

Description fonctionnelle du processus de production et de mise à jour des données géographiques de l'Occupation du Sol Grande Echelle Nouvelle Génération

Table des matières

1. Généralités	3
Processus de constitution initiale.....	4
Synoptique du processus.....	4
Principe de l'apprentissage profond ou <i>deep learning</i>	4
Constitution d'annotations	6
Objectif de l'étape de constitution d'annotations	6
Les filières de production	6
Constitution d'annotations	7
Classification par Machine Learning.....	10
Exploitation de bases de données externes.....	11
Phase d'entraînement des modèles <i>deep learning</i>	14
Phase de détection des modèles <i>deep learning</i>	14
Traitement vecteur.....	15
Vectorisation	15
Préparation des données	17
Constitution d'une ossature de type ORFHB.....	17
Constitution de l'OCS GE	21
Processus de mise à jour	26
Synoptique du processus.....	26
Principes	26
Étapes principales.....	26
Préparation d'annotations	27
Entraînement non supervisé	27
Détection de changements	28
Mise à jour de l'OCS GE	28

1. Généralités

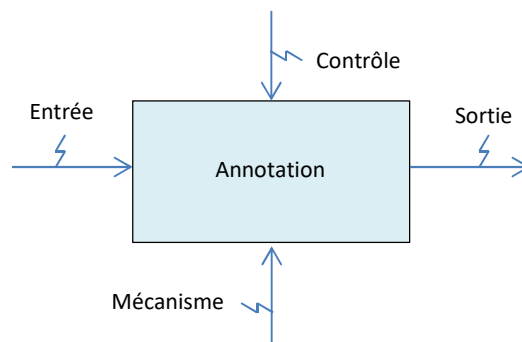
L'Occupation du Sol Grande Echelle dite de « Nouvelle Génération » (OCS GE NG) est baptisée ainsi en raison de son mode de production largement automatisé, s'appuyant notamment sur les techniques d'Intelligence Artificielle d'apprentissage profond, ou *deep learning*. Deux processus distincts entrent en jeu : celui consistant à constituer une première version, dit « **processus de constitution initiale** » ; un second appelé « **processus de mise à jour** ».

La production d'une OCS GE NG commence donc par la mise en œuvre d'un processus de constitution initiale se basant sur des images issues d'une campagne de prises de vues aériennes, afin d'en construire une version millésimée, auquel peuvent succéder plusieurs processus de mise à jour pour la fabrication de versions d'autres millésimes (antérieurs ou ultérieurs).

A ce stade des travaux, les caractéristiques du produit ne correspondent pas rigoureusement à celles de l'OCS GE classique. Le choix a été pris d'assurer la qualité des classes d'objet utiles pour la mesure de l'évolution de l'artificialisation des sols.

Les processus décrits sont ceux mis en œuvre ou évalués en date de la rédaction de ce document.

Les fonctions décrites sont représentées graphiquement sous formes de diagrammes SADT simplifiés :



Processus de constitution initiale

Synoptique du processus

Principe de l'apprentissage profond ou *deep learning*

Par rapport au processus classique de production de l'OCS GE, une partie est confiée à un système de classification automatique utilisant des techniques d'apprentissage profond (*deep learning*). Une instance d'un tel système est appelée ici « modèle », dont le cycle de vie comprend :

- une phase dite « **d'entraînement** » (ou « d'apprentissage ») au cours de laquelle des données d'entraînement, encore appelées « annotations », alimentent un processus mathématique ajustant progressivement les paramètres du modèle afin qu'il puisse devenir opérationnel pour la phase suivante ;
- qui alterne avec une phase dite « **de détection** » (ou encore « **d'inférence** »), au cours de laquelle le modèle est capable de générer des « cartes de chaleur », représentations raster brutes des classes spécifiquement choisies au préalable.

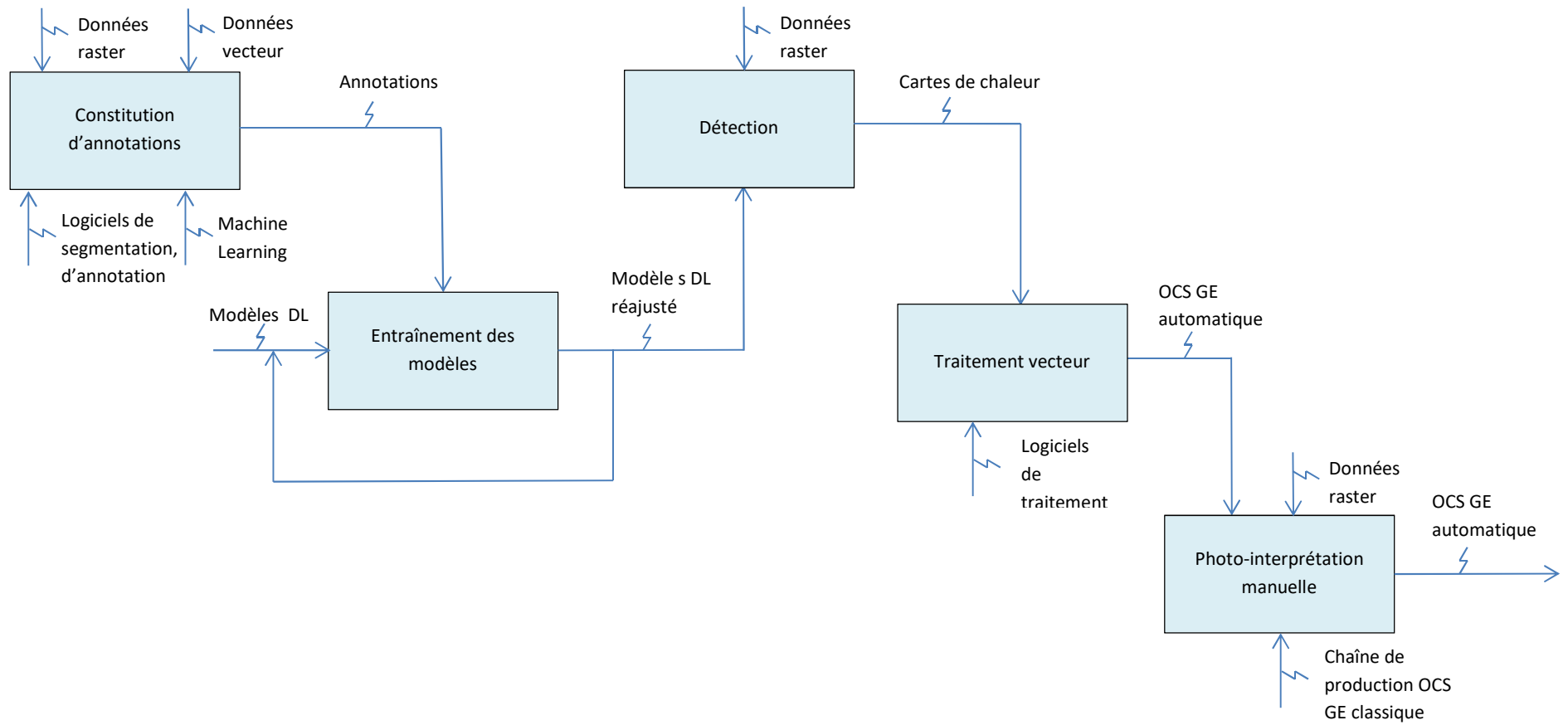
Une première phase d'entraînement doit nécessairement précéder une phase de détection. Toutefois, il serait possible de relancer après une phase de détection une nouvelle phase d'entraînement afin de consolider les performances du modèle, à mesure que l'on dispose de nouvelles annotations. Cette option n'a pas encore été mise en œuvre.

Etapes principales

Les données résultant de la détection doivent être post-traitées afin de constituer un produit de type OCS GE. Elles peuvent ensuite être corrigées interactivement afin d'en améliorer la qualité. Les grandes étapes du processus sont donc :

- la constitution de données d'annotations ;
- la phase d'entraînement des modèles *deep learning* ;
- la phase de détection des modèles *deep learning* ;
- le traitement vecteur des données issues du *deep learning* ;
- la photo-interprétation manuelle.

Constitution initiale



Constitution d'annotations

Objectif de l'étape de constitution d'annotations

Le principe d'apprentissage retenu dans la partie « apprentissage profond » du processus de constitution initiale est celui d'un apprentissage supervisé. Dans ce cas, les modèles de classification retenus doivent être paramétrés sur la base d'exemples qui leur permettent de s'entraîner, et ainsi de renforcer leur efficacité. Ces annotations (dans ce cas, il s'agit de masques de classes qu'on cherche à détecter) sont le reflet de la vérité du terrain et doivent correspondre à des résultats de classification normalement attendus par les modèles en fonction des ortho-images qui leur sont soumises.

Deux listes de classes à détecter par les modèles *deep learning* ont été dressées :

- **en milieu urbain**, il s'agit des classes suivantes : bâti, eau, bitume, ligneux, herbacée, matériaux minéraux, piscines, sols nus ;
- **en zone naturelle, agricole et forestière (NAF)**, 16 classes différentes doivent être discriminées : coupe, culture, lande ligneuse, pelouse, peuplement de conifères, peuplement feuillus, terre arable, vigne, bâtiment, neige, sol nu, surface d'eau, surface inondable, zone imperméable, zone perméable, autres.

Ces listes ne sont pas figées pour autant. Elles ont vocation à évoluer au fil de l'amélioration des processus.

La première étape consiste donc à préparer l'ensemble des annotations, ou données-vérité, car issues de produits contrôlés, considérés comme conformes à la réalité du terrain. A partir de données disponibles, issues de processus de production tiers, sa finalité consiste à constituer des jeux de données aptes à servir d'exemples pour l'entraînement d'un modèle *deep learning*.

Les filières de production

Les annotations sont issues de trois filières distinctes :

- de la saisie interactive d'annotations ;
- de la classification par Machine Learning ;
- de l'exploitation de bases de données externes au processus : RPG, BD TOPO®.

Les données produites par ces filières sont vectorielles. Elles sont agrégées, puis transformées en images binaires représentant des masques thématiques.



Exemple de masque binaire du thème « végétation haute »



Même zone avec une ortho-image de contexte

Pourquoi n'utilise-t-on pas l'OCSGE déjà produite comme donnée d'apprentissage ?

Parce que l'OCSGE est un produit généralisé (les petites zones sont agrégées avec les zones voisines pour respecter des seuils de surface ou de largeur minimum, des règles de priorité entre classes ou concernant l'usage). Idéalement, on aimerait que le modèle deep learning permette de passer directement de l'image brute à une OC SGE respectant toutes les spécifications du produit (auquel cas des échantillons du produit constitueraient les meilleures données d'apprentissage possible) mais en pratique, le travail de généralisation est encore plus compliqué à "apprendre" que le travail d'interprétation de l'image, et on préfère se focaliser sur ce dernier, ce qui suppose d'avoir des données d'apprentissage non généralisées (aussi proche que possible de l'image brute).

Constitution d'annotations

Création d'une pyramide de segmentations

Elle se fait en 2 étapes. La première s'opère au moyen d'un logiciel automatique développé par l'IGN, Pyram, qui crée une pyramide de segmentations d'ortho-images RVB sur des critères radiométriques. Chaque niveau de la pyramide appartient à une segmentation plus ou moins fine. Les segments de chaque niveau se superposent.



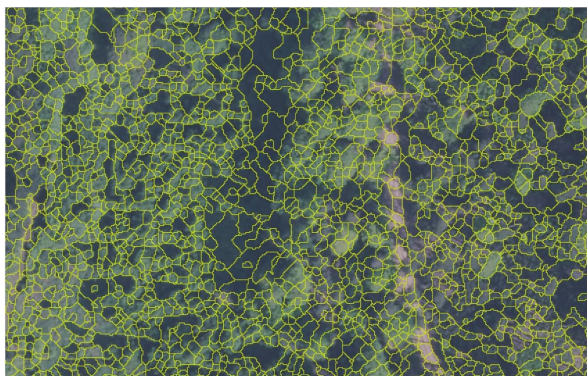
Ortho-image brute



Segmentation de niveau
1 (la plus large)



Segmentation de niveau
4



Segmentation de niveau
6 (la plus fine)

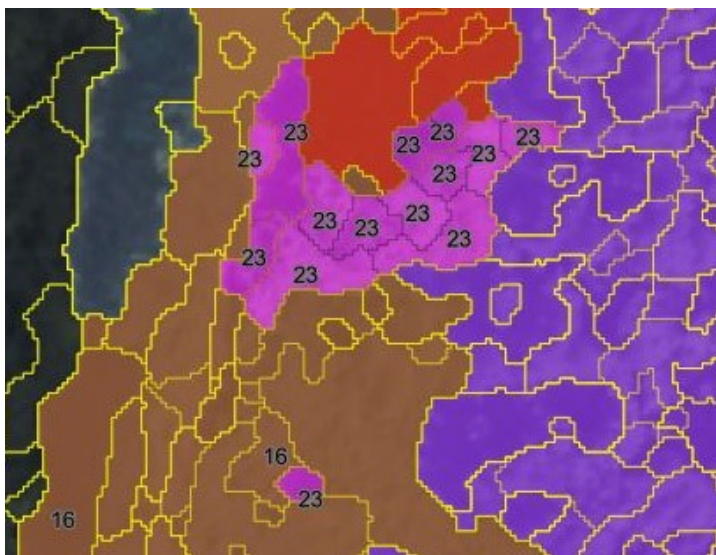
Production des annotations

La pyramide de segmentation sert de trame à la production des annotations. Il s'agit d'affecter aux segments d'images (polygones) dont on peut déterminer la nature un label figurant parmi une nomenclature donnée.

10	Feuillu indéterminé	
11	Peuplier	
12	Châtaignier	
13	Chênes décidus	
14	Chênes sempervirents	
15	Hêtre	
16	Robinier	
17	Autre feuillu pur	
20	Conifère indéterminé	
201	Conifère indéterminé aut...	
21	Douglas	
22	Mélèze	
23	Pin d'Alep	
24	Pin noir, Pin laricio	
25	Pin maritime	
26	Pin sylvestre	
27	Pin à crochets, pin cembro	
28	Sapin - épicéa	
29	Autre conifère pur autre ...	
70	Pin indéterminé	
71	Autre pin pur autre que p...	
30	Ligneux bas (arbustes)	
40	Formation herbacée natu...	
41	Formation herbacée agri...	
42	Sol minéral / chemins	
50	Bâtiment	
51	Route	
52	Eau	
53	VF	

La nomenclature utilisée dépend du contexte de la zone. Dans cet exemple, la nomenclature est adaptée à la forêt. Pour l'OCS GE, la discrimination des essences n'est pas visée, seule la distinction entre feuillus et conifères étant utile.

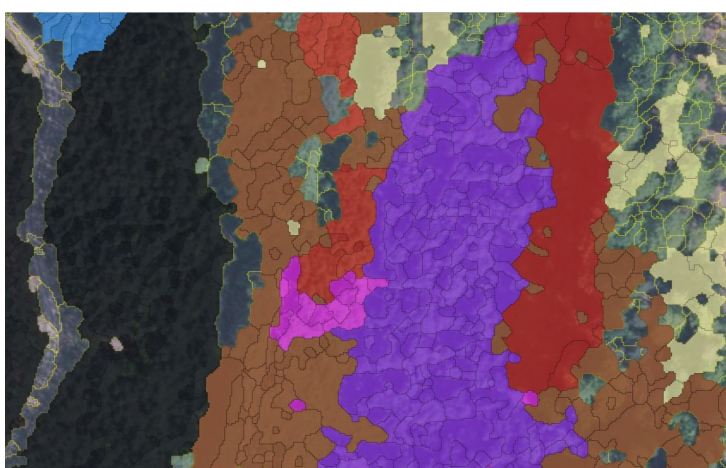
Un photo-interprète labélise un maximum de polygones en leur assignant une valeur sélectionnée dans la nomenclature, au moyen d'une IHM adaptée. Ce travail peut se faire en s'appuyant sur différents niveaux de la pyramide de segmentation.



Labélisation de polygones marqués comme robiniers (valeur 16) et pins d'Alep (valeur 23), au niveau 4 de la pyramide de segmentation.

Les autres polygones, sur lesquels aucun chiffre n'est marqué, ont été labélisés à un niveau de segmentation plus bas (pour lequel la finesse du niveau 4 de la pyramide n'était pas nécessaire).

Une fois le travail de saisie terminé, les différents niveaux de la pyramide sont fusionnés. On obtient un ensemble de segments labélisés, ne recouvrant pas nécessairement entièrement l'ortho-image.



Les segments sans label n'ont pas été identifiés par la photo-interprétation.

Classification par Machine Learning

Cette classification regroupe l'ensemble des procédés expérimentaux mis au point par le Cerema et qui apportent des données d'entraînement complémentaires, basés sur différentes méthodologies en fonction du contexte, soit en milieu urbain, soit dans les espaces NAF et faisant intervenir des techniques d'apprentissage automatique (ou *machine learning*), en utilisant des algorithmes d'arbre décisionnel (« Random Forest »).

Dans les zones urbaines est opérée une classification supervisée par pixel d'ortho-images 4 canaux (RVB+IR) superposée avec un MNH (différence entre MNT et MNS, correspondant au sursol), après apprentissage grâce à des polygones issus de la BD TOPO® croisée avec de l'ORTHO HR®. Les pixels ainsi labélisés sont croisés avec une segmentation de la BD ORTHO (algorithme LargeScaleMeanShift) aboutissant à la création des polygones dont on retient ceux classés comme étant de la végétation basse et ou de la végétation haute.



Polygones de végétation haute (vert foncé) et de végétation basse (vert clair)

Dans les zones NAF (majoritaires), une segmentation d'images SPOT superposées avec un MNH est calculée (algorithme GenericRegionMerging), puis croisée avec des données OSO, base de données d'occupation des sols produite par le Cesbio au moyen d'un algorithme Random Forest appliqué à des images Sentinel 2 (chaîne Iota²). On obtient ainsi des polygones correspondant aux classes « feuillus », « conifères », « agricole » et « naturel ».

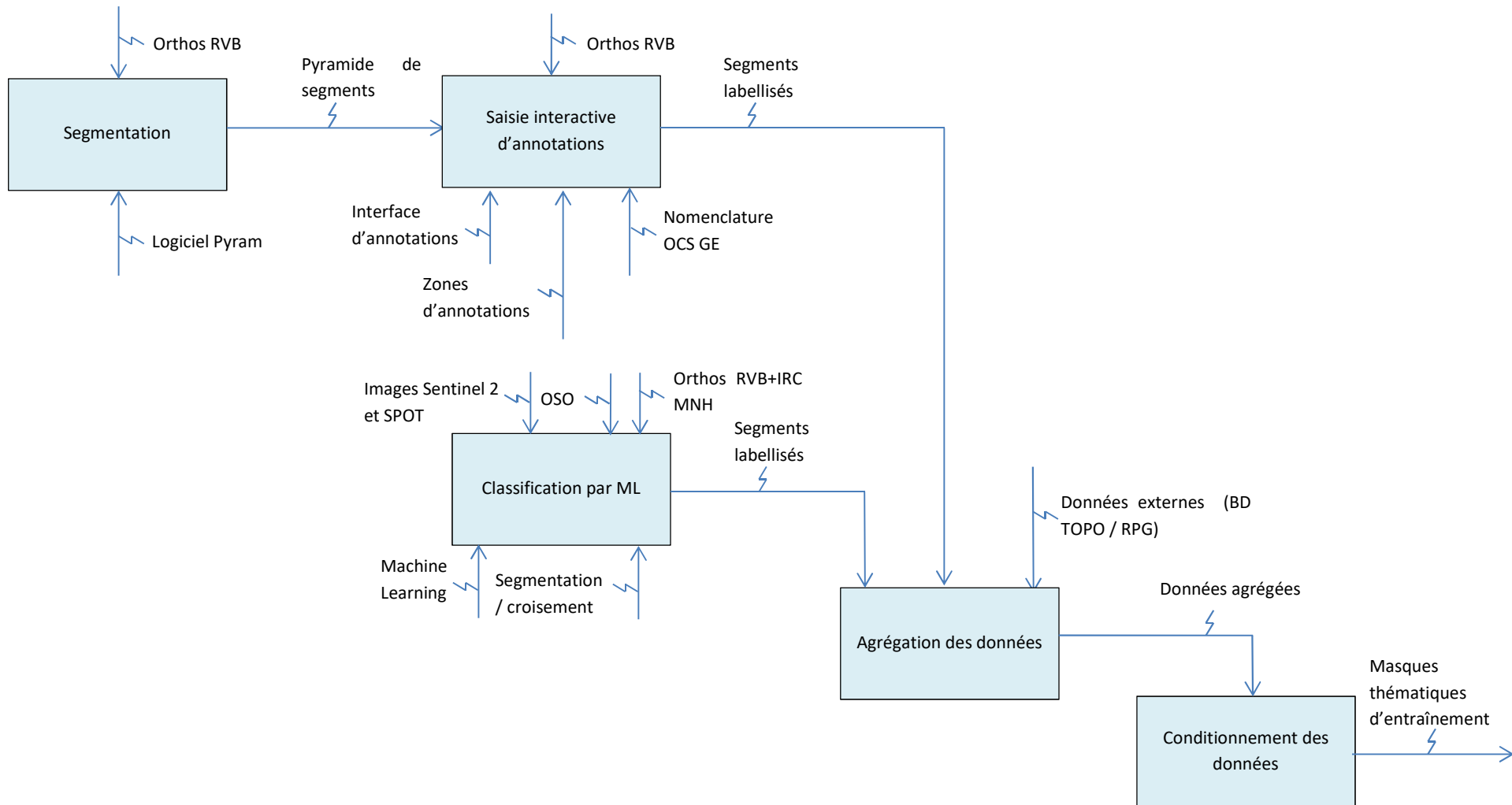


Segments (polygones orange) correspondant aux bâtiments obtenus par ML à partir d'ortho-images 20cm

Exploitation de bases de données externes

En complément, les données BD TOPO® permettent de préparer en milieu urbain des annotations pour l'ensemble des classes (à l'exception de la végétation haute et de la végétation basse). Le RPG est quant à lui utilisé dans les zones NAF pour constituer des données-vérités de type vignes.

Constitution des annotations

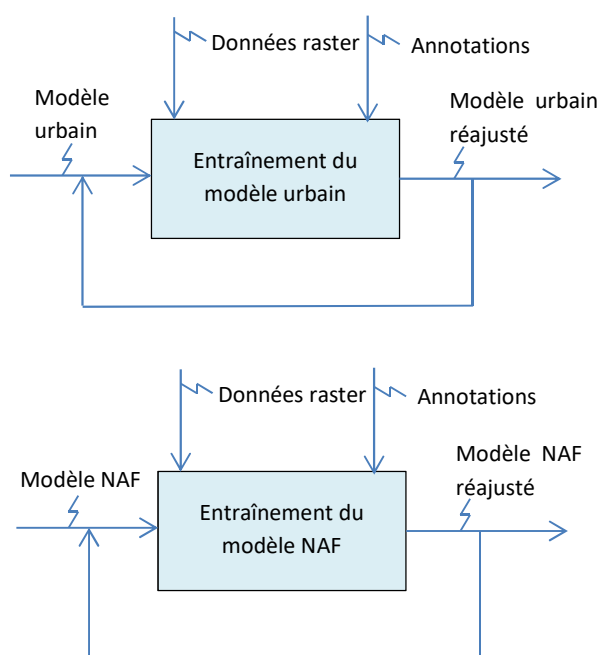


Processus de production de l'OCS GE NG

Phase d'entraînement des modèles *deep learning*

Deux modèles *deep learning* sont utilisés : le premier est adapté aux thèmes urbains ; l'autre aux thèmes rencontrés dans les zones NAF.

On soumet aux modèles les annotations constituées et des données raster : ortho-images RVB+IR et MNH (images de 512x512 pixels de résolution de 20 cm et de 5 canaux de profondeur).

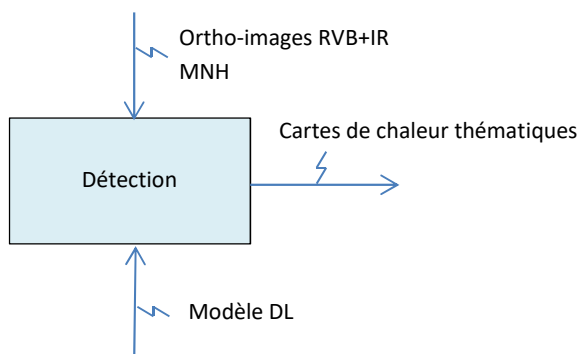


Fonctions d'entraînement du modèle urbain et du modèle NAF. Il serait possible d'instancier davantage de modèles, par exemple, un par type de zone et par classe OCS à discriminer (par exemple, un modèle dédié à la détection de ligneux en milieu urbain).

Les modèles ont vocation à être régulièrement renforcés par soumission de nouveaux jeux d'entraînement (ainsi à l'occasion d'une nouvelle campagne de production).

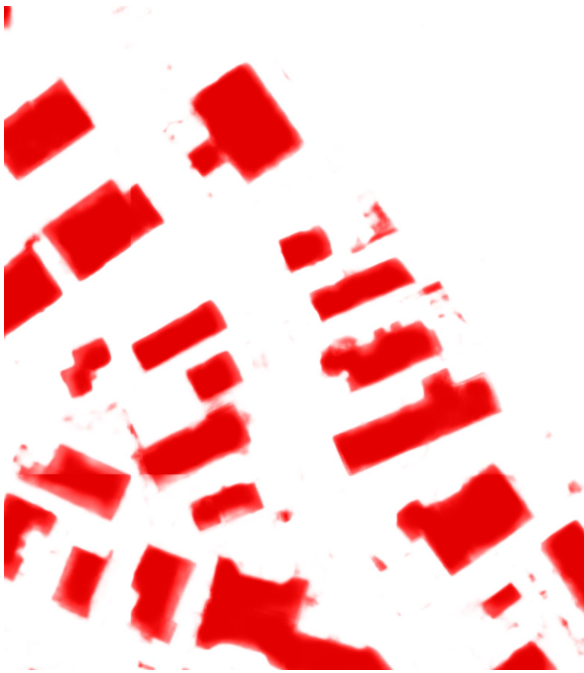
La création de davantage de modèles en fonction des paysages ou des régions pourrait être expérimentée.

Phase de détection des modèles *deep learning*



Fonction de détection des modèles

Les données produites sont des images appelées « carte de chaleur ». Une carte de chaleur est une image mono-canal (constituée par exemple de pixels possédant une valeur parmi 256 possibles). La valeur de chaque pixel correspond à la probabilité pour que ce pixel soit couvert par une classe de couverture donnée de l'OCS GE.



Sur cet exemple, plus un pixel est foncé, plus la probabilité qu'il se situe sur du bâti est élevée.

La résolution du pixel d'une carte de chaleur est de 20 cm, pour une emprise de 5x5 km.

Les cartes de chaleur pourraient faire l'objet de contrôles.

Traitement vecteur

Vectorisation

Les cartes de chaleur générées dans la phase de détection doivent d'abord être vectorisées. Le seuillage des images permet de les transformer en images binaires (les pixels ne pouvant alors prendre que 2 valeurs) : chaque pixel de valeur inférieure à un seuil est mis à 0 ; les autres sont mis à 1. La valeur du seuil est fixée interactivement et empiriquement en comparant l'effet du seuillage sur quelques cartes de chaleur à des données-vérité.

Ce seuillage pourrait être déterminé automatiquement, soit en l'intégrant à la phase de détection, soit par l'analyse des cartes de chaleur d'autres thèmes et de données-vérité.

Carte de chaleur (valeurs de pixels comprises entre 0 et 255)

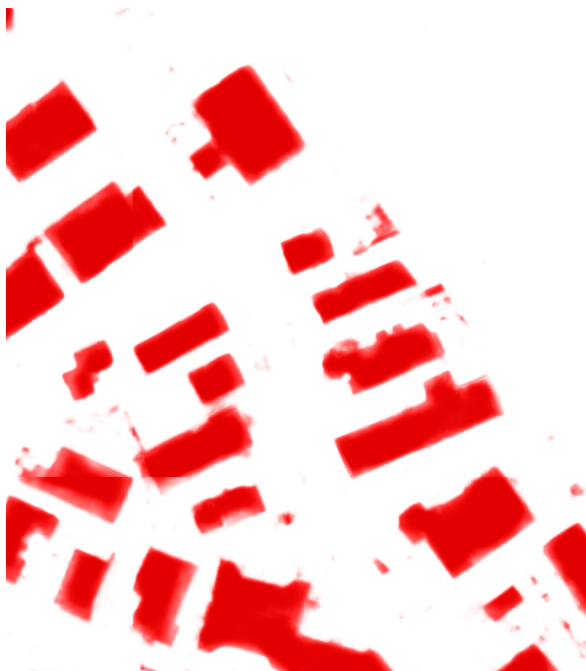


Image binaire obtenue avec un seuil de 127



Dans les zones NAF, ces images sont simplifiées (au moyen d'un programme fourni par le Cerema) dans le but de rendre les données vecteur plus facilement manipulables.

La vectorisation s'appuie sur des algorithmes également mis au point par le Cerema. On dispose alors d'un fichier de type shapefile par classe et par dalle (5x5 km). L'ensemble est importé dans une base de données (une table par couche).

Image binaire



Polygones obtenus par vectorisation



Préparation des données

Outre les résultats de détection par *deep learning*, les sources de données utilisées sont :

- des extractions datées de la BD TOPO® : réseaux routier et ferré, zones d'activités, vignes ;
- les parcelles agricoles issues du RPG millésimé ;
- le référentiel géographique forestier (BD Forêt®) ;
- les fichiers fonciers : fichiers de la DGFIP enrichis par le Cerema (consistant notamment à agréger des parcelles afin de constituer des unités parcellaires).

Constitution d'une ossature de type ORFHB

La production de l'OCS GE démarre par la production d'une ossature, structure constituée à partir de données vecteur externes, généralement les réseaux principaux (routier et ferré) auxquels on applique une « bufferisation » (dilatation). Elle est censée représenter une trame de l'OCS relativement stable dans le temps. Toutefois, des travaux structurants, telle qu'une construction d'autoroute, peuvent parfois la changer significativement. Elle est complétée et précisée par d'autres données pour aboutir à un produit intermédiaire : l'**ORHFB (Ossature-Réseaux-Forêt-Hydro-Bâti)**.

Création de l'ossature



Ossature créée par dilatation des tronçons du réseau principal de la BD TOPO®.

Après intégration du réseau principal, on insère le réseau secondaire (routier et ferré) dilaté. En affectant aux objets les attributs de l'OCS GE, on obtient une structure nommée **ORHF** (Ossature Réseau/Hydro/Ferré).



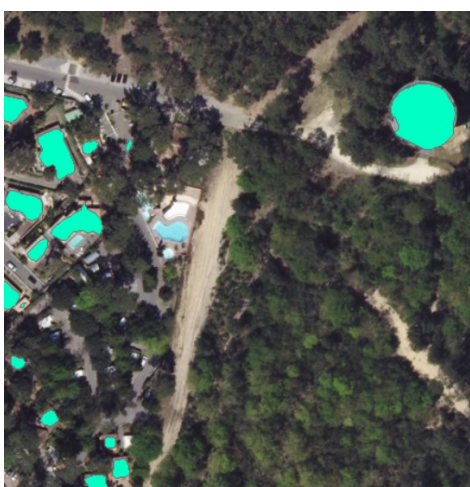
On ajoute les portions en rouge aux bâtis proches des routes.

Sélection du bâti

Puis, le bâti de la BD TOPO®, dont l'actualité ne correspond pas nécessairement aux dates des prises de vues, et donc au millésime du produit, est mis à jour par croisement avec les masques de bâti issu du *deep learning*. Les bâtiments inférieurs à 20m² sont filtrés (conformément aux spécifications de l'OCS GE).



Sur cet exemple, l'objet situé à droite (château d'eau) est absent de la BD TOPO®.



Le masque du bâti issu du *deep learning* comporte en revanche cet objet. D'autres bâtiments, masqués par les arbres, n'ont pas été détectés.



Le bâti sélectionné est représenté en jaune, avec la géométrie de la BD TOPO® si elle existe, sinon avec celle dérivée du *deep learning*. Les bâtiments trop petits, ou situés sous les arbres, représentés en rouge, ont été éliminés.

Généralisation du bâti

Les bâtiments distants de moins de 10 m les uns des autres sont alors agrégés ; ceux de moins de 200m² sont « gonflés » (selon les spécifications de l'OCS GE).



Les polygones bleus représentent le résultat de la généralisation.

On peut déduire des éléments obtenus (ossature et bâti généralisé) une délimitation entre la zone urbaine et la zone NAF, délimitation utile dans la suite du processus.

Zone urbaine →

Zone NAF →



Sélection de l'hydrographie

De même que pour le bâti, les éléments hydrographiques de la BD TOPO® sont croisés avec les masques issus du *deep learning*. On filtre les objets de petite taille.

Mise en cohérence des zones bâties avec les réseaux

Afin de respecter les spécifications de l'OCS GE, les zones bâties sont mises en cohérence avec les données précédemment calculés (Ossature, complément routier et ferré, Hydrographie et Bâti). La couche **ORHFB** est alors constituée.

Intégration du bitume

Les masques de bitume obtenus par *deep learning* sont rajoutés aux endroits où il fait défaut dans la BD TOPO®.



Sont représentés en mauve les éléments de l'ossature dérivés des réseaux BD TOPO®, auxquels on a adjoint les parties en rouge, calculées par *deep learning*. On obtient le produit **ORHFB Merge**.



Autre aperçu du produit OFRHB Merge après insertion d'autres objets tels que les terrains de sport.

A ce stade, il est possible de calculer le périmètre urbain plus finement.

Constitution de l'OCS GE

Couche usage

L'analyse des fichiers fonciers combinée avec des données BD TOPO® permet de créer une couche d'usage OCS GE en émettant une hypothèse sur l'usage des unités parcellaires (primaire, secondaire, etc.)

Cet algorithme s'appuie sur un traitement développé par le Cerema.

Les unités parcellaires dont les usages sont similaires sont alors agrégées.

Unités parcellaires issus des fichiers fonciers



Unités parcellaires d'usage similaire agrégées



Une couche d'usage est également créée avec les objets BD TOPO®



La fusion de l'ensemble permet d'obtenir une couche d'usage plus complète.



Couche couverture

L'ORHFB précédemment calculée est très morcelée. Il est nécessaire de la remplir afin d'obtenir une couche couverture qui recouvre totalement le territoire, grâce aux masques thématiques fournis par le *deep learning*.



En zone urbaine (délimitée par le trait bleu), les masques du ligneux (vert clair) contribuent au remplissage de la couche couverture.



On finit par obtenir, en zone urbaine, une couverture représentée ici avec les couleurs conventionnelles de l'OCS GE.

De même, l'ORHFB est remplie en zone NAF avec des masques thématiques calculés par *deep learning*, mais aussi par les données de la BD Forêt® et du RPG.

D'ultimes corrections géométriques sont apportées aux objets pour mieux se conformer aux spécifications.

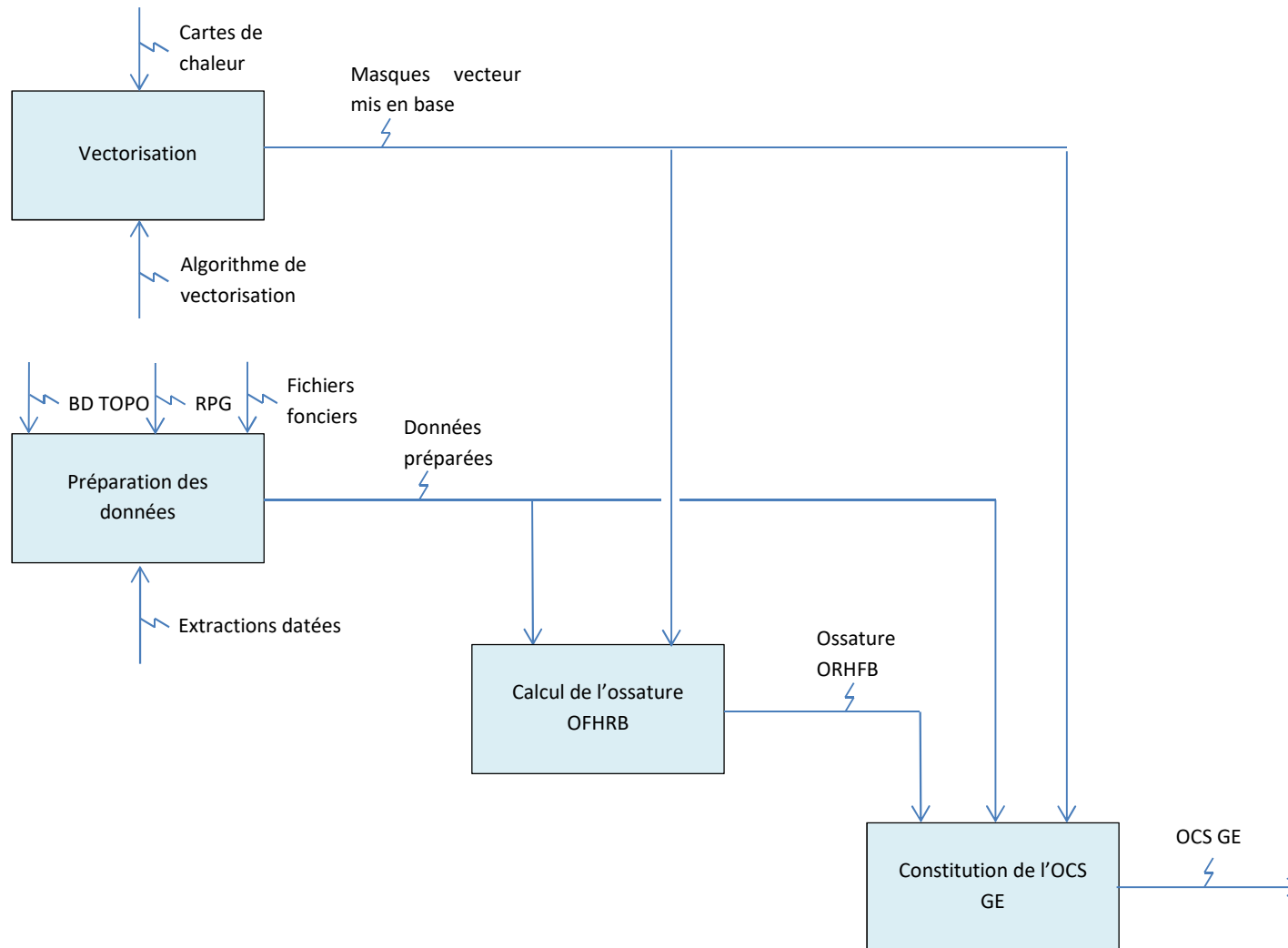
Avant corrections



Après corrections



Traitement vecteur



Processus de production de l'OCS GE NG

Processus de mise à jour

Synoptique du processus

Principes

Le processus de mise à jour doit garantir un certain niveau de stabilité des objets afin de permettre de comparer deux millésimes et d'en dériver des indicateurs pertinents (tels que l'évolution de l'artificialisation des sols). L'idée générale n'est pas de calculer complètement une OCS GE pour chaque millésime, mais de détecter les changements entre deux millésimes avant de mettre à jour l'OCS GE localement uniquement sur les zones concernées.

La zone à produire est découpée en dalles ou « vignettes » carrées de 64 m de côté. La phase de détection automatique doit repérer les vignettes contenant du changement. Les alertes sont donc générées « à la vignette ». Elles sont ensuite exploitées par des méthodes classiques afin de mettre à jour l'OCS GE localement. Il faut noter que le changement est un phénomène diffus (par exemple, seulement 0.35% de la surface de la région Pays de la Loire a changé en moyenne par année sur la période 2013-2016).

A terme, la saisie du changement dans les zones de changement pourrait également être automatisée.

La partie automatisée du processus devrait permettre de détecter au moins 95% des zones de changements, tout en minimisant la sur-détection de changements. Le but est de s'affranchir d'un parcours exhaustif du territoire par des photo-interprètes.

Deux approches ont été combinées pour la mise en œuvre de techniques de *deep learning* pour la détection de changement. La première est supervisée, comme pour la détection multithématique en constitution initiale, donc s'appuyant sur de grandes quantités d'annotations. La seconde est dite « non-supervisée ».

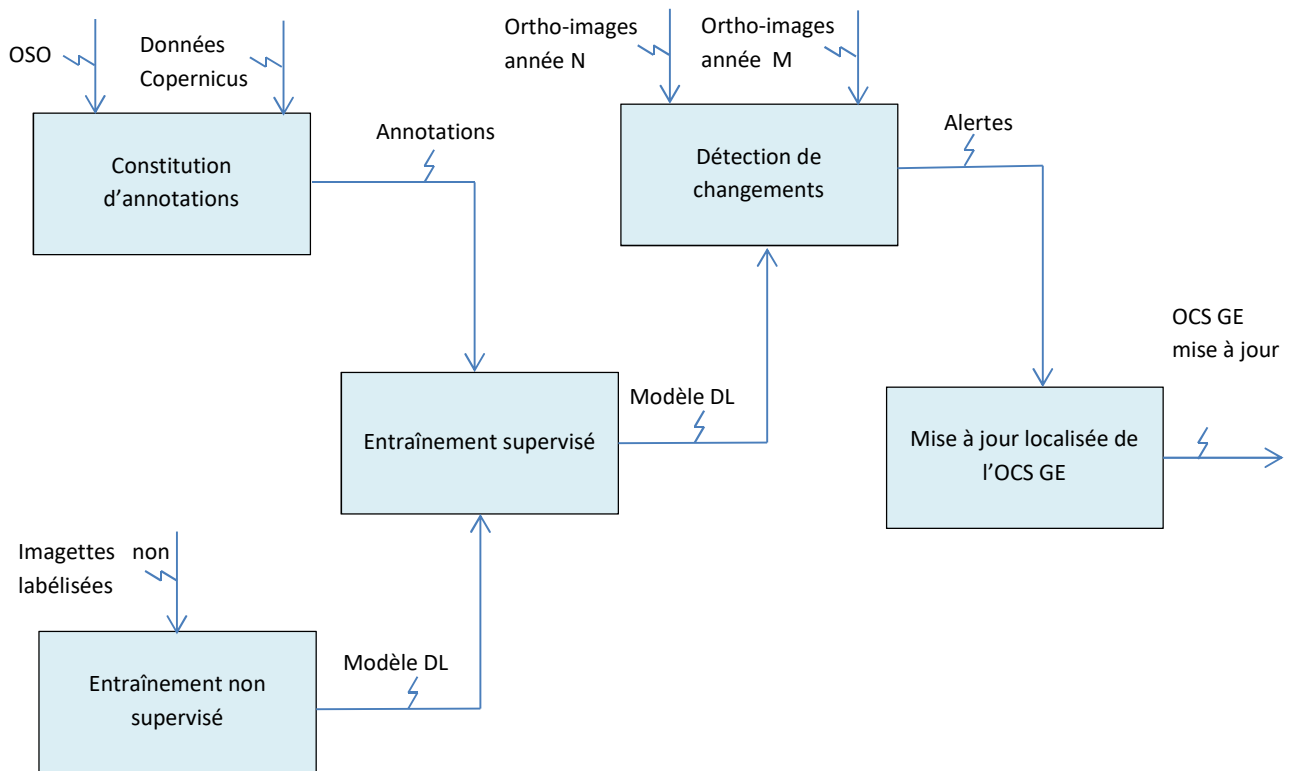
Etapes principales

Les grandes étapes du processus sont donc, quelle que soit l'approche utilisée pour le *deep learning* :

- la constitution des annotations ;
- la phase d'entraînement non-supervisé des modèles *deep learning* ;
- la phase d'entraînement supervisé des modèles *deep learning* ;
- la mise à jour localisée de l'OCS GE.

Les parties automatisées sont encore au stade expérimental.

Mise à jour



Préparation d'annotations

Les données d'OCS GE existantes (déjà produites via un processus classique) peuvent constituer une source de données intéressante. Toutefois, elles ne recouvrent qu'environ 7% du territoire en version bi-millésimées, non représentative du territoire, avec des dynamiques de changement assez faible. On risquerait donc se heurter à des problèmes de généralisation si on se contentait d'exploiter ces seules données. On cherche donc à les compléter par d'autres données.

On s'appuie donc sur des couches produites par le programme européen Copernicus :

- Corine Land Cover ;
- Urban Atlas ;
- High Resolution Layer (HRL).

En les croisant avec des données historiques de la BD TOPO®, on arrive à construire des polygones censés recouvrir des zones de changement, là où ces différentes sources indépendantes convergent. *Cette base vérité contient inévitablement du bruit. Il faut donc la fiabiliser manuellement.*

Par ailleurs, on peut également constituer des masques de « non-changement », concrètement des paires d'images sur lesquelles il n'y a pas de changement ; pour cela, on exploite les cartes d'occupation du sol OSO du Cesbio.

Entraînement non supervisé

Dans de ce type d'approche, le modèle n'utilise pas d'annotations pour s'entraîner, au contraire de l'apprentissage supervisé auxquels on soumet des vignettes labélisées, c'est-à-dire connues pour contenir ou non du changement.

Ici, on soumet au modèle un nombre important de paires d'images d'une même zone, mais de différentes dates, sans spécifier si elles sont censées contenir des changements.

Cela permet d'obtenir une version pré-entraînée des modèles *deep learning* et pallier la faiblesse des annotations.

Détection de changements

Trois canaux d'ortho-images de chaque millésime de 50 cm de résolution sont exploités : infra-rouge, rouge et vert.

Elles sont soumises à un réseau de convolution « siamois ». L'encodage obtenu est aplati (fonction « flat ») lui-même soumis à un réseau de type perceptron multi-couches (PMC).

Le résultat est une classification binaire « changement / non changement » des vignettes partitionnant le territoire. Ces images doivent donc être interprétées comme des alertes indiquant potentiellement des changements.

Mise à jour de l'OCS GE

A ce stade, on dispose d'un certain nombre de vignettes réputées marquer un changement.

Disposant de l'OCS GE d'un millésime N, on peut le mettre à jour pour le millésime M suivant deux stratégies distinctes à étudier.

La première stratégie consiste :

1. à calculer le millésime M de façon complète selon le processus de constitution initiale ;
2. à « trouser » le millésime N à l'endroit des vignettes de changement ;
3. à remplir les trous ainsi faits par de la donnée du millésime M ;
4. à contrôler et à corriger l'ensemble de ces remplacements.

La deuxième stratégie consiste :

1. à initialiser l'OCS GE du millésime M en copiant celle du millésime N ;
2. à faire saisir la photo-interprétation systématiquement sur les vignettes.