



# Analyse automatique du couvert végétal pour la gestion du risque végétation en milieu ferroviaire à partir d'imagerie aérienne

■ Hélène ROUILLON

*La végétation et les risques qu'elle peut comporter pour les infrastructures ferroviaires et leurs usagers constituent un enjeu majeur pour SNCF<sup>1</sup> Réseau. Il est donc nécessaire pour l'entreprise ferroviaire de connaître et maîtriser ce risque végétation. L'objet de ce projet de fin d'études (PFE) est l'analyse automatique du couvert végétal en milieu ferroviaire à partir d'imagerie aérienne RVB<sup>2</sup> acquise par drone, hélicoptère ou avion. Pour répondre à cette problématique, un réseau de neurones destiné à la segmentation sémantique des images est mis en œuvre. En effet, une fois bien entraînés, les réseaux de neurones, par leur capacité d'apprentissage, sont en mesure de classifier efficacement toute nouvelle image. Trois classes ont été définies en fonction des risques que pouvaient présenter la végétation : "arbres", "reste de la végétation" et "non-végétation". Une chaîne de traitement complète des données a été proposée permettant, sur la base de ces images, une cartographie SIG<sup>3</sup> de la végétation. Cette connaissance, aisément déployable sur des lignes entières, doit permettre au mainteneur de cibler les zones prioritaires et d'optimiser ses plans d'élagages.*

## MOTS-CLÉS

Couvert végétal, réseaux de neurones, segmentation sémantique, imagerie aérienne, risque végétation, cartographie

exploiter afin d'estimer le risque végétation.

Pour cela, il faut disposer de trois éléments : la cartographie des infrastructures, la cartographie de la végétation et la définition des modèles de danger indiquant à partir de quel moment on considère qu'il y a un risque. SNCF Réseau possède déjà une cartographie de l'infrastructure ferroviaire (caténaires, voies, etc.) dans un SIG sur l'ensemble du territoire. De plus, la définition exacte des différents types de "dangers végétation" est établie par SNCF Réseau. En somme, le dernier élément manquant pour l'évaluation du risque végétation est la position et l'étendue de cette dernière, qui devra être consignée sous forme de carte. C'est pour cette raison que nous avons choisi d'avoir recours à des orthophotographies, car elles sont corrigées géométriquement et celles de l'entreprise sont géoréférencées. Elles deviennent ainsi un véritable document cartographique sur lequel il est possible de localiser des éléments par leurs coordonnées. Elles peuvent être utilisées dans un SIG.

Le travail que j'ai réalisé a donc pour objectif la mise en place d'une solution permettant de connaître la position et l'étendue de la végétation de manière générale dans des orthophotographies RVB, avec un intérêt plus particulier pour les arbres dans cette étendue de végétation, car ils peuvent présenter un risque direct. Ainsi, le rendu final envisagé consiste en une ou plusieurs couches vectorielles relatives aux classes de

<sup>1</sup> Société nationale des chemins de fer français.

<sup>2</sup> Rouge, vert, bleu.

<sup>3</sup> Système d'information géographique.

## Introduction

Altametrus est une filiale de SNCF Réseau. L'entreprise collecte des données à des fins de topographie, surveillance, inspection, etc. Elle réalise aussi leur traitement pour ses clients partenaires dans un but d'optimisation de la gestion de leurs actifs.

Le suivi et la maîtrise de la végétation constituent le deuxième poste de dépenses pour SNCF Réseau, soit 120 millions d'euros chaque année. Cette tâche de gestion du "risque végétation" doit permettre de limiter le risque d'incidents (chutes d'arbres, patinage de train, incendies, etc.) liés à la présence de végétation à proximité du réseau ferré national (RFN). De nombreux enjeux sont liés à une gestion adéquate du couvert végétal : la sécurité, l'état des installations, l'écologie, des enjeux financiers et des enjeux sociaux et sociétaux.

Aujourd'hui, la principale méthode utilisée pour connaître le risque végétation aux abords des voies ferrées consiste en des tournées pédestres de reconnaissance. Cette méthode est fiable. Toutefois, elle est chronophage et donc coûteuse pour SNCF Réseau. Ainsi, SNCF Réseau et Altametrus cherchent aujourd'hui à développer une nouvelle méthode d'analyse du risque végétation. Différents tests ont déjà été effectués. En 2015 notamment, un PFE portant sur le recours au drone pour la surveillance du RFN, c'est-à-dire des infrastructures ferroviaires uniquement, a révélé les atouts du drone notamment en matière de coût, de qualité des résultats et de couverture du réseau. Par ailleurs, Altametrus est amenée à générer un très grand nombre d'orthophotographies RVB dans le cadre de projets portant entre autres sur la maintenance du RFN. Ce sont ainsi des données opportunes. L'entreprise souhaiterait maintenant les





végétation. À terme, la solution développée devra permettre d'améliorer la connaissance du risque végétation pour optimiser la maintenance de la végétation sur le terrain et ainsi réduire les coûts dans une perspective de gestion efficace des actifs.

L'étude de l'état de l'art nous a permis de prendre connaissance des différentes approches de classification d'objets, et plus particulièrement de la végétation, dans des images aériennes RVB. Cette étude nous a amenés à choisir une approche basée sur les réseaux de neurones, car présentant les résultats les plus convaincants pour une automatisation et généralisation du processus de classification d'images.

## La méthode utilisée : un réseau de neurones

### ■ Pertinence des réseaux de neurones à convolution (CNN<sup>4</sup>) : généralités

Avant l'apparition des réseaux de neurones à convolution, les caractéristiques des images traitées étaient extraites manuellement. Les réseaux comportent eux des couches spécifiques qui vont extraire et apprendre ces caractéristiques automatiquement. Ensuite a lieu une phase de reconnaissance à partir de ces caractéristiques [Heudin, 2016]. Ainsi, lorsque les réseaux de neurones sont bien entraînés, ils sont en mesure de généraliser, c'est-à-dire d'effectuer correctement la tâche qu'on leur confie sur de nouvelles images, qui n'ont pas été vues pendant l'entraînement et ce, de façon automatique. Ils sont donc très intéressants lorsque l'on souhaite automatiser un processus pour l'appliquer sur un grand nombre de données.

Toutefois, cette méthode comporte tout de même quelques inconvénients. Le principal défaut réside dans la quantité de données nécessaire pour entraîner le réseau de neurones. La base de données doit en effet être suffisamment grande et diversifiée (et représentative des futures données à traiter) pour que le réseau soit en mesure de générali-

<sup>4</sup> Convolutional Neural Network qui signifie réseau de neurones à convolution.

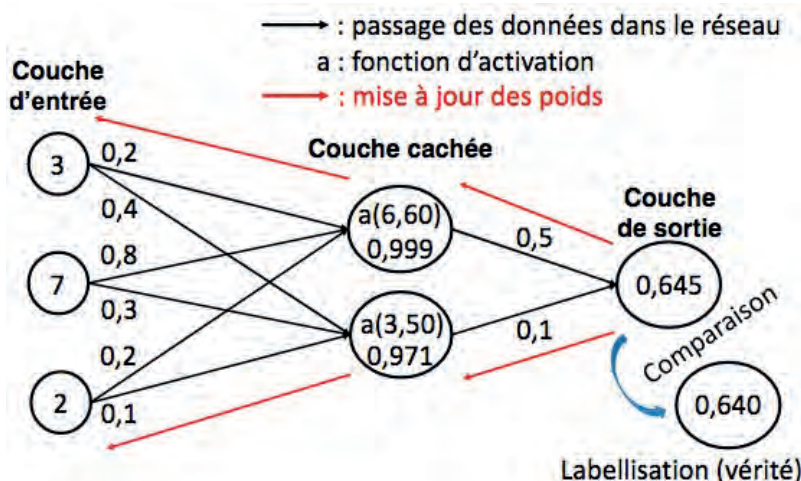


Figure 1. Représentation schématique simplifiée de la structure et du fonctionnement d'un réseau de neurones.

ser. Par ailleurs, si l'on souhaite coder entièrement le réseau de neurones, il est nécessaire d'avoir de bonnes connaissances en programmation informatique. Enfin, l'entraînement du réseau de neurones est une partie délicate qui demande de l'expérience.

Un réseau de neurones est composé de plusieurs couches, elles-mêmes constituées de plusieurs neurones. Les neurones d'une couche à l'autre sont reliés par des connexions auxquelles sont associés des poids. Ceux-ci sont les paramètres du réseau. L'entraînement du réseau va consister à ajuster ces poids pour qu'il apprenne à résoudre la problématique posée. Pour faire cet entraînement, un grand nombre d'images préalablement annotées (considérées comme vérité terrain) vont lui être données et passées à travers chaque couche. Le réseau va effectuer une prédiction pour chaque image et cette prédiction sera comparée à la labellisation. La différence entre labellisation et prédiction, appelée perte, est calculée. Puis le réseau va ajuster les poids de ses connexions de façon à diminuer cette perte. Le lecteur pourra se rapporter à la figure 1 pour une meilleure compréhension. L'ensemble du jeu de données est ainsi passé plusieurs fois dans le réseau et les poids sont donc modifiés de façon itérative, en fonction de l'évolution de la perte. Une fois l'entraînement terminé (lorsque la perte a été minimisée), le réseau devrait être capable de faire de bonnes prédictions pour les nouvelles images qui lui seront présentées.

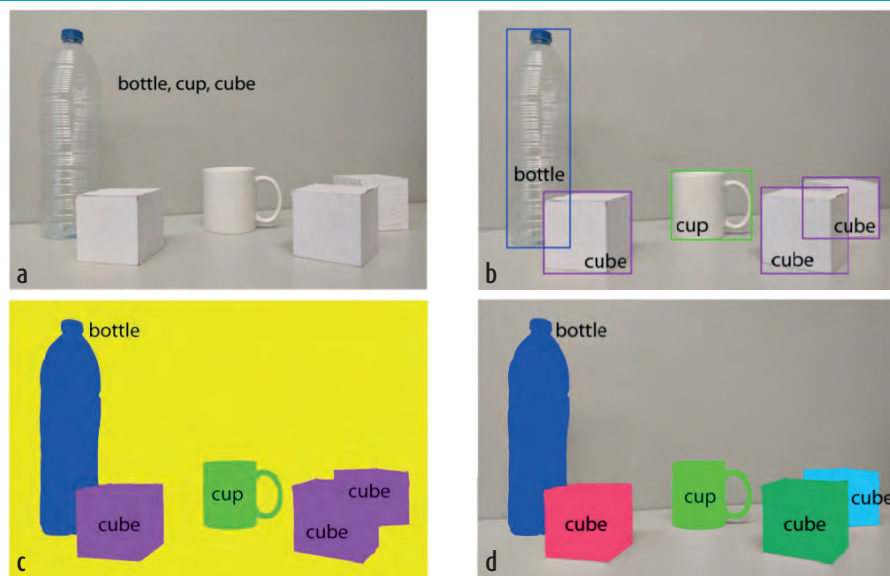
### ■ L'architecture de réseau de neurones à convolution retenue

Le réseau de neurones de base utilisé pour le traitement d'images est appelé réseau de neurones à convolution, en référence aux opérations de convolution qu'il effectue pour extraire automatiquement les caractéristiques. Il permet d'indiquer les classes contenues dans une image. Mais la grande famille des réseaux de neurones à convolution comporte plusieurs types de réseaux neuronaux, destinés à différentes tâches comme indiqué en figure 2. Dans notre cas, nous ne souhaitons pas seulement connaître les classes présentes dans l'image, mais nous voulons indiquer à quel endroit se trouvent les classes dans l'image et quelles étendues elles occupent, de façon à avoir une représentation du couvert végétal. Cela revient à faire (dans le cadre des réseaux de neurones) de la segmentation sémantique, c'est-à-dire à attribuer une classe à chaque pixel. Chaque classe correspondant à une couleur, en sortie l'image obtenue comporte des "masques colorés" associés aux différentes classes. Différentes architectures de réseaux de neurones à segmentation sémantique existent. L'état de l'art a montré que le réseau de neurones SegNet convenait pour notre problématique.

### La mise en place de la solution : le réseau de neurones SegNet

#### ■ Le choix du code informatique

Étant donné les limites de la méthode des réseaux de neurones précédemment



**Figure 2.** Quatre types d'opérations de traitement d'images, en a) Classification, b) Localisation d'objet (ou détection), c) Segmentation sémantique, d) Segmentation d'instance, [Garcia-Garcia et al., 2018].

évoquées, nous avons d'abord vérifié sur Internet s'il n'existait pas un projet ayant vocation à réaliser une segmentation sémantique d'images aériennes. Les recherches ont abouti sur le code informatique de [Audebert et al., 2017], qui recourt à une architecture SegNet pour la segmentation sémantique d'images d'observation de la Terre.

Plus précisément, dans son cas d'usage initial, le réseau de neurones proposé permet de classer des orthophotographies IRRV<sup>5</sup> de résolution 9 cm et prises dans la ville de Vaihingen (Allemagne). Plusieurs classes intéressent les auteurs de la recherche, dont deux relatives à la végétation (végétation basse et arbres). Bien que les images utilisées dans notre cas se trouvent dans l'espace colorimétrique RVB, l'évaluation qualitative des résultats obtenus pour ces classes a été l'occasion d'avoir une première idée des résultats qui pouvaient être obtenus pour la végétation avec ce réseau. De plus, nous avons sélectionné ce code car la résolution des orthophotographies utilisées par les auteurs est proche de celle des données à disposition pour entraîner le réseau (généralement de 2 à 4 cm). Enfin, le code proposait déjà un réglage de différents éléments du réseau, ce qui permettait d'avoir à la fois une première idée des résultats que l'on pouvait obtenir avec ces réglages et un ordre de grandeur des réglages à effectuer dans la perspective de leur modification pour entraîner notre réseau.

<sup>5</sup> Infrarouge, Rouge, Vert.

### ■ La création de la base de données

La création de la base de données commence par la sélection d'images (en l'occurrence d'orthophotographies RVB) à labelliser. Nous avons sélectionné des images de résolution 2 cm et dimension 5000 x 5000 pixels, comportant des types de végétation (arbres, arbustes, buissons, arbrisseaux, etc.) variés et pris à des moments de l'année divers (végétation verte, jaune-orangé, marron, sans feuilles). Par ailleurs, pour évaluer si la quantité de données est suffisante, il est possible de procéder de façon empirique. En outre, nous invitons le lecteur à se reporter à l'article de [Castillo-Navarro et al., 2019] pour plus d'informations relatives à l'évaluation de la quantité de données nécessaires pour l'entraînement des réseaux de neurones à segmentation sémantique pour le traitement d'images d'observation de la Terre.

Après avoir sélectionné ces images, nous sommes passés à l'étape de labellisation. Il a d'abord fallu déterminer le nombre de classes que nous voulions retenir pour la problématique posée. Conformément au besoin exprimé par l'entreprise, la classification suivante a été choisie :

- la classe "arbres" : elle comporte tous les arbres ainsi que les arbustes ;
- la classe "reste de la végétation" : elle contient les arbrisseaux, buissons, ronciers, haies, plantes herbacées qui ne sont pas au ras du sol et qui sont parfois envahissantes, etc. ;



**Figure 3.** Extrait de la labellisation d'une image, avec les traits réalisés manuellement et le résultat de la labellisation après application de l'algorithme des bassins versants à l'ensemble de l'image, avec en vert foncé la classe "arbres", en vert anis la classe "reste de la végétation", en gris la classe "non-végétation".

- la classe "non-végétation" : elle comprend tous les éléments qui ne sont pas de la végétation (rails, bâtiments, etc.) et de l'herbe rase, sèche ou clairsemée, qui ne présente aucun risque.

L'objectif de cette classification est de pouvoir connaître la position et l'étendue des arbres tout en ayant une connaissance générale de la position et de l'étendue des autres types de végétation qui pourraient présenter un risque, même moins important.

Ensuite, nous avons recherché un outil qui nous permette de réaliser une labellisation efficace, c'est-à-dire de qualité et rapide. Nous avons sélectionné l'outil *Pixel Annotation Tool* [Breheret, 2017]. En effet, il ne nécessite pas de labelliser toute l'image manuellement puisque l'opérateur est aidé par l'algorithme des bassins versants<sup>6</sup>, comme indiqué sur la figure 3.

La base de données totale (pour l'entraînement et l'évaluation du réseau) contient 30 images de dimension 5000 x 5000 pixels et résolution 2 cm, soit 361 images Full HD<sup>7</sup>. Cette base de données est divisée en trois sous-jeux :

- le jeu d'entraînement : 80 % du jeu total,

<sup>6</sup> Cet algorithme se base sur la visualisation de l'image (en niveaux de gris) comme s'il s'agissait d'un paysage comportant des reliefs. À chaque point de l'image est associée une intensité du pixel correspondant. Cette dernière fournit l'altitude du point. Plus l'intensité du pixel est élevée, plus le point correspondant est haut (sommets). Pour chaque minimum local (vallée), l'algorithme cherche à quel bassin versant ce minimum appartient. Ainsi, la position des lignes de crêtes, frontières entre deux classes, se dessine progressivement. On obtient finalement une image segmentée [INRA].



- comme l'indique son nom, il permet d'entraîner le réseau de neurones ;
- le jeu de validation : 10 % du jeu total, il est utile pour pré-évaluer les performances du réseau sur un jeu de données non utilisé pendant l'entraînement. Il doit permettre d'affiner les hyperparamètres du réseau (réglages autres que les paramètres, c'est-à-dire autres que les poids des connexions) et ainsi de sélectionner un modèle final ;
  - le jeu de test : 10 % du jeu total, il est employé pour évaluer les performances du réseau retenu sur un jeu de données jusqu'alors jamais vu par le réseau.

### ■ Évaluation des résultats et bilan des erreurs théorique

L'évaluation des résultats de segmentation sémantique peut à la fois être qualitative et quantitative. L'évaluation qualitative s'effectue en confrontant l'image segmentée à l'image labellisée et/ou l'orthophotographie. L'évaluation quantitative se fait grâce à des métriques d'évaluation qui vont permettre de mesurer la justesse des prédictions d'un modèle entraîné. En pratique, cette évaluation est réalisée sur le jeu de test.

Plusieurs notions sont importantes et vont être présentées dans ce qui suit à travers le cas d'une classification binaire. On va donc considérer que la classe positive sera la classe "végétation" et la classe négative "l'ensemble des autres classes". L'exactitude correspond, pour une classe donnée, au pourcentage de pixels bien classifiés par rapport à la labellisation. Le rappel permet pour une classe donnée d'avoir une idée de la quantité de faux négatifs (oublis), tandis que la précision permet d'avoir une idée de la quantité de faux positifs (excès). Le rappel et la précision prennent une valeur comprise entre 0 et 1. Lorsqu'ils valent 1 pour une classe donnée, cela signifie respectivement qu'il n'y a pas de faux négatif et pas de faux positif. Une illustration de ces notions dans le cas d'une détection est fournie en *figure 4*. Enfin, le score F1 correspond à la moyenne harmonique de la précision

<sup>7</sup> Correspond aux dimensions 1920 x 1080 pixels.



**Figure 4.** Pour une détection de chat (classe positive) dans une image, le chat encadré est une détection exacte (vrai positif), le chat oublié (non encadré) est un faux négatif (oubli), le chien détecté est un faux positif (excès), les autres animaux non-détectés sont exacts (vrais négatifs), [Saniterpen, 2018], modifié.

et du rappel<sup>8</sup> ; sa valeur est également comprise entre 0 et 1.

Par ailleurs, il paraît important de faire un bilan des erreurs théorique afin de pouvoir essayer d'expliquer les résultats obtenus. À ce sujet, il faut noter que lorsqu'il utilise une approche neuronale pour faire une segmentation sémantique, le concepteur maîtrise : la base de données, l'architecture du réseau, le réglage de plusieurs éléments. Une fois l'entraînement lancé, l'opérateur ne maîtrise pas les calculs intermédiaires faits au sein des couches par le réseau. Il devient donc plus complexe d'expliquer les résultats obtenus.

Néanmoins, nous pouvons dégager plusieurs motifs pouvant expliquer des résultats de segmentation sémantique décevants :

- une quantité de données d'entraînement insuffisante ;
- une trop grande variabilité au sein des données d'entraînement au regard de la quantité de données de chaque type (arbres, arbrisseaux, haies, etc.) ;
- une labellisation insuffisamment précise ;
- une labellisation insuffisamment exacte ;
- une erreur de distinction entre les classes par l'opérateur et donc erreur de labellisation ;

- un oubli d'éléments de la classe "arbres" ou "reste de la végétation" qui seraient donc injustement associés à la classe "non-végétation" ;
- un taux d'apprentissage trop élevé et/ou un nombre d'époques d'apprentissage trop important qui conduirait à un surapprentissage ;
- un contexte insuffisant ne permettant pas au réseau de discerner correctement les éléments d'une scène, en d'autres termes, une taille de tuile trop faible ;
- une résolution d'image trop faible pour laisser apparaître suffisamment de caractéristiques permettant de distinguer les différentes classes.

Cette liste n'est pas exhaustive. Il est possible que d'autres erreurs pouvant conduire à des résultats peu convaincants existent. Celles qui nous paraissent être les plus évidentes et courantes au regard des connaissances acquises au cours de ce travail sont citées ci-dessus.

### ■ Entraînement du réseau et choix du réseau final

Plusieurs paramètres<sup>9</sup> du réseau de neurones peuvent être modifiés durant la phase d'entraînement pour tenter d'améliorer les performances du réseau. Parmi eux, nous allons expliciter dans

<sup>9</sup> Attention nous ne parlons pas ici des poids du réseau, le terme "paramètres" est ici à prendre au sens commun du terme.

<sup>8</sup>  $Score\ F1 = 2 \times \frac{précision \times rappel}{précision + rappel}$



ce qui suit les paramètres suivants : la résolution, la taille de tuile, la taille de lot, le nombre d'époques.

La résolution correspond à la taille sur le terrain (en réalité) d'un pixel de l'image. Elle peut être modifiée par rééchantillonnage. Par ailleurs, les images ne sont pas passées en entier au réseau de neurones, mais par morceaux. Ces sous-images sont appelées tuiles et leur taille peut être modifiée. Aussi, l'ensemble du jeu de données n'est pas passé en une seule fois dans le réseau dans ce travail, mais par paquets appelés lots. La taille du lot correspond au nombre de tuiles contenues dans un lot et ce nombre peut être modifié. Enfin, l'ensemble du jeu d'entraînement est passé plusieurs fois dans le réseau de neurones. Lorsque le jeu a été passé une fois en entier dans le réseau, on dit qu'une époque d'entraînement a eu lieu. Le nombre d'époques peut également être modifié pour entraîner le réseau.

Pour entraîner une première fois mon réseau, j'ai choisi d'utiliser les réglages suivants :

- résolution des images : 2 cm ;
  - taille des tuiles : 256 x 256 pixels ;
  - taille du lot : 9 ;
  - nombre d'époques d'entraînement : 50.
- Les résultats obtenus sur le jeu de test sont consignés dans le *tableau 1*.

On constate que le score F1 est très élevé (proche de 1, qui est la valeur maximale possible) pour la classe "non-végétation". Cela se traduit par une valeur de précision et une valeur de rappel importantes, indiquant la présence d'une faible quantité de faux positifs et faux négatifs pour cette classe. Ces résultats peuvent s'expliquer à la fois par la représentativité de cette classe dans le jeu de données (qui est plus importante que les autres) et par la grande redondance des éléments se trouvant dans cette classe. En effet, les rails, le sol nu et les toits s'y trouvent en nombre important et avec une apparence (formes et couleurs) très analogue d'une image à l'autre. Le réseau n'a donc pas de difficulté à identifier les constituants de cette classe.

Au contraire, pour la classe "reste de la végétation", le score F1 est très faible avec une valeur de 0,09. Il en est de même pour le rappel qui est de 0,05, ce qui indique la présence d'un très grand

Classe	Exactitude (en %)	Score F1	Précision	Rappel
Non-végétation	85,2	0,93	0,90	0,96
Arbres		0,76	0,73	0,79
Reste de la végétation		0,09	0,45	0,05

**Tableau 1.** Résultats sur le jeu de test avec les premiers réglages.

nombre de faux négatifs. Cela peut s'expliquer par la très grande diversité des éléments appartenant à cette classe au regard de la quantité de données dont on dispose. En effet, la classe "reste de la végétation" contient de la végétation herbacée haute, des haies, des buissons, des ronciers, etc., types de végétation qui ont des formes et un feuillage d'aspects très différents. Le réseau peine donc *a priori* à comprendre les caractéristiques de cette classe.

Enfin, pour la classe "arbres", les résultats obtenus sont relativement satisfaisants pour un premier test, avec un score F1 de 0,76 et un rappel de 0,79. Ces résultats peuvent probablement s'expliquer par le contenu de la classe, qui est relativement redondant puisqu'il consiste exclusivement en des arbres ou des arbustes. Toutefois, les arbres et arbustes, s'ils ont des formes analogues d'une image à l'autre, n'ont ni la même texture ni la même taille, ce qui peut rendre l'identification de cette classe moins triviale que la classe "non-végétation". De plus, nous avons moins de données d'entraînement pour la classe "arbres" que pour la classe "non-végétation".

Ces premiers résultats sont relativement satisfaisants pour la classe "arbres", mais nous souhaiterions essayer de diminuer encore la quantité d'oublis (faux négatifs). Pour la classe "reste de la végétation", les résultats obtenus sont trop faibles. Nous avons donc procédé de façon empirique pour sélectionner les réglages du réseau qui nous semblaient être les plus pertinents pour l'entraînement, de sorte à améliorer les résultats de segmentation sémantique pour ces deux classes et particulièrement pour la classe "reste de la végétation". Nous avons donc modifié certains des réglages précédemment évoqués un à un. Nous avons ensuite rassemblé les réglages les plus pertinents dans un dernier entraînement. Nous avons retenu les paramètres suivants :

- résolution des images : 4 cm ;
- taille des tuiles : 512 x 512 pixels ;
- taille du lot : 9 ;
- nombre d'époques d'entraînement : 50.

Les résultats obtenus pour le jeu de test sont présentés dans le *tableau 2*.

Pour la classe "arbres", on constate une légère amélioration, avec notamment une diminution de la quantité de faux négatifs matérialisée par un rappel passant de 0,79 à 0,87. La quantité de faux positifs a également baissé. Cela conduit finalement à une augmentation du score F1, passant de 0,76 à 0,83. Les résultats pour la classe "reste de la végétation" se sont eux très nettement améliorés, le score F1 passant de 0,09 à 0,58, et le rappel de 0,05 à 0,55.

Ces améliorations s'expliquent par deux raisons. D'une part, la diminution de la résolution a permis d'augmenter les résultats pour la classe "reste de la végétation". Il semblerait donc qu'avec une résolution de 2 cm, les détails étaient trop nombreux pour le réseau et le perdaient dans l'extraction des caractéristiques de la classe. D'autre part, l'augmentation de la taille des tuiles permet au réseau d'avoir plus d'informations relatives au contexte. Cela lui a vraisemblablement permis d'avoir une meilleure compréhension de la scène, impliquant une meilleure reconnaissance des éléments qui s'y trouvent.

Après cette comparaison quantitative des résultats, on peut également faire une évaluation qualitative à partir de la *figure 5*. Cette évaluation illustre les constats précédemment faits.

Au vu des résultats, nous avons finalement conservé le dernier réseau de neurones entraîné comme réseau final fourni à l'entreprise.

### ■ Post-traitement des données

Ce projet s'inscrivant dans une perspective de cartographie du risque végétation, nous avons souhaité nous intéresser au post-traitement des



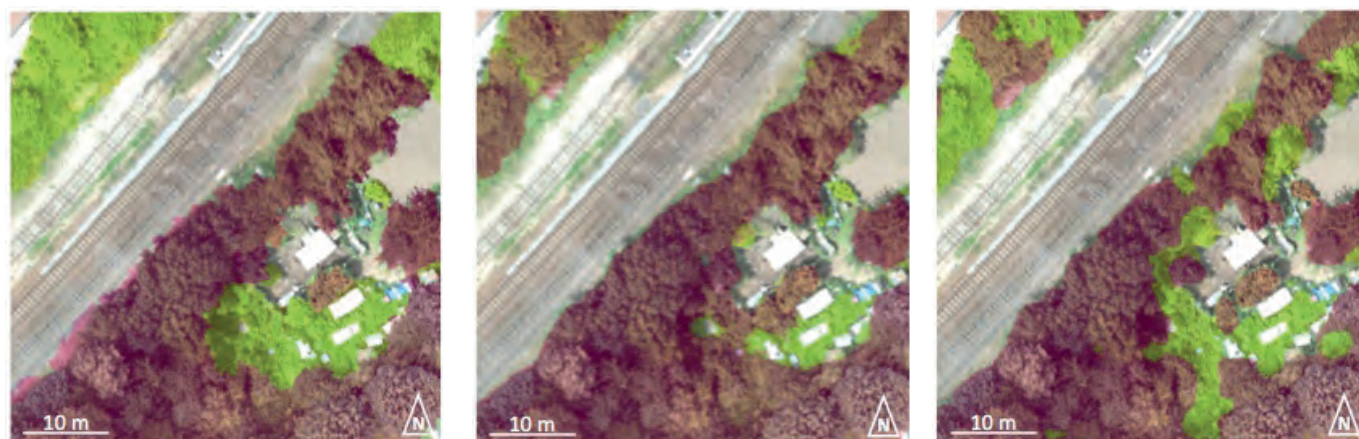


Figure 5. a) Vérité terrain, b) Prédiction à l'issue du premier test, c) Prédiction à l'issue du dernier test, en rose la classe "arbres", en jaune la classe "reste de la végétation", en beige la classe "non-végétation".

Classe	Exactitude (en %)	Score F1	Précision	Rappel
Non-végétation	88,8	0,94	0,95	0,93
Arbres		0,83	0,79	0,87
Reste de la végétation		0,58	0,62	0,55

Tableau 2. Résultats sur le jeu de test avec les derniers réglages.



images segmentées en sortie du réseau. L'objectif final consiste à obtenir des couches vectorielles des classes relatives à la végétation.

On peut constater sur la figure 6 que les images obtenues en sortie du réseau comportent quelques artefacts, des pixels isolés. Dans le but de simplifier et alléger la création future des couches vectorielles, nous avons décidé d'effectuer un débruitage des images par l'application d'opérateurs morphologiques. Un exemple de résultat est visible en figure 7.

Une fois cela réalisé, la deuxième étape du post-traitement consiste à vectoriser les classes relatives à la végétation. Tout d'abord, les images segmentées ont dû être géoréférencées. Pour cela, les

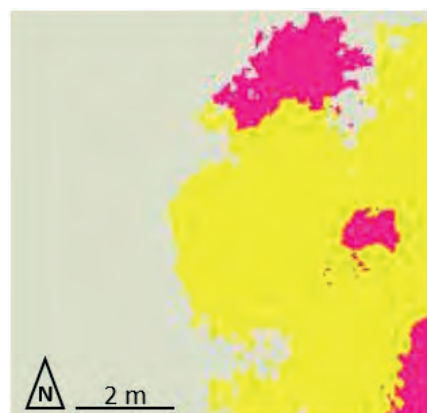
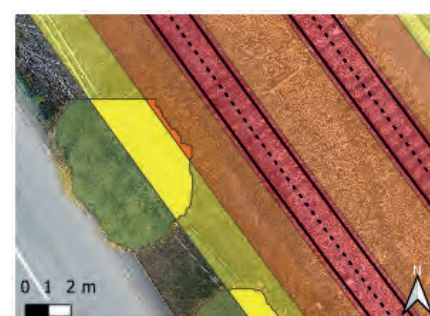


Figure 6. Extrait d'une image test classifiée montrant la présence de pixels isolés, en rose la classe "arbres", en jaune la classe "reste de la végétation", en beige la classe "non-végétation".

fichiers de géoréférencement associés à l'orthophotographie correspondante sont utilisés. Les coordonnées des sommets des polygones de végétation ont alors pu être extraites et copiées dans un fichier .GeoJSON, intégrable dans un SIG et permettant d'avoir une couche vecteur. Enfin, nous avons proposé une représentation du risque végétation en nous basant sur de précédents projets d'Altametrus. Un extrait de cette cartographie est présenté en figure 8.

## Conclusion

L'étude de l'état de l'art a permis de mettre en évidence la pertinence du recours au réseau de neurones SegNet pour la problématique posée. Nos recherches concernant la mise en place



### Légende

- Eléments de l'infrastructure ferroviaire
  - Voies
  - Position de la caténaire
- Echelle du risque végétation
  - Arbres à surveiller
  - Arbres à risque
- Zones pour l'analyse du risque végétation
  - Zone de végétation à surveiller
  - Zone de végétation à risque
  - Zone de végétation dangereuse
- Classes de végétation
  - Classe arbres

Figure 8. Extrait du SIG créé pour cartographier le risque végétation.

de la solution nous ont ensuite conduits à retenir le code informatique proposé par [Audebert et al., 2017] dans le cadre d'un travail portant sur la segmentation



Figure 7. a) Image en sortie du réseau, b) Image après application des opérations morphologiques.

sémantique d'images d'observation de la Terre. Après avoir créé notre base de données labellisée, plusieurs tests ont été menés afin de trouver les réglages les plus optimaux pour notre problématique et notre jeu de données. Nous avons finalement retenu un réseau de neurones aboutissant à des résultats de segmentation sémantique relativement satisfaisants sur le jeu de test pour la classe "reste de la végétation" avec un score F1 de 0,58 et des résultats convaincants pour la classe "arbres" avec un rappel de 0,87. Par ailleurs, pour augmenter la robustesse des résultats dans la perspective d'une industrialisation de la solution, les prédictions issues de plusieurs réseaux de neurones pourront notamment être combinées. Nous avons ensuite proposé un post-traitement des images segmentées en sortie du réseau dans une perspective de cartographie.

Grâce à cette étude, la cartographie SIG du couvert végétal, et plus spécifiquement arbustif, sur la base d'orthophotographies RVB, est rendue possible, simple et automatisée. Ce faisant, elle peut être déployée sur de grands linéaires, offrant immédiatement une première notion de proximité et d'éventuel danger avec les infrastructures ferroviaires. ●

## Contact

Hélène ROUILLON  
helene.rouillon01@gmail.com

## Bibliographie

Pereira, A. (2015). *Valorisation des données acquises par drone sur le réseau ferré national*. Mémoire de fin d'études, Institut National des Sciences Appliquées de Strasbourg.

Heudin, J.-C. (2016). *Comprendre le deep-learning - Une introduction aux réseaux de neurones*. Auto-édition, ISBN : 9791091245449, 188 pages.

[Audebert et al., 2017] Audebert, N., Le Saux, B. et Lefèvre, S. (2017). *Beyond RGB : Very High Resolution Urban Remote Sensing With Multimodal Deep Networks*. ISPRS Journal of Photogrammetry and Remote Sensing, 140:20-32.

[Breheret, 2017] Breheret, A. (2017). *Pixel Annotation Tool*. URL : <https://github.com/abreheret/PixelAnnotationTool>. Consulté le : 20-04-2020.

[Castillo-Navarro et al., 2019] Castillo-Navarro, J., Audebert, N., Boulch, A., Le Saux, B. et Lefèvre, S. (2019). *What data are needed for semantic segmentation in earth observation ?* In 2019 Joint Urban Remote Sensing Event (JURSE), pages 1-4

[Garcia-Garcia et al., 2018] Garcia-Garcia, A., Orts-Escobedo, S., Oprea, S., Villena-Martinez, V., Martinez-Gonzalez, P. et Garcia-Rodriguez, J. (2018). *A survey on deep learning techniques for image and video semantic segmentation*. Applied Soft Computing Journal, 70:41-65.

[INRA] INRA. *Ligne de partage des eaux*. URL : <http://www.pfl-cepia.inra.fr/index.php?page=tutolmg-watershed>. Consulté le 17-05-2020.

[Saniterpen, 2018] Saniterpen (2018). *Saniterpen - Club des Partenaires*. URL : <https://www.saniterpen.fr/club-des-partenaires.html>. Consulté le 30-06-2020.

## ABSTRACT

**Keywords:** Vegetation cover, neural networks, semantic segmentation, aerial imagery, vegetation risk, cartography

*Vegetation and the risk it induces regarding rail infrastructures and users represent a major stake for SNCF Réseau. Today, the railway company needs to control the vegetation risk. Therefore, this study consists in the automatic analysis of the vegetation cover for the purpose of the vegetation risk management in rail environment, by using RGB pictures acquired by UAVs, helicopters or airplanes. Thus, a neural network aimed at a semantic segmentation is implemented. Indeed, thanks to their learning ability, well-trained neural networks are able to classify efficiently any new picture. Three classes were defined according to the risks related to the vegetation: "trees", "rest of the vegetation" and "non-vegetation". A complete processing chain of data was proposed, which allows, on the basis of these pictures, to make a Geographic Information System (GIS) cartography. This expertise is easily deployable on full lines and should allow the maintenance worker to target the priority areas and to optimize his pruning plans.*



**COMMANDEZ "LES RÉFÉRENCES DE TEMPS ET D'ESPACE" AU PRIX DE 35,00 €**  
17 x 24 cm, 484 pages  
frais de port inclus (France)

M/Mme Nom : \_\_\_\_\_

Prénom : \_\_\_\_\_

Société ou organisme : \_\_\_\_\_

Adresse : \_\_\_\_\_

Code postal :

Ville : \_\_\_\_\_

Tél. :

Fax :

Courriel : \_\_\_\_\_

Date : \_\_\_\_\_

Signature \_\_\_\_\_

**Bulletin de commande à retourner accompagné d'un chèque à l'Association francophone de topographie**  
73, avenue de Paris  
94165 SAINT-MANDÉ Cedex  
Tél. : +33 (0) 1 43 98 84 80  
**Achat également sur Internet : [www.aftopo.org](http://www.aftopo.org)**