
- 700
- 800
- 900
- 1000
- 1100
- 1200
DETECTIONS_PER_IMAGE: 100
EVAL_PERIOD: 1000
EXPECTED_RESULTS: []
KEYPOINT_OKS_SIGMAS: []
PRECISE_BN:
  ENABLED: false
  NUM_ITER: 200
VERSION: 2
VIS_PERIOD: 0
```

6.5. Itération 05

6.5.1. Configuration – Implementation I

```
CUDNN_BENCHMARK: false
DATALOADER:
  ASPECT_RATIO_GROUPING: true
  FILTER_EMPTY_ANNOTATIONS: true
  NUM_WORKERS: 5
  REPEAT_THRESHOLD: 0.0
  SAMPLER_TRAIN: TrainingSampler
DATASETS:
  PRECOMPUTED_PROPOSAL_TOPK_TEST: 1000
  PRECOMPUTED_PROPOSAL_TOPK_TRAIN: 2000
  PROPOSAL_FILES_TEST: []
  PROPOSAL_FILES_TRAIN: []
  TEST:
    - walous_sem_seg_val
  TRAIN:
    - walous_sem_seg_train
GLOBAL:
  HACK: 1.0
INPUT:
  CROP:
    ENABLED: false
    SINGLE_CATEGORY_MAX_AREA: 1.0
    SIZE:
      - 512
      - 1024
    TYPE: absolute
```

```
FORMAT: RGB
MASK_FORMAT: polygon
MAX_SIZE_TEST: 512
MAX_SIZE_TRAIN: 512
MIN_SIZE_TEST: 512
MIN_SIZE_TRAIN:
- 512
MIN_SIZE_TRAIN_SAMPLING: choice
RANDOM_FLIP: horizontal
MODEL:
  ANCHOR_GENERATOR:
    ANGLES:
      - -90
      - 0
      - 90
    ASPECT RATIOS:
      - 0.5
      - 1.0
      - 2.0
    NAME: DefaultAnchorGenerator
    OFFSET: 0.0
    SIZES:
      - 32
      - 64
      - 128
      - 256
      - 512
  BACKBONE:
    FREEZE_AT: 0
    NAME: build_resnet_deeplab_backbone
  DEVICE: cuda
  FPN:
    FUSE_TYPE: sum
    IN_FEATURES: []
    NORM: ''
    OUT_CHANNELS: 256
  KEYPOINT_ON: false
  LOAD_PROPOSALS: false
  MASK_ON: false
  META_ARCHITECTURE: SemanticSegmentor
  PANOPTIC_FPN:
    COMBINE:
      ENABLED: true
      INSTANCES_CONFIDENCE_THRESH: 0.5
      OVERLAP_THRESH: 0.5
      STUFF_AREA_LIMIT: 4096
    INSTANCE_LOSS_WEIGHT: 1.0
  PIXEL_MEAN:
    - 24789.973
```

```

- 26193.802
- 24536.253
- 37894.432
- 4.0130
PIXEL_STD:
- 7440.666
- 6546.848
- 6115.053
- 6647.631
- 5.8717
PROPOSAL_GENERATOR:
  MIN_SIZE: 0
  NAME: RPN
RESNETS:
  DEFORM_MODULATED: false
  DEFORM_NUM_GROUPS: 1
  DEFORM_ON_PER_STAGE:
    - false
    - false
    - false
    - false
  DEPTH: 101
  NORM: SyncBN
  NUM_GROUPS: 1
  OUT_FEATURES:
    - res2
    - res5
  RES2_OUT_CHANNELS: 256
  RES4_DILATION: 1
  RES5_DILATION: 2
  RES5_MULTI_GRID:
    - 1
    - 2
    - 4
  STEM_OUT_CHANNELS: 128
  STEM_TYPE: deeplab
  STRIDE_IN_1X1: false
  WIDTH_PER_GROUP: 64
RETINANET:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_WEIGHTS: &id001
    - 1.0
    - 1.0
    - 1.0
    - 1.0
  FOCAL_LOSS_ALPHA: 0.25
  FOCAL_LOSS_GAMMA: 2.0
  IN_FEATURES:
    - p3

```

```

- p4
- p5
- p6
- p7
IOU_LABELS:
- 0
- -1
- 1
IOU_THRESHOLDS:
- 0.4
- 0.5
NMS_THRESH_TEST: 0.5
NORM: ''
NUM_CLASSES: 80
NUM_CONVS: 4
PRIOR_PROB: 0.01
SCORE_THRESH_TEST: 0.05
SMOOTH_L1_LOSS_BETA: 0.1
TOPK_CANDIDATES_TEST: 1000
ROI_BOX_CASCADE_HEAD:
  BBOX_REG_WEIGHTS:
    - - 10.0
    - 10.0
    - 5.0
    - 5.0
    - - 20.0
    - 20.0
    - 10.0
    - 10.0
    - - 30.0
    - 30.0
    - 15.0
    - 15.0
  IOUS:
    - 0.5
    - 0.6
    - 0.7
ROI_BOX_HEAD:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS:
    - 10.0
    - 10.0
    - 5.0
    - 5.0
  CLS_AGNOSTIC_BBOX_REG: false
  CONV_DIM: 256
  FC_DIM: 1024
  NAME: FastRCNNConvFCHead

```

```
NORM: ''
NUM_CONV: 0
NUM_FC: 2
POOLER_RESOLUTION: 7
POOLER_SAMPLING_RATIO: 0
POOLER_TYPE: ROIAAlignV2
SMOOTH_L1_BETA: 0.0
TRAIN_ON_PRED_BOXES: false
ROI_HEADS:
  BATCH_SIZE_PER_IMAGE: 512
  IN_FEATURES:
    - res5
  IOU_LABELS:
    - 0
    - 1
  IOU_THRESHOLDS:
    - 0.5
  NAME: StandardROIHeads
  NMS_THRESH_TEST: 0.5
  NUM_CLASSES: 80
  POSITIVE_FRACTION: 0.25
  PROPOSAL_APPEND_GT: true
  SCORE_THRESH_TEST: 0.05
ROI_KEYPOINT_HEAD:
  CONV_DIMS:
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
  LOSS_WEIGHT: 1.0
  MIN_KEYPOINTS_PER_IMAGE: 1
  NAME: KRCNNConvDeconvUpsampleHead
  NORMALIZE_LOSS_BY_VISIBLE_KEYPOINTS: true
  NUM_KEYPOINTS: 17
  POOLER_RESOLUTION: 14
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAAlignV2
ROI_MASK_HEAD:
  CLS_AGNOSTIC_MASK: false
  CONV_DIM: 256
  NAME: MaskRCNNConvUpsampleHead
  NORM: ''
  NUM_CONV: 4
  POOLER_RESOLUTION: 14
  POOLER_SAMPLING_RATIO: 0
```

```
POOLER_TYPE: ROIAlignV2
RPN:
  BATCH_SIZE_PER_IMAGE: 256
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS: *id001
  BOUNDARY_THRESH: -1
  HEAD_NAME: StandardRPNHead
  IN_FEATURES:
    - res5
  IOU_LABELS:
    - 0
    - -1
    - 1
  IOU_THRESHOLDS:
    - 0.3
    - 0.7
  LOSS_WEIGHT: 1.0
  NMS_THRESH: 0.7
  POSITIVE_FRACTION: 0.5
  POST_NMS_TOPK_TEST: 1000
  POST_NMS_TOPK_TRAIN: 2000
  PRE_NMS_TOPK_TEST: 6000
  PRE_NMS_TOPK_TRAIN: 12000
  SMOOTH_L1_BETA: 0.0
SEM_SEG_HEAD:
  ASPP_CHANNELS: 256
  ASPP_DILATIONS:
    - 6
    - 12
    - 18
  ASPP_DROPOUT: 0.1
  COMMON_STRIDE: 4
  CONVS_DIM: 256
  IGNORE_VALUE: 255
  IN_FEATURES:
    - res2
    - res5
  LOSS_TYPE: hard_pixel_mining
  LOSS_WEIGHT: 1.0
  NAME: DeepLabV3PlusHead
  NORM: SyncBN
  NUM_CLASSES: 11
  PROJECT_CHANNELS:
    - 48
  PROJECT_FEATURES:
    - res2
  USE_DEPTHWISE_SEPARABLE_CONV: false
WEIGHTS: model_walous_iter-3.RGBH20.pth
```

```
OUTPUT_DIR: output_walous_iter-5.2020only
```

```
SEED: -1
```

```
SOLVER:
```

```
AMP:
```

```
  ENABLED: false
```

```
BASE_LR: 0.001
```

```
BIAS_LR_FACTOR: 1.0
```

```
CHECKPOINT_PERIOD: 5000
```

```
CLIP_GRADIENTS:
```

```
  CLIP_TYPE: value
```

```
  CLIP_VALUE: 1.0
```

```
  ENABLED: false
```

```
  NORM_TYPE: 2.0
```

```
GAMMA: 0.1
```

```
IMS_PER_BATCH: 16
```

```
LR_SCHEDULER_NAME: WarmupPolyLR
```

```
MAX_ITER: 35000
```

```
MOMENTUM: 0.9
```

```
NESTEROV: false
```

```
POLY_LR_CONSTANT_ENDING: 0.0
```

```
POLY_LR_POWER: 0.9
```

```
REFERENCE_WORLD_SIZE: 0
```

```
STEPS:
```

```
- 60000
```

```
- 80000
```

```
WARMUP_FACTOR: 0.001
```

```
WARMUP_ITERS: 1000
```

```
WARMUP_METHOD: linear
```

```
WEIGHT_DECAY: 0.0001
```

```
WEIGHT_DECAY_BIAS: 0.0001
```

```
WEIGHT_DECAY_NORM: 0.0
```

```
TEST:
```

```
AUG:
```

```
  ENABLED: false
```

```
  FLIP: true
```

```
  MAX_SIZE: 4000
```

```
  MIN_SIZES:
```

```
    - 400
```

```
    - 500
```

```
    - 600
```

```
    - 700
```

```
    - 800
```

```
    - 900
```

```
    - 1000
```

```
    - 1100
```

```
    - 1200
```

```
DETECTIONS_PER_IMAGE: 100
```

```
EVAL_PERIOD: 1000
```

```
EXPECTED_RESULTS: []
```



```
KEYPOINT_OKS_SIGMAS: []  
PRECISE_BN:  
  ENABLED: false  
  NUM_ITER: 200  
VERSION: 2  
VIS_PERIOD: 0
```

6.5.2. Configuration – Implementation II

```
CUDNN_BENCHMARK: false  
DATALOADER:  
  ASPECT_RATIO_GROUPING: true  
  FILTER_EMPTY_ANNOTATIONS: true  
  NUM_WORKERS: 5  
  REPEAT_THRESHOLD: 0.0  
  SAMPLER_TRAIN: TrainingSampler  
DATASETS:  
  PRECOMPUTED_PROPOSAL_TOPK_TEST: 1000  
  PRECOMPUTED_PROPOSAL_TOPK_TRAIN: 2000  
  PROPOSAL_FILES_TEST: []  
  PROPOSAL_FILES_TRAIN: []  
  TEST:  
    - walous_sem_seg_val  
  TRAIN:  
    - walous_sem_seg_train  
GLOBAL:  
  HACK: 1.0  
INPUT:  
  CROP:  
    ENABLED: false  
    SINGLE_CATEGORY_MAX_AREA: 1.0  
    SIZE:  
      - 512  
      - 1024  
    TYPE: absolute  
  FORMAT: RGB  
  MASK_FORMAT: polygon  
  MAX_SIZE_TEST: 512  
  MAX_SIZE_TRAIN: 512  
  MIN_SIZE_TEST: 512  
  MIN_SIZE_TRAIN:  
    - 512  
  MIN_SIZE_TRAIN_SAMPLING: choice  
  RANDOM_FLIP: horizontal  
MODEL:  
  ANCHOR_GENERATOR:  
  ANGLES:
```

```
- - -90
- 0
- 90
ASPECT_RATIOS:
- - 0.5
- 1.0
- 2.0
NAME: DefaultAnchorGenerator
OFFSET: 0.0
SIZES:
- - 32
- 64
- 128
- 256
- 512
BACKBONE:
  FREEZE_AT: 0
  NAME: build_resnet_deeplab_backbone
DEVICE: cuda
FPN:
  FUSE_TYPE: sum
  IN_FEATURES: []
  NORM: ''
  OUT_CHANNELS: 256
KEYPOINT_ON: false
LOAD_PROPOSALS: false
MASK_ON: false
META_ARCHITECTURE: SemanticSegmentor
PANOPTIC_FPN:
  COMBINE:
    ENABLED: true
    INSTANCES_CONFIDENCE_THRESH: 0.5
    OVERLAP_THRESH: 0.5
    STUFF_AREA_LIMIT: 4096
    INSTANCE_LOSS_WEIGHT: 1.0
PIXEL_MEAN:
- 27108.4928
- 29066.0223
- 26921.2527
- 39606.5559
- 6.0366
PIXEL_STD:
- 7118.1259
- 6383.3102
- 6026.0482
- 6578.2162
- 6.1443
PROPOSAL_GENERATOR:
  MIN_SIZE: 0
```

```
NAME: RPN
RESNETS:
  DEFORM_MODULATED: false
  DEFORM_NUM_GROUPS: 1
  DEFORM_ON_PER_STAGE:
    - false
    - false
    - false
    - false
  DEPTH: 101
  NORM: SyncBN
  NUM_GROUPS: 1
  OUT_FEATURES:
    - res2
    - res5
  RES2_OUT_CHANNELS: 256
  RES4_DILATION: 1
  RES5_DILATION: 2
  RES5_MULTI_GRID:
    - 1
    - 2
    - 4
  STEM_OUT_CHANNELS: 128
  STEM_TYPE: deeplab
  STRIDE_IN_1X1: false
  WIDTH_PER_GROUP: 64
RETINANET:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_WEIGHTS: &id001
    - 1.0
    - 1.0
    - 1.0
    - 1.0
  FOCAL_LOSS_ALPHA: 0.25
  FOCAL_LOSS_GAMMA: 2.0
  IN_FEATURES:
    - p3
    - p4
    - p5
    - p6
    - p7
  IOU_LABELS:
    - 0
    - -1
    - 1
  IOU_THRESHOLDS:
    - 0.4
    - 0.5
  NMS_THRESH_TEST: 0.5
```

```

NORM: ''
NUM_CLASSES: 80
NUM_CONVS: 4
PRIOR_PROB: 0.01
SCORE_THRESH_TEST: 0.05
SMOOTH_L1_LOSS_BETA: 0.1
TOPK_CANDIDATES_TEST: 1000
ROI_BOX_CASCADE_HEAD:
  BBOX_REG_WEIGHTS:
    - - 10.0
      - 10.0
      - 5.0
      - 5.0
    - - 20.0
      - 20.0
      - 10.0
      - 10.0
    - - 30.0
      - 30.0
      - 15.0
      - 15.0
  IOUS:
    - 0.5
    - 0.6
    - 0.7
ROI_BOX_HEAD:
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS:
    - 10.0
    - 10.0
    - 5.0
    - 5.0
  CLS_AGNOSTIC_BBOX_REG: false
  CONV_DIM: 256
  FC_DIM: 1024
  NAME: FastRCNNConvFCHead
  NORM: ''
  NUM_CONV: 0
  NUM_FC: 2
  POOLER_RESOLUTION: 7
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAlignV2
  SMOOTH_L1_BETA: 0.0
  TRAIN_ON_PRED_BOXES: false
ROI_HEADS:
  BATCH_SIZE_PER_IMAGE: 512
  IN_FEATURES:
    - res5

```

```
IOU_LABELS:
- 0
- 1
IOU_THRESHOLDS:
- 0.5
NAME: StandardROIHeads
NMS_THRESH_TEST: 0.5
NUM_CLASSES: 80
POSITIVE_FRACTION: 0.25
PROPOSAL_APPEND_GT: true
SCORE_THRESH_TEST: 0.05
ROI_KEYPOINT_HEAD:
  CONV_DIMS:
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
    - 512
  LOSS_WEIGHT: 1.0
  MIN_KEYPOINTS_PER_IMAGE: 1
  NAME: KRCNNConvDeconvUpsampleHead
  NORMALIZE_LOSS_BY_VISIBLE_KEYPOINTS: true
  NUM_KEYPOINTS: 17
  POOLER_RESOLUTION: 14
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAlignV2
ROI_MASK_HEAD:
  CLS_AGNOSTIC_MASK: false
  CONV_DIM: 256
  NAME: MaskRCNNConvUpsampleHead
  NORM: ''
  NUM_CONV: 4
  POOLER_RESOLUTION: 14
  POOLER_SAMPLING_RATIO: 0
  POOLER_TYPE: ROIAlignV2
RPN:
  BATCH_SIZE_PER_IMAGE: 256
  BBOX_REG_LOSS_TYPE: smooth_l1
  BBOX_REG_LOSS_WEIGHT: 1.0
  BBOX_REG_WEIGHTS: *id001
  BOUNDARY_THRESH: -1
  HEAD_NAME: StandardRPNHead
  IN_FEATURES:
    - res5
  IOU_LABELS:
    - 0
```

```
- -1
- 1
IOU_THRESHOLDS:
- 0.3
- 0.7
LOSS_WEIGHT: 1.0
NMS_THRESH: 0.7
POSITIVE_FRACTION: 0.5
POST_NMS_TOPK_TEST: 1000
POST_NMS_TOPK_TRAIN: 2000
PRE_NMS_TOPK_TEST: 6000
PRE_NMS_TOPK_TRAIN: 12000
SMOOTH_L1_BETA: 0.0
SEM_SEG_HEAD:
  ASPP_CHANNELS: 256
  ASPP_DILATIONS:
    - 6
    - 12
    - 18
  ASPP_DROPOUT: 0.1
  COMMON_STRIDE: 4
  CONVS_DIM: 256
  IGNORE_VALUE: 255
  IN_FEATURES:
    - res2
    - res5
  LOSS_TYPE: hard_pixel_mining
  LOSS_WEIGHT: 1.0
  NAME: DeepLabV3PlusHead
  NORM: SyncBN
  NUM_CLASSES: 11
  PROJECT_CHANNELS:
    - 48
  PROJECT_FEATURES:
    - res2
  USE_DEPTHWISE_SEPARABLE_CONV: false
  WEIGHTS: model_walious_iter-3.RGBIH20.pth
OUTPUT_DIR: output_walious_iter-5.AFF
SEED: -1
SOLVER:
  AMP:
    ENABLED: false
  BASE_LR: 0.001
  BIAS_LR_FACTOR: 1.0
  CHECKPOINT_PERIOD: 2500
  CLIP_GRADIENTS:
    CLIP_TYPE: value
    CLIP_VALUE: 1.0
    ENABLED: false
```

```
NORM_TYPE: 2.0
GAMMA: 0.1
IMS_PER_BATCH: 8
LR_SCHEDULER_NAME: WarmupPolyLR
MAX_ITER: 10000
MOMENTUM: 0.9
NESTEROV: false
POLY_LR_CONSTANT_ENDING: 0.0
POLY_LR_POWER: 0.9
REFERENCE_WORLD_SIZE: 0
STEPS:
- 60000
- 80000
WARMUP_FACTOR: 0.001
WARMUP_ITERS: 1000
WARMUP_METHOD: linear
WEIGHT_DECAY: 0.0001
WEIGHT_DECAY_BIAS: 0.0001
WEIGHT_DECAY_NORM: 0.0
TEST:
AUG:
  ENABLED: false
  FLIP: true
  MAX_SIZE: 4000
  MIN_SIZES:
  - 400
  - 500
  - 600
  - 700
  - 800
  - 900
  - 1000
  - 1100
  - 1200
DETECTIONS_PER_IMAGE: 100
EVAL_PERIOD: 2500
EXPECTED_RESULTS: []
KEYPOINT_OKS_SIGMAS: []
PRECISE_BN:
  ENABLED: false
  NUM_ITER: 200
VERSION: 2
VIS_PERIOD: 0
```