

Développement d'un indice de fréquentation touristique grâce aux camera traps et aux réseaux neuronaux convolutionnels

Auteur : Guidosse, Quentin

Promoteur(s) : Dufrêne, Marc

Faculté : Gembloux Agro-Bio Tech (GxABT)

Diplôme : Master en bioingénieur : gestion des forêts et des espaces naturels, à finalité spécialisée

Année académique : 2018-2019

URI/URL : <http://hdl.handle.net/2268.2/7829>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.



LIÈGE université
Gembloux
Agro-Bio Tech

DÉVELOPPEMENT D'UN INDICE DE FRÉQUENTATION TOURISTIQUE GRÂCE À L'USAGE DES CAMERA TRAPS ET DES RÉSEAUX NEURONAUX CONVOLUTIONNELS

QUENTIN GUIDOSSE

**TRAVAIL DE FIN D'ÉTUDES PRÉSENTÉ EN VUE DE L'OBTENTION DU DIPLÔME DE
MASTER BIOINGÉNIEUR EN GESTION DES FORÊTS ET DES ESPACES NATURELS**

ANNÉE ACADÉMIQUE 2018-2019

PROMOTEUR : PR. MARC DUFRÈNE

Le présent document n'engage que son auteur

Toute reproduction du présent document, par quelque procédé que ce soit, ne peut être réalisée qu'avec l'autorisation de l'auteur et de l'autorité académique de Gembloux Agro-Bio Tech.

**DÉVELOPPEMENT D'UN INDICE DE
FRÉQUENTATION TOURISTIQUE GRÂCE À
L'USAGE DES CAMERA TRAPS ET DES
RÉSEAUX NEURONAUX CONVOLUTIONNELS**

QUENTIN GUIDOSSE

TRAVAIL DE FIN D'ÉTUDES PRÉSENTÉ EN VUE DE L'OBTENTION DU DIPLÔME DE
MASTER BIOINGÉNIEUR EN GESTION DES FORÊTS ET DES ESPACES NATURELS

ANNÉE ACADÉMIQUE 2018-2019

PROMOTEUR : PR. MARC DUFRÈNE

Remerciements

Je tiens tout d'abord à remercier le laboratoire de Biodiversité et Paysage et mon promoteur le Pr. Marc Dufrêne pour son expertise, ses conseils avisés et son encadrement dans la réalisation de ce travail de fin d'études. Il a toujours su me rediriger lorsque l'objectif final devenait flou et lorsque mes recherches prenaient des directions périlleuses. Je remercie également Johanna Breyne, doctorante travaillant sur le projet AGRETA qui m'a suivi tout au long de mon TFE. Elle m'a tout d'abord fourni les données nécessaires à l'élaboration de ce document et m'a également apporté de nombreuses autres ressources telles que de la documentation, des conseils et de précieuses relectures.

Je suis reconnaissant envers l'aide que m'a apportée la plateforme CAMI de l'ULiège GxABT grâce à laquelle j'ai pu analyser de nombreuses photographies sur un serveur facile d'accès. Je remercie particulièrement Nicolas Gengler, qui m'a accordé cet accès et a pris de son temps pour m'expliquer les règles de fonctionnement de cet outil et ne m'a pas tenu rigueur de quelques petites erreurs de manipulation. Merci aussi à Bernard Segers qui a configuré mon environnement informatique avec patience et résolu chacun de mes soucis liés à ce domaine avec une grande efficacité.

Débutant complet dans le deep learning au départ de ce travail, je remercie par ordre de consultation Stijn Van Hoey (INBO, Instituut voor Natuur- en Bosonderzoek), François Laurent (Aerospacelab) et le Pr. Marc Vandroogenbroeck (département d'électricité, électronique et informatique, Institut Montefiore, ULiège) qui ont accéléré grandement mon apprentissage des bases de cette discipline. Chacun avec des idées différentes, ils m'ont permis d'évaluer la quantité de travail nécessaire pour créer et manipuler ce type d'algorithme et m'ont aiguillé vers des outils à ma portée.

Je remercie aussi David Nunes que j'ai rencontré en ligne pour avoir pris de son temps afin d'adapter certains packages Python à mes besoins et à David Colignon pour avoir pris de son temps afin de m'expliquer le fonctionnement de la plateforme CÉCI (Consortium des Équipements de Calcul Intensif)

Merci aux membres de mon jury pour le temps accordé à la lecture de mon TFE et l'attention portée à mon travail. Individuellement, merci au Pr. Jacques Hébert pour ses conseils et sa bienveillance lors de ses communications par mail, au Pr. Hélène Soyeurt pour ses conseils lors de mon apprentissage du langage Python et son cours sur le machine learning et enfin au Pr. Philippe Lejeune pour ses cours pratiques ainsi que sa rapidité de réponse lors de la recherche d'un GPU.

Merci à Wissal Selmi d'avoir relu mon TFE et de m'avoir donné des conseils avisés.

Enfin, merci à ma famille pour son amour et son soutien sans failles durant ces huit années d'études supérieures ainsi que pour la tolérance dont ils ont fait preuve envers ma mauvaise humeur fréquente lors des périodes de stress.

Résumé

Les services écosystémiques récréatifs sont généralement mal évalués ou ignorés. Cependant, une meilleure évaluation de ces services devrait permettre aux gestionnaires de comprendre plus finement les dynamiques des visiteurs dans ces milieux. Cela aiderait à mieux lutter contre la dégradation des milieux naturels, mais aussi à mettre certains écosystèmes plus en valeur. Dans ce travail, un prototype d'une nouvelle technique d'évaluation de la fréquentation touristique est élaboré. Cette technique utilise les camera traps pour collecter des données sur le nombre de visiteurs et sur leur comportement. Le travail d'analyse de ces images, souvent conséquent, est confié à un algorithme de deep learning nommé Mask R-CNN. D'une part, les résultats montrent que la position des caméras traps doit être homogénéisée pour faciliter le fonctionnement de l'algorithme. D'autre part, le modèle détecte $89,0 \pm 4,2$ % du temps les personnes et les classe correctement $97,1 \pm 1,0$ % du temps sur des photos floues. Néanmoins, les autres classes (chiens, vélos, véhicules, sacs à dos) ne sont pas encore correctement détectées. Le modèle doit être partiellement entraîné sur des données en extérieur et un prétraitement des images par segmentation devrait être utilisé. Une étude de cas réalisée sur le projet AGRETA montre que, même si la technique doit être améliorée, il est déjà possible d'obtenir certaines données précises sur le long terme. De plus, pouvoir consulter les images lors d'événements suspects dans les résultats est une source d'informations inattendue, mais riche.

Mots-clés : services écosystémiques culturels, réseaux neuronaux convolutionnels, CNN, Mask R-CNN, piège caméra, fréquentation touristique, nombre de visiteurs, écotourisme, Ardenne, parcs nationaux

Abstract

Recreational ecosystem services are generally poorly valued or ignored. However, a better evaluation of these services should allow managers to better understand the dynamics of visitors in these environments. This would help to better prevent the degradation of natural environments but also to highlight certain ecosystems. In this work, a prototype of a new technique for visitor monitoring is being developed. This technique uses camera traps to collect counting and behavioural data of visitors. The analysing task of these images, often substantial, is carried out by a deep learning algorithm called Mask R-CNN. The results show, on the one hand, that the position of the camera traps must be homogenized to facilitate the operation of the algorithm. On the other hand, the model detects people $89.0 \pm 4.2\%$ of the time and classifies them correctly $97.1 \pm 1.0\%$ of the time on blurred photos. However, the other classes (dogs, bicycles, vehicles, backpacks) are not correctly detected yet. The model must be partially trained on outdoor data, and image pre-processing by segmentation should be used. A case study of the AGRETA project shows that although the technique needs to be improved, it is already possible to obtain some accurate data over the long term. In addition, the ability to view images during suspicious events in the results is an unexpected but rich source of information.

Keywords: cultural ecosystem services, convolutional neural network, CNN, Mask R-CNN, camera trap, visitor monitoring, visitor number, ecotourism, Ardenne, national parks

Table des matières

REMERCIEMENTS.....	I
RÉSUMÉ.....	II
ABSTRACT.....	III
TABLE DES MATIÈRES.....	IV
INTRODUCTION.....	1
CADRE CONCEPTUEL	3
I. LES SERVICES ÉCOSYSTÉMIQUES	3
1) DÉFINITIONS	3
a. <i>Les services d’approvisionnement ou de production</i>	4
b. <i>Les services de régulation</i>	5
c. <i>Les services socioculturels</i>	5
2) COMPARAISONS ET INTERDEPENDANCES ENTRE SERVICES.....	6
3) FOCUS SUR LES SERVICES ÉCOSYSTÉMIQUES RÉCRÉATIFS.....	7
II. LES APPROCHES D’ÉVALUATION DES SERVICES ÉCOSYSTÉMIQUES RÉCRÉATIFS.....	10
1) ASPECTS ÉCONOMIQUES DÉCOULANT DES INDICES DE FRÉQUENTATION TOURISTIQUE	10
2) ASPECTS ÉCOLOGIQUES DÉCOULANT DES INDICES DE FRÉQUENTATION TOURISTIQUE	11
3) LES DIFFÉRENTES TECHNIQUES D’ÉVALUATION DE LA FRÉQUENTATION TOURISTIQUE	11
a. <i>Les techniques quantitatives</i>	12
b. <i>Les techniques qualitatives</i>	13
c. <i>Les nouvelles technologies</i>	14
III. CAMERA TRAPPING ET AUTOMATISATION	16
1) DÉFINITION, HISTOIRE ET FONCTIONNEMENT DES CAMERA TRAPS	16
2) AVANTAGES ET LIMITES DES CAMERA TRAPS.....	17
a. <i>Avantages</i>	17
b. <i>Limites</i>	17
3) CAMÉRA TRAPPING ET TOURISME	19
a. <i>Camera trapping et respect de la vie privée</i>	20
4) RÉSEAUX NEURONAUX CONVOLUTIONNELS ET AUTOMATISATION.....	21
a. <i>Réseaux neuronaux convolutionnels (CNN)</i>	22
MATÉRIEL ET MÉTHODE.....	23
I. LE PROJET AGRETA.....	24
1) CONTEXTE SPATIAL DE L’ÉTUDE	25
a. <i>La réserve naturelle des Hautes Fagnes</i>	26
b. <i>Le parc naturel des Deux Ourthes</i>	26
c. <i>La Grande Forêt de Saint-Hubert</i>	27
d. <i>La Grande Forêt d’Anlier</i>	27
II. CHAÎNE DE TRAITEMENT	28

Table des matières

1) ACQUISITION DES DONNÉES	29
2) STOCKAGE DES DONNÉES	31
3) TRAITEMENT DES DONNÉES	31
III. EVALUATION DE LA QUALITÉ DU MODÈLE	31
IV. CRÉATION DE LA BASE DE DONNÉES.....	32
V. APPLICATION DE L'INDICE SUR LES DONNEES COLLECTEES PAR AGRETA.....	34
1) COMPARAISON AVEC LA METHODE DE LUPP ET AL. (2016)	34
2) QUELLES INFORMATIONS PEUVENT ETRE TIREES DE CE NOUVEL INDICE ?.....	34
RÉSULTATS.....	35
I. ETUDE DU NOMBRE DE PHOTOS OBTENUES.....	35
1) RÉSULTATS ET VARIATIONS TEMPORELLES	35
2) RELATIONS ENTRE CAMERAS	40
II. ETUDE DU NOMBRE DE PERSONNES COMPTÉES.....	41
1) EVALUATION DE L'EFFICACITÉ DU MODÈLE	41
2) RÉSULTATS ET VARIATIONS TEMPORELLES	45
3) RELATIONS ENTRE CAMÉRAS	49
4) ETUDE DE LA TAILLE DES GROUPES.....	50
III. LIENS ENTRE LE NOMBRE DE PHOTOS ET LA FREQUENTATION.....	52
1) COMPARAISON AVEC LA METHODE DE LUPP ET AL. (2016)	52
2) COMPARAISON ENTRE LE NOMBRE DE PHOTOS ET LE NOMBRE DE PERSONNES TOTAL.....	54
DISCUSSION.....	56
I. LA FRÉQUENTATION TOURISTIQUE EN ARDENNE.....	56
II. COMPARAISON À LA MÉTHODE DE LUPP ET AL. (2016).....	59
III. QUALITÉ DES RÉSULTATS.....	59
1) CONFIGURATION DES CAMÉRAS	59
2) QUALITÉ DU MODÈLE.....	61
IV. AUTRES PISTES D'OPTIMISATION DE LA MÉTHODE	63
CONCLUSION.....	64
REFERENCES BIBLIOGRAPHIQUES.....	65
ANNEXES.....	76
ANNEXE 1 : HISTOIRE ET FONCTIONNEMENT DES RESEAUX NEURONAUX CONVOLUTIONNELS	76
1) INTELLIGENCE ARTIFICIELLE ET AUTOMATISATION	76
a. Réseaux neuronaux artificiels.....	76
b. Réseaux neuronaux convolutionnels (CNN).....	77
c. Différentes méthodes d'apprentissage.....	80
ANNEXE 2 : PROTOCOLE D'UTILISATION DE L'INDICE ET PISTES D'AMELIORATIONS	81
1) ACQUISITION DES PHOTOS A ANALYSER ET PREPARATION DES OUTILS.....	81
2) LE SCRIPT PYTHON.....	90
3) CRÉATION BASE DE DONNÉES	93

Table des matières

ANNEXE 3 : MASK R-CNN.....	99
1) DÉTECTION DES OBJETS	100
2) SEGMENTATION DES OBJETS	101
ANNEXE 4 : LISTE DES CLASSES CONNUES PAR MASK R-CNN	102
ANNEXE 5 : RESULTATS INTERMEDIAIRES DE L'ETUDE DES RELATIONS ENTRE LES CAMERAS PAR RAPPORT A LEUR NOMBRE DE PHOTOS.....	103
ANNEXE 6 : RESULTATS DU TEST DE LA QUALITE DU MODELE	104
ANNEXE 7 : RESULTATS INTERMEDIAIRES DE L'ETUDE DES RELATIONS ENTRE CAMERAS POUR LE COMPTAGE DU NOMBRE DE PERSONNES.....	107
ANNEXE 8: RESULTATS DU TEST DE LA METHODE DE SUBSAMPLING ET EXTRAPOLATION DE LUPP ET AL. (2016) SUR LES DONNEES D'AGRETA.....	108
ANNEXE 9 : RESULTATS COMPLEMENTAIRES CONCERNANT L'ETUDE DE LA TAILLE DES GROUPES.....	111
ANNEXE 10 : CRITIQUE DE LA POSITION DE CHAQUE CAMERA	112
ANNEXE 11 : COMPARAISON ENTRE LES DONNEES CAMERAS ET PROXIMUS.....	113

Introduction

Comme l'a récemment à nouveau démontré le rapport planétaire de l'IPBES (2019), l'état des ressources naturelles est alarmant. Le facteur qui a le plus d'impact sur les écosystèmes est le changement d'affectation des terres et des mers. En 2005, la fragmentation de l'habitat était déjà la principale source d'inquiétude (Millennium Ecosystem Assessment, 2005). Bien que nous vivions tous sur la Terre qui est un espace limité, l'augmentation démographique et ses impacts multiples ne semblent pas avoir de fin. D'ici 2030, la démographie mondiale devrait augmenter de 10 %, ce qui ferait grimper la population à 8.5 milliards de personnes (United Nations, 2019). Il est donc indispensable de trouver des solutions afin d'agencer cet espace de façon optimale et durable pour réduire l'empreinte de l'Homme de façon à la fois locale et globale.

La notion de service écosystémique est née d'une approche anthropocentriste de préservation de la nature. Les services écosystémiques (SE) sont l'ensemble des biens et services que la nature prodigue à l'homme (Costanza et al., 1997; Millennium Ecosystem Assessment, 2005). Les services écosystémiques de production sont facilement quantifiables puisqu'il existe des marchés économiques pour la plupart. Les services de régulation peuvent majoritairement être évalués par des mesures biophysiques. Concernant les services socioculturels, l'évaluation est souvent plus compliquée à cause de leur intangibilité (Gómez-Baggethun et al., 2013). Par manque d'indicateurs, ce groupe de services écosystémiques est souvent mis de côté lors des décisions, ce qui masque tout une partie de ces relations (Chan et al., 2012). En remettant tous les groupes de services écosystémiques au même plan, les aspects multifonctionnels de la forêt et des autres milieux pourraient être mieux remplis. Par exemple, remplacer les peuplements purs par des peuplements mélangés augmente l'importance des services écosystémiques de régulation et socioculturels (Carnol et al., 2010). Les besoins économiques qui pèsent sur les forêts productives pourraient même continuer à être assouvis en considérant davantage le potentiel économique des services écosystémiques culturels et particulièrement récréatifs (Schägner et al., 2017).

Ces services écosystémiques récréatifs sont sous-évalués ou ignorés lorsqu'ils ne sont pas directement liés à une valeur économique (Daniel et al., 2012), c'est-à-dire lorsque l'entrée dans le milieu naturel n'est pas payante. La fréquentation touristique diffuse (non liée à des infrastructures) des milieux naturels est généralement étudiée via des interviews et des comptages sur de courtes périodes de temps (Muhar et al., 2002). Ces informations sont alors extrapolées à l'état général du milieu et engendrent une gestion incohérente des espaces naturels.

L'objectif de ce travail est d'explorer une nouvelle technique d'évaluation de la fréquentation touristique des forêts et des milieux ouverts grâce aux pièges caméras. Cette technique doit permettre de fournir des informations fiables sur le long terme aux gestionnaires d'espaces naturels. Elle doit renseigner sur le nombre de personnes fréquentant un milieu, mais aussi sur leur comportement de façon à avoir une compréhension plus fine de la dynamique des visiteurs. Les pièges caméras constituent une solution économique pour rassembler ces informations sur le long terme. L'une des raisons principales pour laquelle elles ne sont pas largement utilisées est la quantité d'images à analyser après la récolte de données. Aujourd'hui,

Introduction

grâce à la démocratisation du « deep learning », il est désormais possible d'automatiser le traitement de ces données. Les algorithmes les plus performants à ce jour pour ce type de tâche sont les réseaux neuronaux convolutionnels. L'architecture Mask R-CNN (He et al., 2017) est utilisée dans ce cas.

Ce travail consiste à créer un prototype de chaîne de traitement de données issues de ces caméras. Pour élaborer ce premier test, les photos récoltées dans le cadre du projet AGRETA (Ardenne Grande Région, Eco-Tourisme et Attractivité) sont utilisées. AGRETA est un projet Interreg se déroulant de janvier 2017 à décembre 2020. Il vise à développer l'attractivité transfrontalière sur le massif ardennais tout en améliorant sa compétitivité économique et la gestion durable de ses ressources. Dans ce cadre, 20 caméras ont été placées sur quatre massifs de l'Ardenne belge à partir de juillet 2018. Il s'agit des massifs de la réserve naturelle des Hautes Fagnes-Eifel, du parc naturel des Deux Ourthes, de la grande forêt de Saint-Hubert et de la grande forêt d'Anlier. Les caméras ont été placées sur des chemins fréquentés par différents types de visiteurs. Pour respecter la loi sur la vie privée, une bande adhésive a été collée sur la lentille pour flouter l'image.

La première partie de ce document définit un cadre conceptuel, en analysant brièvement les différents types de services écosystémiques. Ensuite, les différentes approches d'évaluation de la fréquentation touristique sont décrites avant d'aborder l'usage des pièges caméras. Les avantages et les défauts de cet outil sont discutés de manière générale puis en lien avec le tourisme. Finalement, l'analyse des images par intelligence artificielle est décrite de manière théorique en expliquant le fonctionnement des réseaux neuronaux convolutionnels. Dans la seconde partie, les différentes étapes de la chaîne de traitement sont explicitées. En partant du contexte spatio-temporel de l'étude, la chaîne de traitement est parcourue jusqu'à la création d'une base de données non-redondante en passant par l'explication du fonctionnement du modèle utilisé. Finalement, la qualité des résultats du modèle ainsi qu'une critique de la méthode est réalisée. Le prototype ainsi créé est alors testé sur un cas d'étude en analysant les données de fréquentation d'AGRETA. Ensuite, il est comparé à une méthode plus classique utilisée par Lupp et al. (2016) qui consiste à analyser manuellement un sous-échantillon des photos récoltées et à extrapoler les comptages de personnes sur l'ensemble du dataset par régression linéaire.

Cadre conceptuel

Cette première partie délimite un cadre conceptuel en définissant la notion de service écosystémique (SE). Après avoir détaillé de manière succincte les différents groupes de SE et passé en revue les différentes typologies existantes, l'attention est davantage portée sur les services écosystémiques récréatifs, faisant partie des SE socioculturels. Les aspects économiques et écologiques qui en découlent sont aussi discutés. Enfin, les différentes techniques permettant d'évaluer ce type de service écosystémique sont détaillées et commentées. Une attention particulière est portée à la méthode employée dans le cadre de ce travail. La fin de ce chapitre traite donc en détail de l'utilisation des caméras traps adaptée à l'étude de la fréquentation touristique ainsi que de l'utilisation des réseaux neuronaux convolutionnels dans le but de résoudre un des principaux défauts de cet outil : la charge de travail nécessaire pour traiter les données récoltées.

I. Les services écosystémiques

1) Définitions

Les services écosystémiques (SE) sont l'ensemble des biens et services que la nature prodigue à l'homme (Costanza et al., 1997; Millennium Ecosystem Assessment, 2005). Ils forment un flux de matériaux, d'énergie et d'information provenant de la nature et se combinent ou non au capital et à la confection humaine afin d'augmenter son bien-être (Costanza et al., 1997). Ces SE ont été évalués de manière globale dans le Millenium Ecosytem Assessment (2005) qui est un rapport de recherche international destiné à rendre compte de l'état des écosystèmes dans le monde et des changements qu'ils subissent dans le but d'en évaluer les

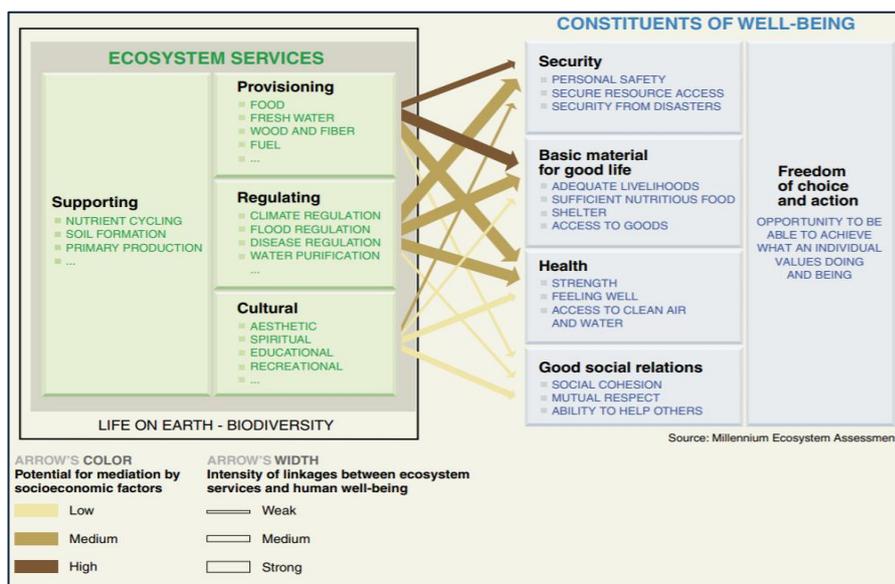


Figure 1: Les différents liens entre les services écosystémiques et leur impact sur le bien être humain. La couleur des flèches indique les liens plus ou moins forts avec les facteurs socioéconomiques. Leur largeur indique l'importance du SES par rapport au bien (MEA, 2005)

Cadre conceptuel

conséquences et d'en dégager des actions. Cette notion de service écosystémique met en avant l'importance de la structure et du fonctionnement des écosystèmes naturels pour l'humanité. À l'échelle mondiale, le capital naturel vaudrait entre 16 et 44 trillions (10^{12}) d'US dollars avec une moyenne estimée à 33 trillions (10^{12}) d'USD, soit 1,8 le RNB mondial (Costanza et al., 1997). La dégradation incessante et accrue de ces écosystèmes a donc un prix conséquent pour la société (Schneiders et al., 2012; Vo et al., 2012).

Parmi les différentes typologies, on retrouve au niveau international le MEA et « The Economics of Ecosystems and Biodiversity » (TEEB, 2010) qui en découle. En 2012 sortait l'équivalent du GIEC pour la biodiversité et les services écosystémiques, l'« Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services » (IPBES) qui a été créé afin d'évaluer au niveau mondial l'état de la biodiversité et des services écosystémiques. Contrairement au Millenium Ecosystem Assessment, l'IPBES est une structure pérenne. Le rapport de l'IPBES (2019) a défini dans son cadre conceptuel la notion de « contribution de la nature envers les personnes » (NCP) qui rejoint l'approche des services écosystémiques du Millenium Ecosystem Assessment tout en mettant davantage l'accent sur les aspects socioculturels de la nature et la reconnaissance de l'importance des populations locales et indigènes dans la compréhension de la nature.

Au niveau européen, la typologie du Common International Classification of Ecosystem Services (CICES) (Haines-Young et al., 2018) a été développée dans le but d'adapter cette approche aux pays densément peuplés. De plus, il existait un besoin de standardisation entre les pays voisins qui facilite non seulement l'évaluation des SE à grande échelle mais aussi la communication entre ces pays.

En Wallonie, la plateforme Wal-ES (2016) joue le rôle d'interface entre les pouvoirs publics et les universités dans le but de faire valoir les travaux scientifiques et les outils mis au point par les scientifiques en faveur des services écosystémiques. Cette typologie sépare les services écosystémiques en trois grands groupes : les services de production, de régulation et les services culturels. Au sein de ces trois groupes, sont répertoriés 60 types de services écosystémiques différents.

Il ressort de ces différentes typologies trois principaux groupes de services écosystémiques¹ (Figure 1) :

a. Les services d'approvisionnement ou de production

Ce sont l'ensemble des produits tangibles que fournit la nature. On y retrouve des aliments (végétaux, animaux, mycètes) mais aussi l'eau, les matériaux ou encore les ressources génétiques créées par la nature. Les usages sont multiples. L'Homme s'en sert pour se nourrir, comme matériaux de construction ou de confection ou encore comme combustible pour créer de l'énergie. Dans ce groupe on retrouve la plupart des SE pour lesquels il existe un marché financier préexistant. Ils sont donc faciles à évaluer économiquement. Au fil de l'histoire, les hommes se sont toujours tournés vers ce type de services en négligeant les autres afin de maximiser leur survie (Rodríguez et al., 2006; Kareiva et al., 2007).

¹ Aujourd'hui, les services écosystémiques de soutien ou de support sont appelés « fonctions écologiques » et supportent le fonctionnement des écosystèmes.

b. Les services de régulation

Les services écosystémiques de régulation sont assurés par les fonctions écologiques qui maintiennent la vie de l'Homme possible sur Terre. Ces services sont par exemple le maintien du climat (Mooney et al., 2009), la pollinisation, la dépollution de l'air, de l'eau, du sol, le stockage du carbone, la protection contre les catastrophes naturelles (van Loon-Steensma et al., 2013), la lutte biologique (Civitello et al., 2015) ou encore le maintien de la fertilité des sols. Bien qu'ils soient indispensables pour l'Homme, leur lien indirect avec eux fait qu'ils sont seulement reconnus lorsqu'ils sont mis en péril ou perturbés (de Groot et al., 2002).

c. Les services socioculturels

Ce sont l'ensemble des avantages immatériels que les humains peuvent tirer de la nature. On y retrouve entre autres les aspects esthétiques, spirituels, récréationnels et d'appartenance (Kumar et al., 2008). On pourrait aussi classer dans cette rubrique l'ensemble des connaissances que la nature nous enseigne pour nous permettre d'évoluer, notamment au niveau scientifique. On parle de biomimétisme.

La Figure 1 *Error! Not a valid bookmark self-reference.* montre que les services écosystémiques socioculturels sont considérés par le Millenium Ecosystem Assessment (2005) comme ceux qui impactent le moins le bien-être humain, mais surtout comme ayant le moins de potentiel économique. Or, ceux-ci comprennent en partie le secteur du

Le tourisme est la troisième catégorie d'exportation mondiale après la chimie et les carburants. Il est passé de 25 millions de touristes en 1950 à 1.240 millions en 2016 dont 182 millions en Europe occidentale (UNWTO 2018)

tourisme. Les touristes se déplaçant pour la nature sont de plus en plus nombreux (Bell et al., 2007). Entre 1980 et 2009, ils sont passés de 2 à 20 % dans le monde (Buckley, 2009). Cela montre donc que l'intérêt apporté aux différents types de services écosystémiques pourrait être en réalité déséquilibré et que les services écosystémiques socioculturels sont trop peu étudiés

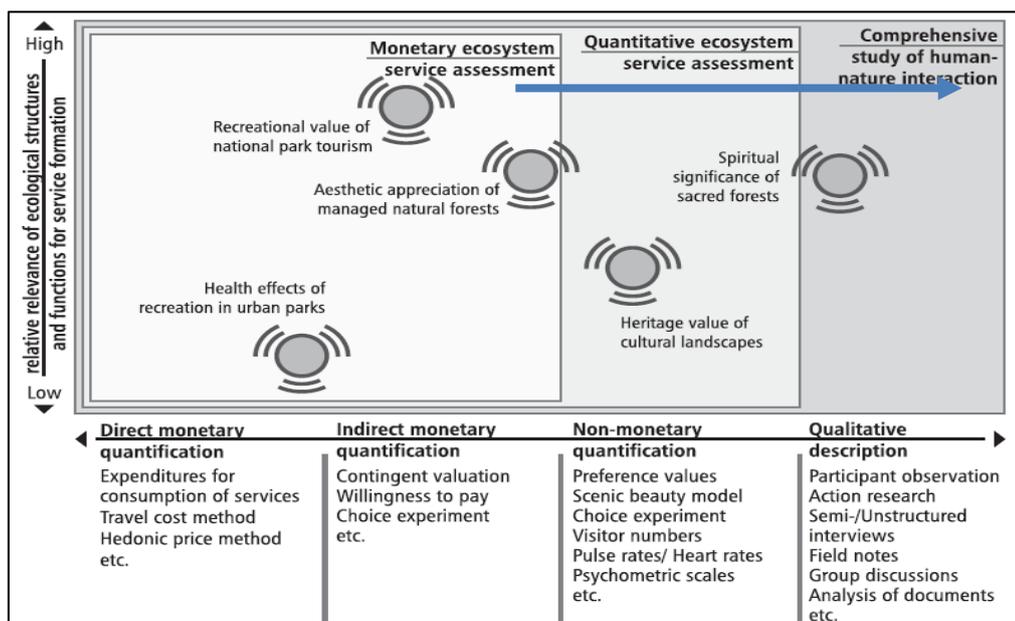


Figure 2: Exemples de services écosystémiques culturels en fonction de leurs liens avec les structures écologiques et leur transposabilité économique. Tiré de Daniel et al. (2012). La ligne bleue désigne les aspects qui devraient être davantage pris en compte dans l'étude de la fréquentation touristique des espaces naturels

Cadre conceptuel

(Norton et al., 2012). D'autres types de services écosystémiques socioculturels peuvent aussi avoir un impact économique significatif, bien que moins direct. À titre d'exemple, une étude de l'inventaire national des forêts du Royaume-Uni montre que l'accès aux espaces verts permettrait d'économiser environ 2,1 milliards de livres sterling (2,3 milliards d'euros) en soins et qu'un comportement 1 % moins sédentaire réduirait d'1,4 milliards ces dépenses (1,5 milliards d'euros) par an (Forest Research, 2010). De nombreux autres articles mettent en avant les bénéfices des écosystèmes pour la santé et les économies que cela génère (Hartig et al., 2003; Karmanov et al., 2008; Bowler et al., 2010).

Ce travail concerne ce groupe de services écosystémiques et particulièrement les services socioculturels récréatifs. La Figure 2 montre que ces services sont très importants pour la structure écologique des écosystèmes. En revanche, les aspects quantitatifs non-monétaires et qualitatifs doivent aussi être pris en compte.

2) Comparaisons et interdépendances entre services

Puisque l'économie est la discipline qui a toujours été maître dans la question de l'évaluation (Maillefert et al., 2017), elle semble pour beaucoup la plus adaptée pour comparer les services écosystémiques. Certains services écosystémiques ont un impact direct sur la prospérité de l'Homme, tels que la production de bois ou encore la dépollution de l'air. Ces bénéfices sont facilement transposables en valeur monétaire, selon le cours du bois sur les marchés ou selon la somme d'argent qu'il faudrait investir pour dépolluer une certaine quantité d'air. Cependant, l'évaluation des services écosystémiques de manière générale est souvent vue avec scepticisme (Brockington, 2011). La principale raison est que certains sont très difficiles à identifier et à quantifier précisément à cause de leur intangibilité (Richards et al., 2018), surtout en ce qui concerne les SE socioculturels (Lupp et al., 2016).

La notion de service écosystémique est née d'une vision anthropocentrée et utilitariste de la nature, vue comme un support pour la vie humaine (Vallet, 2009). En opposition, la vision biocentriste de la nature met l'Homme au second plan et lutte contre l'instrumentalisation de la nature (Cobb et al., 1976)

Les différents SE ne sont pourtant pas indépendants. Chaque fonction écologique remplie par un écosystème peut influencer sur un ou plusieurs services écosystémiques à la fois et cette globalité doit être prise en compte (Costanza et al., 1997). Pour cette raison, la sous-représentation des services écosystémiques socioculturels au profit des services de production et de régulation bloque non seulement des opportunités d'actions à l'échelle du milieu, mais cache aussi tout une partie des relations existantes entre l'ensemble des services (Daniel et al., 2012).

3) Focus sur les services écosystémiques récréatifs

Les SE récréatifs sont une partie importante des SE socioculturels. Ils peuvent prendre différents noms selon les typologies :

- MEA and TEEB: “*Recreation and ecotourism*”
- CICES v4.3 (2013): “*Physical and experiential interactions*”
- Wal-ES : n'apparaît pas clairement, mais pourrait s'apparenter à la catégorie “*Environnement pour les loisirs*”.

L'écotourisme, fait donc partie des définitions du MEA et du TEEB qui en découle. Il s'agit d'un concept relativement récent dont les premières publications à ce sujet sont sorties aux alentours des années 1990 (Ventura, 2015). L'écotourisme est défini la première fois par Ceballos-Lascuráin (1996) comme un « *voyage environnementalement responsable et la visite d'habitats naturels relativement non perturbés, pour apprécier et profiter de la nature (et de tout autre aspect culturel, passé et présent) qui promeut la conservation, avec un faible taux de visiteurs et qui prodigue une amélioration socioéconomique active pour les populations locales* » (traduction libre). Le concept d'écotourisme est cependant généralement associé à Budowski (1976) qui est le premier à traiter d'une symbiose possible entre le tourisme et la conservation. Le challenge majeur est de rester dans le concept de tourisme durable, sans basculer vers une espèce de tourisme nature de masse (Ventura, 2015). Le tourisme nature de masse est plutôt tourné vers les activités de détente (plage, mer, soleil, ...) et d'aventure, avec des composantes de durabilité et de conservation moins marquées (Weaver, 2001). Colson (2009) sépare également le tourisme encadré par des infrastructures ou localisé à un endroit du tourisme diffus, réparti à l'échelle du milieu et où les visiteurs peuvent venir pour différentes raisons. Dans ce travail, l'accent est mis sur le tourisme diffus.

Depuis bien trop longtemps, l'évaluation de ces services écosystémiques récréatifs ont été inexistantes ou imprécises, se basant sur des estimations grossières (Watson et al., 2000; Muhar et al., 2002). Prendre davantage en compte les services écosystémiques récréatifs pourrait améliorer la compréhension de la dynamique des flux touristiques par les gestionnaires et combler le fossé entre la gestion durable de ces espaces naturels et ces flux (Muhar et al., 2002; Campbell, 2010; Bartczak et al., 2012; Bartczak, 2015). De plus, gérer ces espaces naturels de façon plus adaptée aux attentes des visiteurs pourrait permettre de développer davantage l'économie locale et de réinjecter ces revenus supplémentaires dans la conservation et la promotion des milieux (Whitelaw et al., 2014). À l'inverse, des flux de visiteurs mal gérés dégradent inévitablement ces espaces (Stabler, 1997; Weaver, 2001).

Selon Eagles (2014), pour bien gérer une aire protégée, il faut porter une attention particulière à dix facteurs : le suivi des visiteurs, le suivi de l'impact économique du tourisme sur le parc, les compétences professionnelles nécessaires pour la gestion du tourisme, la construction d'une structure d'appui au publique, la satisfaction des visiteurs, les accès légaux pour le tourisme, la politique de paiement, la capacité de gestion et la gouvernance du parc.

Les critères liés au suivi des visiteurs et à l'étude de leur satisfaction sont sûrement les plus importants puisqu'ils permettent de mieux appréhender les autres. En général, les services écosystémiques récréatifs sont facilement évaluables lorsque les gens paient pour avoir accès à certains espaces naturels, mais cet accès n'est que rarement payant (Campbell, 2010). Pourtant, les personnes qui se rendent dans ces milieux y investissent du temps, mais surtout de l'argent

Cadre conceptuel

en transport, en restauration et en logement. Pour aller plus loin, il peut aussi s'agir de coûts non dépensés par l'amélioration de la santé des personnes grâce à la nature. Il y a donc de nombreuses contributions indirectes à prendre en compte dans l'évaluation de ce type de service écosystémique. En 2005, l'ensemble des parcs nationaux des Etats-Unis générait plus de 10 milliards de dollars dont seulement 10 % découlaient des frais d'entrées. Les 90 % restant étaient le logement (28 %) les restaurants (25 %), le shopping (16 %), carburant (12 %) et les courses de supermarché (9 %) (Stynes, 2005). D'autres exemples de montants conséquents générés indirectement grâce à la nature existent tels que les revenus générés par les parcs nationaux en Allemagne (Mayer et al., 2010) ou encore le rapport de O'Connor et al. (2009) qui évalue les bénéfices mondiaux de l'observation des baleines.

À une échelle plus locale, ces tendances se retrouvent aussi en Wallonie et plus particulièrement en région ardennaise.

Les services écosystémiques récréatifs en Wallonie et en Ardenne belge

En Wallonie, le tourisme représente 4 à 6 % du PIB (Commissariat Général au Tourisme - SPW (CGT), 2017), ce qui représente cinq à dix fois plus que les revenus générés par la production agricole, la production forestière et la chasse (Breyne, 2017). Les attractions payantes génèrent 12 millions d'euros par an, dont 30 % sont en lien avec la nature (CGT, 2017) en grande majorité grâce au parc zoologique Pairi Daiza. Cependant, ces chiffres pourraient encore grandement augmenter si la fréquentation touristique « diffuse », qui n'est pas le fruit d'infrastructures ou d'espaces localisés et qui est donc gratuite était mieux prise en compte (Colson, 2009).

La fréquentation touristique est majoritairement localisée dans le sud et l'est de la Wallonie (donc l'Ardenne), alors que les forêts du nord et nord-ouest sont plus fréquentées par les locaux (Colson, 2009). La forêt wallonne est aujourd'hui considérée par l'état comme multifonctionnelle et donc incluant les activités récréatives (gouvernement wallon, 2004). Généralement, ces mesures politiques sont davantage prévues pour contrôler les activités touristiques en forêt plutôt que pour les développer (Bell et al., 2007).

Des études ont déjà été réalisées en 2017 et 2018 par Gembloux Agro-Bio Tech et l'ASBL Ressources Naturelles Développement (RND) afin de caractériser la fréquentation touristique au sein du massif ardennais (Breyne et al., 2018). Ces travaux s'ajoutent à la thèse de Colson (2009). Des questionnaires en ligne ont été envoyés aux opérateurs touristiques de la région ardennaise transfrontalière (Belgique, Grand-Duché du Luxembourg, France). 598 opérateurs ont répondu dans le but d'évaluer l'importance de la nature de l'offre touristique ardennaise. Il ressort de ce rapport les éléments ci-dessous :

- Une grande majorité des établissements touristiques sont favorables aux activités socioculturelles « douces » dans les espaces naturels ardennais tels que les balades non motorisées, l'observation et la cueillette sauvage des produits de la forêt, les camps de jeunes et les bivouacs. Ils sont aussi pour la réservation d'espaces naturels protégés pour la biodiversité.
- Concernant l'exploitation du bois, les opérateurs touristiques sont aussi favorables à l'abattage des arbres pour le bois de construction ainsi que pour l'affouage. Ils sont également en faveur de l'abattage de résineux qui ne sont pas en station en vue de recréer

Cadre conceptuel

de la diversité dans le paysage. En revanche, une majorité est contre les activités commerciales massives dans les forêts. Il s'agit de l'enrésinement des forêts feuillues, des coupes à blanc avec broyage et nettoyage du sol et des nourrissages en vue de chasses.

- 336 opérateurs ont classifié leurs préférences quant aux services rendus par la forêt ardennaise. Cette question a révélé des préférences majoritaires pour l'esthétisme, la biodiversité ainsi que le cadre de vie qu'elle fournit. Cependant, leur statut d'opérateur touristique rend logiques ces réponses qui sont des atouts marketing. Il faut alors se pencher vers l'avis des visiteurs eux-mêmes pour savoir ce qu'ils viennent vraiment chercher dans la nature ardennaise. Le rapport de 2017 de l'Observatoire du Tourisme Wallon (OTW) (CGT, 2017) rapporte les informations suivantes : 67,8 % des touristes se rendent en Wallonie pour ses richesses naturelles (dont 55 % des visiteurs fréquentant les établissements touristiques viennent en Ardenne pour le patrimoine naturel et paysager (Breyne et al., 2018)).
- 20,5 % des visiteurs y viennent pour se promener à pied ou à vélo et 5,6 % pour y visiter des sites naturels. Selon Colson (2009), le public recherche en Wallonie en majorité le calme (83 %), un circuit en boucle (78 %), des sentiers (36 %) de terre (40 %), des grands arbres (64 %) feuillus (75 %), un relief vallonné (39 %), des cours d'eau (70 %) et des équipements de base (45 %).

II. Les approches d'évaluation des services écosystémiques récréatifs

L'importance d'avoir des indicateurs pertinents concernant l'évaluation des services écosystémiques récréatifs est double : d'une part, ils permettent de comparer les milieux et les différents services écosystémiques entre eux d'une façon purement monétaire, en quantifiant leur potentiel économique. D'autre part, ils permettent de quantifier la pression exercée sur ces milieux, qui peut être mise en relation avec leur état de conservation. Tout d'abord, ces aspects économiques et écologiques sont détaillés. Ces deux premières parties mettent en avant la nécessité de mettre en place des techniques qui rendent compte de la qualité des services écosystémiques récréatifs de façon précise, mais aussi qui soient capables d'emmagasiner un grand éventail d'informations sur de longues périodes. Ces techniques doivent permettre de résumer d'un point de vue récréatif l'état d'un milieu non seulement à un instant précis, mais aussi de manière continue afin de développer un réel cadre de biomonitoring et de combler le manque d'informations à long terme dans ce milieu (Campbell, 2010). Enfin, les différentes techniques existantes sont divulguées, les informations qu'elles permettent d'obtenir ainsi que leurs avantages et leurs inconvénients sont mis en avant.

La notion de « fréquentation touristique » utilisée dans ce chapitre englobe à la fois la fréquentation quantitative de la nature, c'est-à-dire le comptage des personnes, mais aussi la fréquentation qualitative, la façon dont les visiteurs se déplacent dans les forêts et les espaces naturels. Il peut par exemple s'agir de leurs différents moyens de déplacement, du matériel qu'ils emportent ou encore de la configuration de leurs groupes.

1) Aspects économiques découlant des indices de fréquentation touristique

L'avènement du développement durable a apporté avec lui toute une panoplie de nouvelles disciplines scientifiques dont l'économie durable (Baumgärtner et al., 2010). Elle permet d'évaluer les avancements et le progrès en matière de durabilité grâce à des outils économiques (Pintér et al., 2012). Une liste des méthodes d'évaluation de la durabilité sont présentées dans l'étude de Gasparatos et al. (2012). On y retrouve différents indices d'évaluation monétaires tels que (Farber et al., 2002; TEEB, 2010):

- **Les prix des marchés financiers** : méthode la plus universelle pour les biens matériels, les prix reflètent la loi de l'offre et de la demande.
- **La méthode des coups de transport** : lorsque des personnes se déplacent pour visiter un site, elles doivent subir le coût du déplacement jusqu'à cet endroit. Ce coût est donc implicitement attribué comme valeur apportée au site. Il s'agit d'une préférence révélée de la part du visiteur.
- **L'évaluation hédonique** : cette méthode consiste à étudier la différence de valeur attribuée à un bien marchand, lorsque celui-ci présente certaines qualités (ici, environnementales) ou non.
- **L'évaluation contingente** : en présentant un scénario aux enquêtés, cette méthode permet d'évaluer leur consentement à payer ou à recevoir dans une situation donnée. Ces deux façons d'aborder la question donnent souvent des résultats différents en raison d'une multitude de biais.

En combinant ce type d'évaluation économique avec des méthodes de fréquentation touristique prenant en compte le nombre d'individus visitant un espace naturel, il est dès lors possible d'estimer de façon relativement précise la valeur de ce milieu. À terme, la traduction

Cadre conceptuel

des aspects matériels et intangibles d'un espace naturel en valeur monétaire doit permettre aux différents décideurs et gestionnaires de ces espaces de prendre des décisions en bonne connaissance de cause.

2) Aspects écologiques découlant des indices de fréquentation touristique

La quantification des flux touristiques est indispensable dans les cadres conceptuels de gestion actuels (Cessford et al., 2003; Pettebone et al., 2010; Moore, 2012). Parmi ces cadres, on retrouve le Pressure-State-Response Model (OECD, 1993; Hammond et al., 1995) qui permet d'avoir une vision claire sur la stratégie à adopter par rapport aux aspects environnementaux, sociaux et économiques quant à la gestion des ressources naturelles. En 1999, le modèle évolue en DPSIR : Driving forces, Pressure, State, Impact, Response (EEA, 1999). Comme tous les indicateurs, ils sont destinés à rendre la communication claire sur l'état d'un milieu, mettent en évidence les priorités à définir par l'identification de certains facteurs de pressions et permettent d'avoir un suivi quant à l'impact de la réponse mise en place (EEA, 1999) (Figure 3).

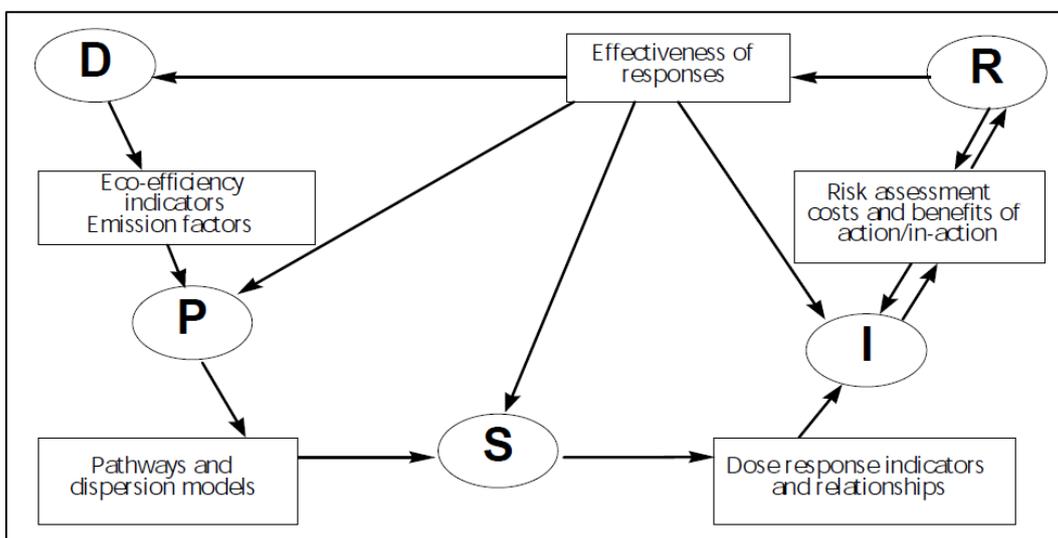


Figure 3: Cadre conceptuel DPSIR : Driving forces, Pressure, State, Impact, Response.

Appliqué aux enjeux du tourisme et de la biodiversité, il permet de diriger, de surveiller et d'évaluer les pressions mises en exergue sur le milieu ainsi que les réponses mises en place. Les événements et les actions (Driving forces, D) exercent des pressions sur le milieu (P, Pressure). Cette pression modifie donc l'état du milieu (State, S) de différentes façons engendrant un impact (Impact, I) plus ou moins marqué sur celui-ci. Cet impact entraîne alors une réponse du milieu (Response, R) dont la dynamique va être modifiée, ce qui peut avoir un effet sur l'ensemble des composantes du système.

3) Les différentes techniques d'évaluation de la fréquentation touristique

Il existe plusieurs techniques de suivi des visiteurs et chacune possède des avantages et des limites. Les paramètres entrant en compte dans le choix de la méthode de suivi sont principalement le prix, le temps disponible, et le type d'information que l'on veut en tirer. Il est aussi possible de combiner différentes méthodes afin de contrebalancer les défauts de l'une avec les avantages de l'autre. Le Tableau 1 montre un aperçu des techniques de suivi des visiteurs disponibles en indiquant le type d'information qu'il est possible d'en obtenir.

Cadre conceptuel

Tableau 1: Panel de techniques de suivi des visiteurs et leurs champs d'applications. Tiré de Muhar et al. (2002).

			visitor numbers	direction of motion	routes	distribution within area	group size	visitor characteristics (age, sex)	visitor characteristics (origin, expectations etc.)	behaviour
direct methods	interviews	oral interviews			x	x	x	x	x	x
		written interviews			x	x	x	x	x	x
	direct observation	roaming observers	(x)	(x)	(x)	(x)	(x)	(x)		(x)
		fixed counting stations	x	x		x	x	x		x
	indirect observation	automatic cameras	x	x		x	x	x		x
		time-lapse video	x	x		x	x	x		x
		erial, satellite imagery	(x)	(x)	(x)	(x)	(x)			
	counting of access permits	tickets sold	x							
		permits issued	x				x			
	counting devices	turnstiles	x	(x)			(x)			
		photoelectric counters	x	(x)			(x)			
		pressure sensitive mats	x	(x)						
		pneumatic tubes	x	(x)						
		inductive loop sensors	(x)	(x)						
	self-registration	trail registers	x	x	(x)		x			
summit books		x	x	x		x				
hut registers		x	x	x		x				
indirect methods	mapping of traces of use	garbage	x			(x)				
		trail deterioration	x			(x)				
		damage to vegetation	(x)			(x)				
		footprints	x	(x)		(x)				
		sandbed	x	(x)						

Aujourd'hui, cette liste s'est grandement allongée et les nouvelles technologies rendent une partie de ces méthodes obsolètes.

a. Les techniques quantitatives

Les comptages sont effectués par des observateurs (Hartmann, 1988; Murphy, 1992; Keul et al., 1997) ou plus récemment par des appareils électroniques (Hornback et al., 1999; Watson et al., 2000; Lynch et al., 2002). Ils peuvent être placés sur les chemins ou bien aux entrées des espaces naturels. S'il n'y a pas d'entrée précise, se placer sur les parkings peut être une alternative (Lynch et al., 2002).

Les appareils électroniques, généralement appelés écompteurs, sont surtout utilisés dans les vastes parcs où il n'est pas possible d'envoyer des personnes chaque jour (Rundle, 2002). Ils permettent d'obtenir un grand nombre d'informations précises sur une longue période de temps dans les endroits les plus reculés (Arnberger et al., 2005). Il en existe différents types. Au Royaume-Uni (Melville et al., 2004) et en Suisse (Rupf et al., 2006), les parcs nationaux peuvent par exemple être équipés de dalles piézométriques. Lorsqu'un individu y applique une pression, l'information est alors directement enregistrée. Des capteurs infrarouges sont aussi utilisés. Il existe aussi d'autres types de capteurs tels que des caméras thermiques² ou les capteurs à boucles inductives pour les objets métalliques comme les vélos (Muhar et al., 2002). Cependant, beaucoup de gestionnaires trouvent les écompteurs plutôt inefficaces (Lynch et al., 2002) car ils ne différencient pas les hommes de la faune (Campbell, 2010). De plus, ils sont plutôt chers (environ 4.000 euros pour un écompteur de la marque du même nom).

² <https://www.eco-compteur.com/produits/citix/citix/>

Cadre conceptuel

Il existe aussi une multitude de méthodes indirectes qui profitent simplement de la structure d'accueil déjà en place. Il s'agit par exemple de l'enregistrement de tickets d'entrée, de permis d'accès, ou d'obligations d'enregistrement sur certains chemins (Lynch et al., 2002; Rundle, 2002; Cessford et al., 2003). L'avantage de ces techniques est le peu de travail qu'elles requièrent pour obtenir des informations exhaustives. Elles sont cependant très rarement disponibles. En effet, une grande minorité d'espaces naturels sont clôturés avec un droit d'accès et généralement ces derniers ont déjà des objectifs touristiques et financiers affirmés qui engendrent l'utilisation de techniques de suivi des visiteurs plus élaborées. Dans la majorité des espaces naturels, ces méthodes peuvent se rabattre sur le comptage indirect des visiteurs grâce aux livres d'or que l'on trouve dans les refuges ou sur les sommets (Bambi et al., 2017). Ces livres peuvent être utiles pour avoir des informations sur les visiteurs passant la nuit dans les réserves, dans les endroits reculés où il y a peu de fréquentation. Cependant, la précision de ces informations est compromise par le fait que les visiteurs ne sont pas obligés de se servir de ces outils.

Les suivis systématiques des usages récréationnels des milieux naturels se font donc rares. Il arrive fréquemment que dans les études liées à la fréquentation des espaces naturels, le comptage des visiteurs soit réalisé durant une courte période et extrapolé. Ces résultats ponctuels sont ensuite impliqués dans la stratégie de gestion des espaces naturels à long terme sans prise en considération de leur réelle pertinence (Muhar et al., 2002).

b. Les techniques qualitatives

De plus en plus d'études se tournent vers une évaluation qualitative de la fréquentation touristique. Cela stimule donc le débat quant au rôle des organismes de recherche dans ces approches méthodologiques en relation avec les opérateurs touristiques (Del Casino et al., 2000). Il est important de connaître le niveau de satisfaction, les préférences et les choix des visiteurs en vue d'une gestion optimale des espaces naturels (Neal et al., 2008; Alegre et al., 2010). De plus, ces informations améliorent la communication entre les gestionnaires et les différents intervenants financiers et politiques (Eagles, 2014). Le niveau de satisfaction des visiteurs donne une information sur le niveau de qualité de l'environnement et sur la qualité des services qui lui sont fournis (Eagles, 2002). Cela peut aussi jouer un rôle dans la survie financière de certaines réserves naturelles (Gursoy et al., 2007). Mieux comprendre les choix que font les visiteurs dans les espaces naturels et leurs préférences doit permettre d'une part d'adapter l'offre à la demande et d'autre part de mieux protéger les zones les plus à risque quant aux perturbations liées à ces visites d'autre part. Cela peut aussi permettre d'adapter certains types de gestion afin de changer le regard des personnes, par exemple sur les travaux de foresterie (Heyman, 2012).

Les interviews et les questionnaires sont d'excellentes techniques qualitatives. Les interviews se font directement in situ, un opérateur interagit directement avec les personnes et peut réorienter les questions si nécessaire. Cela peut aussi rendre les visiteurs plus réticents à répondre certaines questions. Les questionnaires peuvent prendre plusieurs formes. Ils peuvent être délivrés en main propre par un opérateur ou un gestionnaire sur place, être envoyés par e-mail à un public cible ou être disponibles à certains endroits clés d'une réserve naturelle. Par exemple, aux côtés des livres d'or dans les refuges ou sur les sommets (Bambi et al., 2017).

Ces méthodes permettent d'avoir une multitude d'informations précises et sûres sur le comportement des visiteurs. Elles permettent de connaître leur provenance, leur âge, leur situation familiale, leurs revenus, leurs motivations ou encore leur assiduité quant à la visite d'un lieu. Cependant, elles sont intrusives, demandent beaucoup de travail et de temps lors de

Cadre conceptuel

l'élaboration des questions et de l'analyse des réponses. La collecte des données, particulièrement dans le cas des interviews demande aussi beaucoup de temps. Ce qui empêche cette technique d'être durable dans le temps (Kellner et al., 2016).

Il faut aussi différencier les préférences déclarées et révélées. Les préférences déclarées sont celles qui peuvent être exprimées de manière volontaire par une personne. Il est intéressant de compléter cette approche par une analyse des préférences révélées à travers les comportements observés, ce qui peut dévoiler de nouvelles informations que les visiteurs ne souhaitent pas assumer (dégradations, non-respect des règles, ...). Par exemple en zone périurbaine, la présence de chiens sur un tracé dévoile une pression constante des visiteurs sur la zone. Dans ce cas, la visite du milieu ne provient pas de la volonté du visiteur, mais de l'obligation de sortir son chien. De plus, lorsque le chien n'est pas en laisse, cela accroît la pression sur le gibier et donc, la fragmentation du paysage (Arnberger et al., 2003). Les personnes qui marchent hors des sentiers peuvent aussi avoir un impact (Dumont et al., 2005) et pourraient ne pas le déclarer. Or, des mesures de gestion comme l'adaptation de la largeur des sentiers pourrait être envisagée grâce à cette révélation (Wimpey et al., 2010).

De façon plus originale, Heyman (2012) a eu l'idée de réappliquer la technique de (Cherem et al., 1983) qui consiste à donner un appareil photo aux visiteurs (visitor employed photography, VEP) et à leur demander de photographier ce qu'ils ont aimé le plus et ce qui les ont dérangé le plus sur un parcours. Cela permet d'avoir une idée claire de la perception par les usagers des espaces naturels.

c. Les nouvelles technologies

Le tourisme étant l'industrie majeure dans de nombreux pays, il se doit d'être à la pointe de la technologie (Gretzel, 2011). Grâce à celle-ci, une multitude des nouvelles techniques peuvent désormais être employées afin d'évaluer la fréquentation touristique des espaces naturels de façon plus précise, sur de plus longues périodes et sur de plus grands espaces. De plus, ces technologies permettent d'avoir des informations quant au nombre de personne fréquentant un milieu, mais aussi quant à certaines caractéristiques de leur comportement. On peut citer parmi les nouvelles technologies celles liées aux systèmes de communication (Information and Communication Technology, ICT), aux informations géographiques (SIG) ou encore celles liées à l'acquisition d'images par pièges caméras.

Technologies d'information et de communication

Concernant les ICT, les différents opérateurs mobiles possèdent d'énormes quantités d'informations concernant le positionnement de nos téléphones (Debusschere et al., 2017). Il est possible de connaître le nombre de téléphones portables dans une zone, et même la provenance de ceux-ci (Ahas et al., 2008). La provenance des usagers est très importante dans le contexte de globalisation actuel. Ces techniques peuvent paraître discutables en ce qui concerne le respect de la vie privée d'autrui, ce qui peut poser problème lorsque la donnée provient du signal Wi-Fi ou Bluetooth de processeur du téléphone (Abedi et al., 2014) ou du « network probing system » (système d'exploitation du réseau, qui capte chaque activité du signal) (Debusschere et al., 2017). Jusqu'alors, les enregistrements d'appels détaillés (CDR, call detail records) étaient utilisés dans les études de recensement classiques (European Commission, 2014). Cependant, il est aussi possible de tracer les appareils via l'adresse MAC (Media Access Control) propre à chaque téléphone qui peuvent être tracée le long d'un itinéraire. De ce fait, les informations concernant le numéro de téléphone et l'identité de la personne ne sont pas consultables (Abedi et al., 2014). Grâce à cette adresse, il est alors possible d'obtenir des

Cadre conceptuel

données quant au nombre de personnes dans une zone, leur provenance, ainsi que par rapport à la durée de leur fréquentation. Il est également possible d'obtenir une carte révélant les déplacements des téléphones.

Systèmes d'information géographique

Les techniques de suivi par GPS sont de plus en plus employées (Shoval et al., 2007). Elles fournissent des données très fiables et peuvent être utilisées simultanément sur plusieurs sites. Le suivi des personnes par GPS permet d'avoir une information en temps réel sur leurs déplacements. Il est donc possible d'acquérir de précieuses informations quant à la durée de la visite, ou à l'attrait favorable de certains sentiers et points de vue par rapport à d'autres (Wolf et al., 2012). Des cartes de fréquentation peuvent alors être générées à large échelle. Avant l'arrivée des GPS, ces cartes étaient générées soit en demandant aux visiteurs leur itinéraire, ce qui nécessitait qu'ils s'en souviennent, soit en fonction d'indices d'observations indirectes telles que des empreintes ou des déchets sur les chemins (Shoval et al., 2007). Cependant, cette technique est invasive puisqu'elle nécessite d'approcher les personnes afin de leur proposer de porter un récepteur GPS pour leur randonnée (Kellner et al., 2016). Cela peut fausser le comportement des personnes qui savent qu'elles sont traquées (elles pourraient par exemple éviter de quitter les sentiers). De plus, cette technique ne peut être employée que sur une courte période de temps et pour un petit nombre de personnes. Elle est donc intéressante pour une cartographie d'espace naturel, mais ne constitue pas une méthode de suivi à long terme. De plus, le prix des récepteurs est assez conséquent (Wolf et al., 2012), mais ils peuvent être loués (Kellner et al., 2016).

Acquisition d'images

L'enregistrement de vidéos et la prise de photographies par caméra trap sont utilisées depuis plus longtemps (Watson et al., 2000; Arnberger et al., 2003, 2005; Kammler et al., 2004; Guillén et al., 2008). Pour un prix très modeste (en dessous de 100 euros par unité pour l'entrée de gamme), les caméras placées adéquatement sur les chemins peuvent fournir un très grand nombre d'informations à la fois quantitatives, mais aussi qualitatives (Arnberger et al., 2005; Campbell, 2010). Ce genre de dispositif est très modulable et peut être adapté en fonction des besoins. De plus, compte tenu du faible entretien qu'il nécessite, il peut être combiné à des études qualitatives plus approfondies grâce aux interviews ou aux autres techniques de suivi d'itinéraires. Cette technique est décrite plus en détail au point suivant.

III. Camera trapping et Automatisation

Comme discuté au chapitre précédent, les pièges caméras appliqués à l'étude de la fréquentation touristique ont beaucoup de potentiel. Dans ce chapitre, le fonctionnement des pièges caméras est d'abord explicité. Ensuite, les avantages et les inconvénients liés à ce type de prise de données sont détaillés. Dans la liste des inconvénients qui sont liés à cet outil, l'un des plus réfrénant est la quantité de travail nécessaire pour traiter manuellement les données recueillies. Les nouvelles technologies liées à la vision par ordinateur peuvent être la solution à ce problème. En partant de la définition de l'intelligence artificielle, le cheminement est détaillé jusqu'à la présentation des réseaux neuronaux convolutionnels, groupe d'algorithmes les plus efficaces à ce jour pour détecter, segmenter et classifier des objets sur des images.

1) Définition, histoire et fonctionnement des camera traps

L'utilisation des pièges caméra (Camera trapping) consiste à l'origine à placer un appareil photo dans la nature pour capturer des images de la faune sans l'intervention d'un opérateur (Swann et al., 2010). Le concept s'étend de manière plus large, avec notamment la possibilité de faire des vidéos ou des timelapses (Weathers et al., 1998).

L'idée vient à l'origine du photographe américain Georges Shiras à la fin du XIXème siècle. À la base chasseur, il a décidé de vouer sa vie à la protection des animaux en troquant son fusil contre un appareil photo. Afin d'obtenir des clichés d'animaux sauvages durant la nuit, il a inventé un système de câbles grâce auquel lorsque l'animal tirait sur l'un d'eux, il déclenchait le flash et prenait le cliché (Shiras, 1935).

Petit à petit, les méthodes de déclenchement ont évolué. Depuis les années 90, le déclenchement de la caméra se fait généralement grâce à un capteur infrarouge qui détecte les différences de température entre les surfaces et l'arrière-plan (Swann et al., 2004; Meek et al., 2014). La caméra se déclenche donc lors d'un changement de température subit. Le déclenchement peut également se faire la nuit, grâce à un flash ou à des diodes électroluminescentes afin d'obtenir des photos en noir et blanc. Cette dernière caractéristique permet d'obtenir des clichés sans déranger la quiétude de la cible (Séquin et al., 2003; Wegge et al., 2004). Les photos sont enregistrées sur des cartes SD, comme pour les appareils photo classiques.

En ce qui concerne les éléments paramétrables, hormis la date et l'heure, il est possible de modifier la résolution des photographies que l'on veut obtenir (en vue d'un gain de mémoire par exemple). Il est aussi possible de modifier le nombre de captures prises lors d'un seul événement de détection et de modifier l'intervalle de temps minimum entre ceux-ci pour éviter les images redondantes.

2) **Avantages et limites des camera traps**

a. Avantages

Grâce à l'utilisation des camera traps, il est possible d'accumuler des données pendant une durée conséquente (plusieurs semaines à plusieurs mois) et 24 heures/24, 7 jours/7. De plus, hormis lors de la mise en place et du relevé des images, la demande en main d'œuvre sur le terrain est inexistante. Applicable aussi de nuit, elle permet de capturer des sujets particulièrement rares et/ou discrètes (Vine et al., 2009). Il est également possible d'avoir des informations sur le comportement des espèces en face de la caméra, mais aussi des déplacements de celles-ci lorsque les individus ou les groupes d'individus sont identifiables. Il est dès lors possible de modéliser les dynamiques de déplacements des individus ou des populations (Trolliet et al., 2014). Le prix de la technique est aussi très avantageux, surtout pour des études destinées à être menées sur de longues périodes de temps (Roberts, 2011).

b. Limites

En revanche, la méthode présente aussi son lot de défauts qui ont été maintes fois discutés depuis la fin du XXème siècle (Rice et al., 1995; Peterson et al., 1998; Cutler et al., 1999) . Selon Gomez Villa et al. (2017), il y a trois challenges : les conditions environnementales, le comportement du sujet et les limitations liées au matériel.

Les conditions environnementales et la longévité

Les caméras doivent être résistantes aux intempéries. Elles doivent donc être étanches et résister au froid (Griffiths et al., 1993; Cronin, 2010). La neige, la pluie et le brouillard peuvent rendre les photos floues lorsque l'eau se dépose sur la lentille (Kays et al., 2009; Rowcliffe et al., 2011). À l'inverse, sur certains modèles, la sensibilité du détecteur peut diminuer lorsque la température augmente (Swann et al., 2004). Puisque l'un des avantages de la méthode est qu'elle peut couvrir une longue période, il faut aussi pouvoir couvrir les besoins en énergie du matériel. Ceci est possible avec des panneaux solaires, mais la plupart du temps, entre 8 et 12 piles alcalines AA sont utilisées (Wearn et al., 2017).

Les limitations techniques liées au matériel

Il existe aussi le risque de faux positifs et de faux négatifs. Le faux négatif désigne le fait que l'appareil ne prenne pas de photo alors qu'un sujet passe devant celle-ci. Cela semble être dû davantage au temps de latence (Campbell, 2006), à la position de la caméra ainsi qu'à la taille et la rapidité du sujet plutôt qu'à la sensibilité du détecteur (Swann et al., 2004). Le faux positif, beaucoup plus fréquent (Swann et al., 2004), désigne un déclenchement non désiré de l'appareil photo. Il peut être dû à des branches qui passent dans le champ du détecteur de mouvement à cause du vent (Towerton et al., 2008) ou encore au changement d'orientation de la lumière au cours de la journée, particulièrement lorsque les rayons lumineux sont orientés vers l'appareil photo (Fairfax et al., 2014). Il est donc très important de prendre en compte ces paramètres lors de la mise en place du matériel et d'effectuer tout une batterie de tests avant de le laisser en pleine nature au risque de découvrir une carte mémoire vide ou pleine d'images sans intérêt lors du relevé des données.

Cadre conceptuel

Le risque de vol et de dégradation

Il faut enfin noter le risque de vol et de dégradation du matériel. Pour contrer cela, certains scientifiques mettent en place des moyens de protection, comme des cadenas ou même des cages protégeant les caméras (Meek et al., 2013). Il est aussi possible de les placer en hauteur si l'étude le permet.

Traitement des images

Le problème du traitement des données se pose également. En effet, la quantité de photos prise est généralement de l'ordre de dizaines voire de centaines de milliers de photos dans des études de plusieurs mois. Ces données doivent pouvoir être stockées et analysées (Gomez Villa et al., 2017). Le stockage étant uniquement dépendant des ressources informatiques, l'analyse quant à elle requiert un temps conséquent puisqu'elle est en général faite manuellement. Elle demande aussi de l'expertise notamment dans le cas d'études de faune afin de pouvoir classer les espèces sur les photos. Il faut ajouter à cela que les individus à détecter sont rarement positionnés de façon optimale devant la caméra et qu'il peut y avoir un certain nombre de facteurs qui rendent cette classification difficile, voire impossible.

L'une des solutions les plus utilisées pour palier au temps conséquent que nécessite le traitement des images est la technique des sciences citoyennes. Cette technique consiste à recruter des volontaires, de façon physique ou via une plate-forme en ligne, afin de diviser la charge de travail par le nombre de personnes. Le projet Snapshot Serengeti se sert de cette technique pour analyser les photos recueillies par 250 caméras placées dans la savane africaine (Swanson et al., 2015). Le projet eMammal utilise aussi les sciences citoyennes, mais d'une façon différente : ils permettent aux volontaires d'uploader leurs photos issues de piège caméra sur leur site afin de mettre toutes les informations en commun afin de créer une grande base de données sur les mammifères et de bénéficier de l'expertise de chacun (Chen et al., 2014). Le gain de temps est réel. De plus, cela permet de sensibiliser les gens à certaines thématiques et/ou de s'impliquer dans des projets scientifiques qui leur tiennent à cœur. Cependant, ces personnes n'ont pas d'expertise, ce qui pourrait biaiser les résultats. Il est donc nécessaire d'investir un certain temps dans la formation de celles-ci et d'en tenir compte.

3) **Caméra trapping et tourisme**

La méthode du comptage par caméra trap afin de quantifier la fréquentation touristique des milieux naturels est encore aujourd'hui très peu utilisée. Cependant, placer des caméras aux entrées de réserves naturelles, sur des chemins de randonnée ou encore aux abords de points de vue peut donner des informations précieuses sur la fréquentation de ceux-ci pour un coût en argent et en temps relativement faible. Comme cité plus haut, la principale raison à cela est le travail de traitement des données en aval qui demande un temps considérable. Certains fournissent l'effort nécessaire (Conlon, 2014; Fairfax et al., 2014; Bambi et al., 2017) et d'autres contournent cet aspect. Ils peuvent choisir de se tourner vers les sciences citoyennes, ou de se cantonner au comptage du nombre d'images capturées et faire des extrapolations quant à la fréquentation du site (Lupp et al., 2016). Cette dernière méthode est comparée à l'indice développé dans le cadre de ce travail (Voir Comparaison avec la méthode de Lupp et al. (2016)).

Les informations obtenues par cette technique demeurent plus précises que des sessions de comptages sporadiques et des estimations grossières (Lupp et al., 2016). Ce qui permet de gérer de façons plus adéquates les infrastructures pour des utilisateurs spécifiques (marcheurs nordiques, skieurs de fonds, ...) ou indésirables (chiens sans laisse, véhicules motorisés) (Arnberger, 2006).

Les sciences citoyennes sont surtout employées pour la détection et la classification d'animaux sur les photos, l'équivalent avec des personnes serait sans doute moins attrayant et il pourrait être nécessaire de fournir une petite contrepartie pour motiver les volontaires. De plus, la question du respect de la vie privée se pose (voir point suivant).

En résumé, les caméras peuvent permettre d'acquérir des données pendant de longues périodes de temps, pour un prix modeste. Elles fournissent un grand éventail d'informations sur le comportement révélé des personnes. Les seules informations qui pourraient manquer sont liées aux motivations et à la provenance des personnes ainsi que par rapport à leurs déplacements. Ces deux lacunes pourraient être résolues par les ICT. Les défauts majeurs de la méthode sont les problèmes techniques récurrents, le respect de la vie privée, et le post traitement des données. Ces deux derniers points sont abordés ci-dessous.

a. Camera trapping et respect de la vie privée

Contrairement au monitoring de la faune, la méthode du camera trapping appliquée à des personnes doit respecter la loi quant au respect de la vie privée et plus précisément la loi caméra du 21 mars 2007³, mise à jour le 31 mai 2018⁴ en raison du règlement européen n° 2016/679 (Règlement Général sur la Protection des Données, RGPD)⁵ et applicable à partir du 25 mai 2018.

Le règlement européen, contrairement à la directive européenne, est applicable dans les 20 jours de son entrée en vigueur de façon obligatoire par les états et sans aucune transcriptions nationales. Le règlement européen n°2016/679 vise à consolider les droits des personnes quant à leurs données personnelles, mais aussi à renforcer le pouvoir de l'Europe quant à ces données.

Pour se faire, les personnes doivent donner leur consentement et l'opérateur se doit d'informer ces individus de la collecte des données les concernant. Il faut aussi communiquer pourquoi ces données sont collectées et pour combien de temps. Ces informations doivent être stockées de manière documentée et sécurisée suivant le niveau d'importance de celles-ci. Les personnes sont supposées avoir accès aux données que l'on traite par rapport à elles et pouvoir les rectifier ou les supprimer si elles le désirent. Elles peuvent aussi ordonner que leurs informations soient uniquement stockées sans être traitées, particulièrement de façon automatisée, par un algorithme. Si ces informations sont supprimées, modifiées ou piratées, l'opérateur doit notifier cet événement à l'individu. L'opérateur doit quant à lui conserver un registre du traitement des données ainsi qu'une preuve du consentement de la personne. Enfin, les données acquises doivent l'être de façon absolue, il faut pouvoir les tracer, connaître leur nature et savoir qui peut les consulter de façon précise. Lorsqu'elles ne sont plus nécessaires, celles-ci sont supprimées (droit à l'oubli). Il faut aussi pouvoir signaler une faille dans le système dans les 72 heures et nommer un prestataire en charge de ces données privées.

Concernant la loi caméra, avant 2007, toute personne qui installait une caméra devait simplement respecter la législation en matière de vie privée. La loi caméra est apparue lorsque l'on s'est rendu compte qu'une loi spécifique à cette technologie était nécessaire. Elle porte sur les modalités d'installation et d'utilisation des caméras tout en respectant le RGPD (depuis 2018). Cette loi concerne principalement les caméras de surveillance. Cependant, dans cette loi il est stipulé qu'une caméra de surveillance a pour finalité de prévenir, constater ou déceler des incivilités et/ou des infractions et de maintenir l'ordre public. Les camera traps appliquées à des fins de monitoring ne tombent donc pas sous cette dénomination et doivent simplement suivre les règles du RGPD.

³ 21 MARS 2007. - Loi réglant l'installation et l'utilisation de caméras de surveillance. - http://www.ejustice.just.fgov.be/cgi_loi/change_lg.pl?language=fr&la=F&cn=2007032139&table_name=loi

⁴ 21 MARS 2018. - Loi modifiant la loi sur la fonction de police, en vue de régler l'utilisation de caméras par les services de police, et modifiant la loi du 21 mars 2007 réglant l'installation et l'utilisation de caméras de surveillance, la loi du 30 novembre 1998 organique des services de renseignement et de sécurité et la loi du 2 octobre 2017 réglementant la sécurité privée et particulière (1) - http://www.ejustice.just.fgov.be/cgi/article_body.pl?language=fr&caller=summary&pub_date=2018-04-16&numac=2018011626

⁵ Règlement (UE) 2016/ 679 du parlement européen et du conseil - du 27 avril 2016 - relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/ ce (règlement général sur la protection des données) - <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:32016R0679&from=FR>

4) Réseaux neuronaux convolutionnels et automatisation

Jusqu'à récemment, la démocratisation des techniques d'analyse d'images par ordinateur était principalement freinée par la quantité de données nécessitant une puissance informatique qui n'était pas disponible pour le grand public (Debuschere et al., 2017). L'essor des technologies liées au Big Data de cette dernière décennie a permis de s'y réintéresser de manière plus efficace. Désormais, possible d'automatiser la détection et la classification de sujets sur des images prises par camera traps. Cette délivrance permet de se libérer de la contrainte liée au temps que cela nécessite lorsque les images sont analysées manuellement comme dans la plupart des études actuelles (Conlon, 2014; Bambi et al., 2017). Cependant, faute d'avoir trouvé la technique adaptée à l'ensemble des problèmes de ce type, de nombreux scientifiques se relaient dans l'amélioration de cette méthode (Yu et al., 2013).

Selon LeCun et al. (2015), les termes « deep learning » et « machine learning » sont souvent confondus. Le premier est en fait un sous-ensemble du second.

Dans les algorithmes classiques, on insère des paramètres pour obtenir une réponse en fonction des règles encodées. Dans le machine learning, on insère les paramètres et la réponse, l'algorithme en déduit alors les règles qu'il pourrait réappliquer par la suite. Le deep learning ne gère plus une base de données avec des attributs bien spécifiques en entrée du modèle, mais des données qui vont être insérées dans des réseaux de neurones artificiels, inspirés des réseaux de neurones biologiques. L'algorithme se sert de ces données brutes pour créer lui-même de nouveaux attributs qui lui conviennent mieux pour s'entraîner et créer des résultats probants (Figure 4).

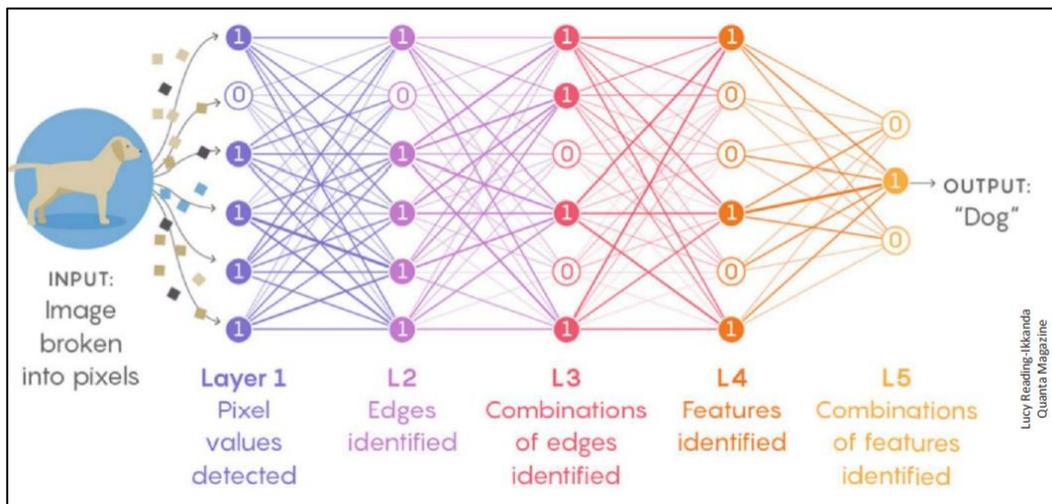


Figure 4: Perceptron multicouche. Exemple de l'architecture de base d'un réseau neuronal artificiel (source: web)

a. Réseaux neuronaux convolutionnels (CNN)

Les scientifiques se sont finalement rendu compte que pour étudier des données sous forme d'images, de vidéos, de sons ou de textes, les résultats étaient trop dépendants de l'architecture du perceptron multicouche (LeCun et al., 2015). Les réseaux neuronaux convolutionnels ont alors été créés. Ils permettent de faciliter l'entraînement du modèle et d'obtenir des résultats plus généralisables principalement grâce au principe de convolution (Voulodimos et al., 2018). L'idée principale du réseau neuronal convolutionnel est qu'il est plus simple de trouver des correspondances entre des morceaux d'images qu'entre des images entières. Un réseau neuronal convolutionnel fonctionne par alternance de trois types de couches.

Tout d'abord, les couches convolutionnelles permettent de résumer l'information présente sur l'image. Grâce à une série de filtres propres à l'entraînement du modèle, des cartes de correspondance entre l'image et les filtres sont créées (la taille de l'image est donc réduite proportionnellement à la taille du filtre).

Deuxièmement, une couche de normalisation modifie les valeurs en sortie, afin de faciliter les opérations mathématiques. Dans la Figure 5, cette couche est appelée « Rectification Linear Unit » et arrondi simplement à zéro toutes les valeurs négatives.

Ensuite, des couches d'agrégation (« Pooling ») permettent de réduire la quantité d'information à traiter par le modèle. Elle permet aussi d'éviter le surajustement du modèle qui serait à des effets de position du sujet dans l'image (Voulodimos et al., 2018). Cela crée donc une invariance compte tenu de l'échelle et de la distorsion entre les images (LeCun et al., 1990).

Enfin, ces cartes de correspondance passent dans des couches entièrement connectées, typique des réseaux neuronaux classiques, qui vont assigner une probabilité d'appartenance à l'une des classes en fonction de la valeur de chaque pixel.

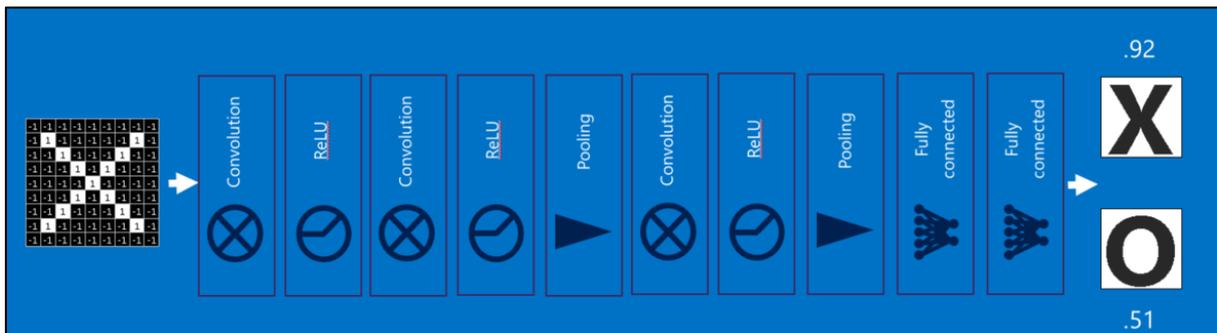


Figure 5: Exemple d'architecture d'un réseau neuronal convolutionnel. Ici, le réseau doit classifier l'image en entrée dans l'une des deux catégories : « X » ou « O ». Les couches de convolution, d'activation (ReLU) et de Pooling s'alternent avant de passer à la phase de classification par les couches entièrement connectées.

Les réseaux neuronaux artificiels, convolutionnels ainsi que les différentes méthodes d'apprentissage sont décrits plus en détails dans l'Annexe 1.

Matériel et méthode

En se basant sur l'utilisation des camera traps, l'objectif principal de ce travail est de faire un premier pas vers la création d'un indice de fréquentation touristique qui prenne en compte l'ensemble des données qui sont récoltées. De cette façon, la précision des résultats devrait être supérieure à celle obtenue grâce aux techniques étudiant un sous-échantillon des données et extrapolant les résultats à l'ensemble. Afin d'améliorer la gestion économique et écologique des espaces naturels, l'indice doit pouvoir être applicable à long terme et sa mise en place doit être bon marché. Il doit aussi donner des informations à la fois quantitatives (nombre de personnes), mais aussi qualitatives (comportement des visiteurs). Le succès et la propagation de cet indice reposent principalement sur la résolution du défaut majeur lié à cette technique : le traitement d'énormes quantités de données. Grâce aux réseaux neuronaux convolutionnels, le travail d'analyse des photos doit pouvoir être fait par un algorithme afin d'alléger grandement la charge de travail pour les gestionnaires qui devrait se contenter d'analyser des tableaux et graphiques de résultats en sortie de la chaîne de traitement.

Afin d'explorer les différentes pistes en vue de l'élaboration de cet indice, les données photographiques du projet AGRETA sont utilisées. La présentation du projet AGRETA ainsi que la configuration spatiotemporelle de l'étude est décrite ci-dessous. Ensuite, une vue globale de la chaîne de traitement est présentée. Sachant que ce travail est une première ébauche, une chaîne de traitement optimisée est décrite parallèlement à celle suivie dans le cadre de ce document. Les méthodes d'acquisition et de stockage des données sont aussi expliquées. La partie suivante est consacrée à l'explication du fonctionnement de l'architecture Mask R-CNN, qui est l'algorithme utilisé. Il permet à la fois détecter et segmenter les objets sur les images, mais aussi de leur assigner une classe précise. Enfin, l'évaluation de la qualité des résultats obtenus et les modalités de création de la base de données formée grâce à ceux-ci sont détaillés.

La dernière partie de ce chapitre décrit un exemple de cas d'étude d'utilisation de l'indice pour répondre à différentes questions de fréquentation touristique dans le cadre du projet AGRETA.

I. Le projet AGRETA

Le projet AGRETA (Ardenne Grande Région, Eco-Tourisme et Attractivité) se concentre sur les provinces de Liège et du Luxembourg en Belgique et sur les Ardennes luxembourgeoises au Grand-Duché du Luxembourg. Le projet est prévu pour une période de quatre ans, entre le 1^{er} janvier 2017 et le 31 décembre 2020. Il vise à développer l'attractivité transfrontalière sur le massif ardennais tout en améliorant sa compétitivité économique et la gestion durable de ses ressources. La stratégie adoptée par AGRETA se base principalement sur l'écotourisme comme moteur économique en adoptant un marketing avec une vision centrée sur l'innovation et les réseaux sociaux⁶. Le projet AGRETA comporte sept actions concernant entre autres le marketing, la communication et la sensibilisation. La faculté de Gembloux Agro-Bio Tech, en partenariat avec l'INRA de Nancy, entre en jeu dans l'action trois, qui vise à diagnostiquer et évaluer l'attractivité de l'Ardenne par rapport au développement durable et au tourisme vert.



Grâce à plusieurs campagnes d'interviews sur site et de questionnaires par mail ainsi qu'aux travaux de Colson (2009), les attentes des visiteurs ont été ciblées. Il faut désormais mesurer l'usage récréatif de différents sites phares de l'Ardenne belge afin de pouvoir enfin quantifier la valeur des espaces naturels ardennais.

Pourquoi l'Ardenne ?

Internationalement connue, la région de l'Ardenne est située à la frontière entre quatre pays : la Belgique, l'Allemagne, la France et le Grand-Duché du Luxembourg. La majorité de la superficie de l'Ardenne se trouve au sud-est de la Belgique, en Wallonie (Figure 6).

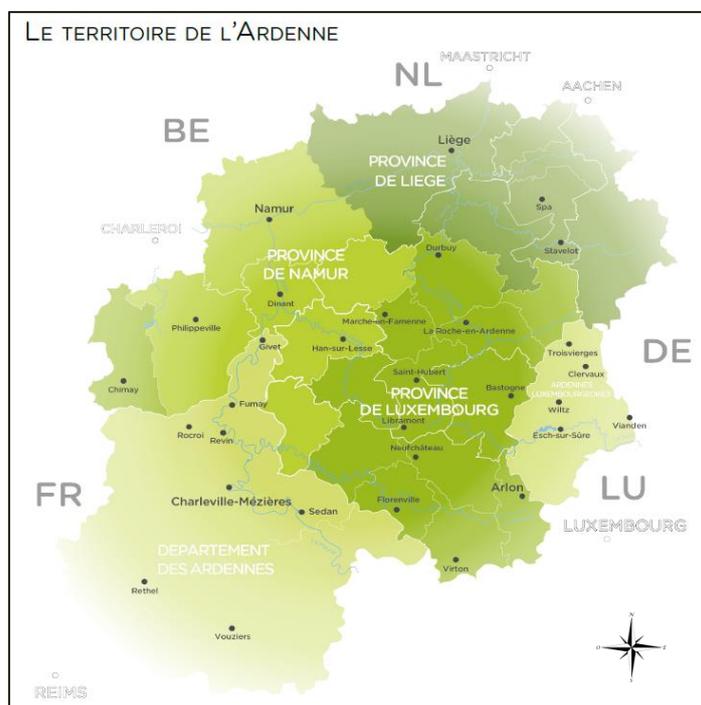


Figure 6: Territoire couvert par la région ardennaise transfrontalière (GEIE Destination Ardenne 2018)

⁶ <http://interreg.visitardenne.com/index.php/fr/agreta>

Matériel et méthode

Le paysage de l'Ardenne est principalement composé de forêts denses : il abrite 60 % des forêts wallonnes. Ces forêts recouvrent en effet 58 % du territoire ardennais belge (Inventaire Permanent des Ressources Forestières de Wallonie). Son altitude atteint 694 mètres et est aussi le point culminant de Belgique, situé au lieu-dit Signal de Botrange. Il se trouve dans le massif des Hautes Fagnes, réserve naturelle depuis 1957 composée en grande partie de tourbières.

L'Ardenne transfrontalière est une région de choix pour ce type de financement, sachant que l'Europe connaît depuis plusieurs années des difficultés socio-économiques principalement dans les zones rurales. Avec l'évolution toujours croissante du tourisme mondial de masse, l'impact de plus en plus compétitif des pays émergents dans le secteur et les facilités de transport actuelles (Esu et Ebitu, 2010), il est indispensable de mettre en avant les atouts de l'Ardenne qui est un point central de l'Europe avec 11 millions de personnes se situant à moins de 4 heures 30 de celle-ci, tous moyens de transport confondus (GEIE Destination Ardenne, 2017).

1) Contexte spatial de l'étude

L'étude de la fréquentation touristique se focalise principalement sur quatre massifs, tous situés en Ardenne belge. Il s'agit de la réserve naturelle des Hautes Fagnes- Eifel, de la grande forêt d'Anlier, de la grande forêt de Saint-Hubert et du parc naturel des Deux Ourthes (Figure 7). Ces sites ont été désignés dans le cadre du projet général et par rapport à leur attrait pour le grand public. De plus, ils sont répartis uniformément sur la région et comportent une densité élevée de chemins de randonnées. On y retrouve également des faciès paysagers bien différents, ce qui pourrait permettre d'observer des comportements préférentiels pour des types de paysages bien particuliers. Cinq emplacements ont été désignés sur chaque massif. Ils ont été choisis en fonction de la proximité des points d'intérêts et de la fréquentation des chemins grâce à l'expertise des gestionnaires des parcs naturels et du Département de la Nature et des Forêts (DNF).



Figure 7: Les quatre massifs répartis en Ardenne belge sur lesquels sont disposés les 20 sites d'étude.

a. La réserve naturelle des Hautes Fagnes

Le plateau des Hautes Fagnes, paysage emblématique de la région ardennaise est, en plus de ses forêts, est un endroit significatif d'un point de vue économique. La réserve naturelle (Figure 8) s'étend sur 4.500 hectares, ce qui en fait la plus grande du pays et attire environ 200.000 écotouristes chaque année (Plunus et al., 2012). Hormis grâce à ses vastes tourbières, les Hautes Fagnes rapportent de l'argent grâce à la chasse et à la production de bois d'épicéa (*Picea abies* (L.) H. Karst). En effet, en Hautes Fagnes, les plantations de résineux ont pris le contrôle du paysage avec de plus en plus de vigueur. À la fin du XVIIIe siècle, les peuplements résineux composent un tiers du massif. Pour pouvoir implanter ces essences, il fallait creuser des drains afin d'assécher ces milieux naturellement humides. Déjà dans les années 60, les écologues étaient conscients des déséquilibres que cela causait sur l'environnement (Froment, 1968). Aujourd'hui, dans une logique de résilience et de durabilité, un projet de restauration de ces milieux humides sensibles a été mené grâce au métaprojet LIFE (regroupement de six projets LIFE à l'échelle du massif ardennais belge, visant à restaurer et reconnecter les massifs tourbeux ardennais avec pour résultat la création de 2.700 ha de nouvelles zones ayant un statut de protection de longue durée, ce qui fait grimper la surface restaurée à plus de 4.500 ha au lieu des 2.800 ha escomptés (Portail Biodiversité Wallonie (SPW), 2014).

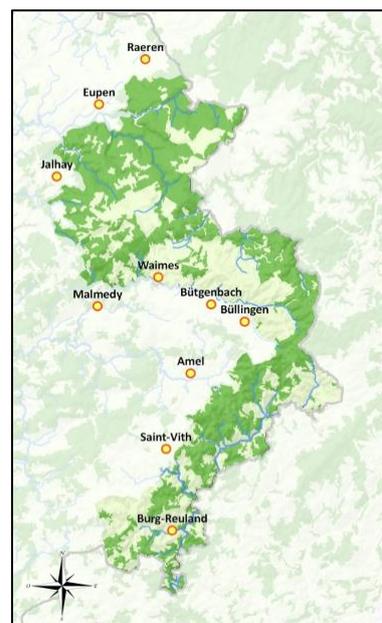


Figure 8: Parc Naturel Hautes Fagnes-Fifel

b. Le parc naturel des Deux Ourthes

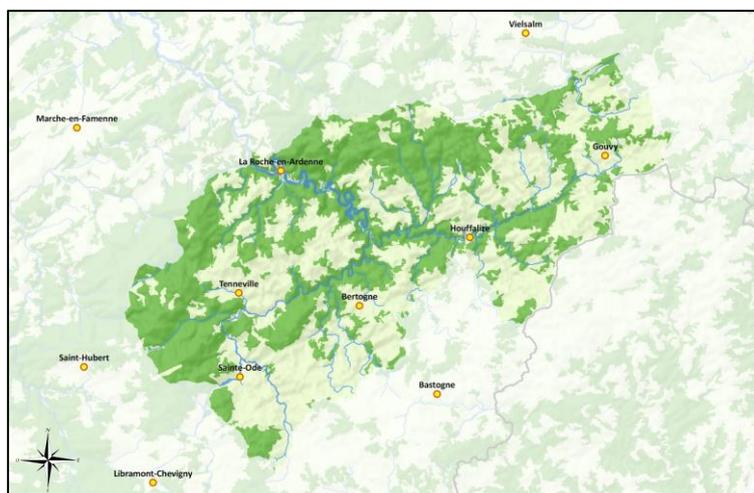


Figure 9: Parc Naturel des Deux Ourthes

Le parc naturel des Deux Ourthes (Figure 9) doit son nom à sa situation. L'Ourthe orientale prenant sa source dans la commune de Gouvy, se joint à l'Ourthe occidentale qui naît à Sainte-Ode. Cette fusion se situe au sein du parc naturel des Deux-Ourthes sur le site d'Hérou et se prolonge alors vers la Roche-en-Ardenne. Le parc se situe également sur les roches les plus anciennes de Belgique composées de schiste rhénan. L'eau et la pierre forment ainsi un

Matériel et méthode

paysage vallonné composé de hauts plateau au nord et à l'ouest du parc, oscillant autour des 600 mètres d'altitude. On y rencontre des milieux tourbeux et des landes. Les parties les moins élevées sont composées de forêts profondes et alluviales.

Créé en 2001, le parc s'étend sur une superficie de 76.000 ha et montre une volonté de développement économique et de sensibilisation au patrimoine local. Le tout de façon durable. L'attrait du paysage pour les touristes est d'une importance capitale pour le parc. La restauration de différents habitats tels que les tourbières, les fonds de vallées et les zones humides y sont réalisées.

c. La Grande Forêt de Saint-Hubert



Figure 10: La Grande Forêt de Saint-Hubert

La Grande Forêt de Saint-hubert (Figure 10) présente des paysages très variés. Elle concilie la chasse, le tourisme et la production de bois. Ses forêts mixtes en font une grande réserve de biodiversité. D'une superficie de plus de 100.000 hectares dont plus de la moitié sont forestiers, la forêt de Saint-Hubert présente plus de 1.500 km de chemins balisés

C'est l'endroit le plus giboyeux de Belgique. Il est idéal pour l'observation de la faune ardennaise et plus particulièrement du cerf élaphe (*Cervus elaphus*). Des milliers de visiteurs y viennent chaque année pour écouter le brame du cerf. Plusieurs sites d'observations destinés au grand public sont disposés dans les zones les plus denses.

d. La Grande Forêt d'Anlier

Constituée d'un bloc compact de 7.000 ha de forêt feuillue et culminant à 517 mètres, la Grande Forêt d'Anlier (Figure 11) abrite la plus grande hêtraie de Belgique. Son couvert continu permet d'abriter une large gamme de mammifères ardennais. Forêt située à la pointe sud-est de la Belgique, elle abrite une diversité remarquable et arbore pour emblème la cigogne noire (*Ciconia nigra*) qui y nidifie. Ce massif dispose néanmoins de très peu d'infrastructures comparé aux trois autres.

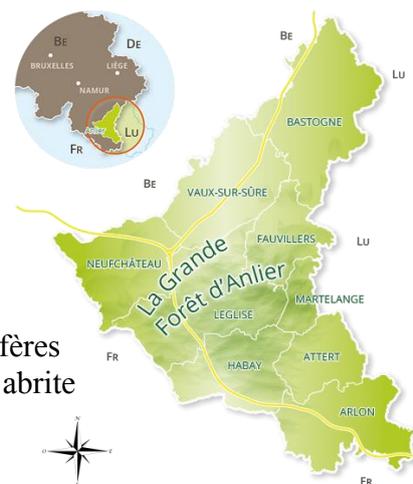


Figure 11: La Grande Forêt

II. Chaîne de traitement

Le développement de l'indice de fréquentation touristique nécessite de prendre en compte l'entièreté des étapes de la chaîne de traitement des données. À partir du placement et de la configuration des caméras, jusqu'à l'obtention de résultats sous forme de graphiques facilement interprétables. Chaque partie doit être optimisée afin de créer une information sûre, précise et compréhensible pour la personne qui souhaite l'obtenir, tout en minimisant la charge de travail qu'elle doit y consacrer. La Figure 12 résume cette chaîne de traitement. La partie supérieure désigne le processus tel qu'il est réalisé dans le cadre de ce travail, la partie inférieure représente une chaîne de traitement optimisée, vers laquelle il faut tendre. Chaque étape est développée dans les différents points ci-après et en Annexe 2.

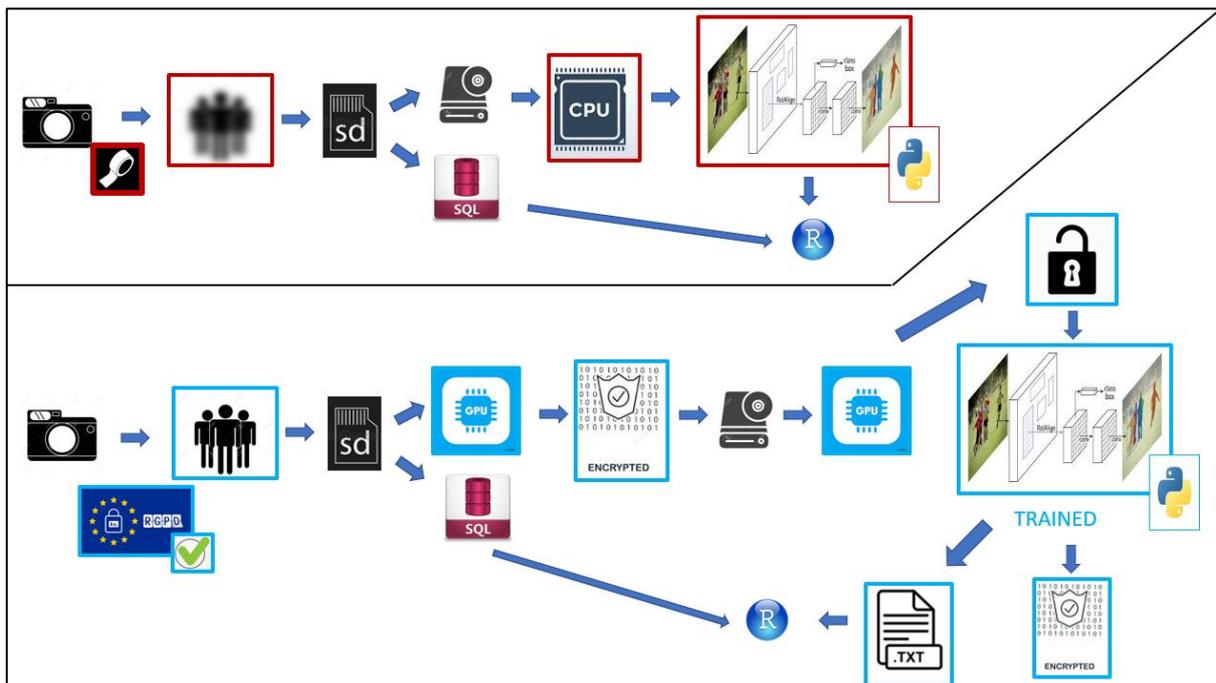


Figure 12: Workflow représentant la chaîne de traitement des données, de l'acquisition des images à l'analyse des résultats. La partie supérieure représente le processus tel qu'il est réalisé dans le cadre de ce travail. La partie inférieure représente le processus optimal qui doit être visé. Les pictogrammes encadrés en rouge sont les étapes qui devraient être remplacées ou supprimées à terme. En bleu, il s'agit des étapes qui ne sont pas exécutées dans le présent travail, mais qui devraient être développées.

Le processus actuel consiste d'abord à acquérir des données. Dans ce cas, elles sont sous forme de photographies des personnes se déplaçant sur les chemins de randonnée et passant devant les pièges caméras. Afin de respecter la loi, du ruban adhésif transparent est placé devant la lentille de la caméra pour rendre les personnes impossibles à identifier. Les photos sont alors automatiquement stockées sur la carte SD de la caméra et récupérées tous les mois. Ensuite, elles sont transférées sur un disque dur externe conservé par les membres de l'équipe. Les métadonnées de ces photos (date, heure, identifiant) et les informations liées à l'emplacement de la caméra sont stockées dans une base de données SQL.

Matériel et méthode

Puisque nous ne disposons pas des ressources nécessaires pour traiter les photos, celles-ci sont envoyées sur le serveur⁷ de la plateforme de calcul et de modélisation informatique (CAMI) de Gembloux Agro-Bio Tech. Ce serveur est destiné au calcul intensif permettant de rendre le traitement environ dix fois plus rapide qu'avec un ordinateur basique. Via ce cluster distant, l'algorithme de deep learning est utilisé afin d'analyser chaque photo individuellement. Cet algorithme (décrit p.99) détecte, classe et segmente les objets sur l'image. Un fichier JPEG de la photo traitée est alors créé ainsi qu'un fichier texte qui va permettre de constituer une base de données générale et analysable.

Un dossier d'analyse d'impact relative à la protection des données (AIPD) est actuellement en cours. En partant de l'hypothèse que l'algorithme utilisé possédera un meilleur taux de réussite sur des photos nettes, si ce dossier est accepté, alors toute la chaîne de traitement devrait être plus efficace. Pour l'instant, l'algorithme utilisé est appliqué par « transfer learning ».

Pour pouvoir obtenir ces photos nettes, l'accès aux données doit être très restreint et celles-ci doivent être chiffrées. Le mieux serait qu'à partir de la carte SD, les photos passent directement dans un algorithme afin d'être chiffrées et donc inaccessibles, sauf aux personnes possédant la clé de chiffrement. Ensuite, ajuster le modèle d'analyse des images pour qu'il demande à l'utilisateur la clé de chiffrement en entrée devrait permettre à l'ordinateur de déchiffrer l'image, de la traiter et de la rechiffrer sans que l'opérateur n'ai besoin de visualiser la photo. L'application pourrait donc être partagée avec des utilisateurs non agréés.

Pour rendre ce procédé viable, un autre point majeur à résoudre est la vitesse de traitement. Même si un cluster puissant en CPU peut augmenter la vitesse d'analyse, il sera vite dépassé par une carte graphique équipée d'un GPU décent. Selon différents experts consultés, une carte graphique à moins de 1.000 euros pourrait déjà fournir des résultats nettement plus rapides.

1) Acquisition des données

Pour collecter les photos, 20 caméras ont été placées sur les quatre massifs (cinq par site), de juillet 2018 à aujourd'hui. Les données utilisées dans le cadre de ce travail s'arrêtent au relevé de janvier 2019. Cela permet donc de couvrir une période de sept mois, sur trois saisons différentes.

Le modèle des caméras est le Dörr Snapshot Limited Black 5.0 S⁸. En activité de jour comme de nuit, ces caméras basiques permettent de capturer les objets en mouvements grâce à la détection infrarouge qui fonctionne jusqu'à 15 mètres. D'une résolution de trois mégapixels, cette dernière est amoindrie en raison du film adhésif transparent placé devant la lentille de la caméra. Afin d'éviter les vols et de ne pas perturber la quiétude des promeneurs, les caméras ont été placées entre deux et trois mètres de hauteur. Ensuite, la sensibilité du détecteur a été réglée. La végétation qui parasitait l'image a été retirée.

⁷ Serveur Lenovo Thinksystem SR850, 4 CPU de type Intel(R) Xeon(R) Platinum 8176 CPU @ 2.10GHz, 3 To de mémoire RAM DDR4

⁸http://www.produktinfo.conrad.com/datenblaetter/1600000-1699999/001661504-an-01-ml-DOERR_SNAPSHOT_LIMITED_5_0S_de_en_fr_it.pdf

Matériel et méthode

A chaque détection, la caméra prend deux photos en rafale (temps de réaction de 0,9 secondes). Le délai minimum entre deux séquences est paramétré à 10 secondes (une séquence correspond donc à deux photos). Ci-dessous, un exemple de l'angle de vue de chaque caméra (Figure 13).

L'emplacement de la caméra joue donc un grand rôle quant à la qualité de la donnée. Cet aspect doit être commenté et critiqué afin d'optimiser les modalités de placement des caméras. Comme cela est évalué en lien avec la capacité de l'algorithme à détecter et classifier les objets sur les images qui en découlent, cette critique sera réalisée dans un chapitre commun (voir p.59).

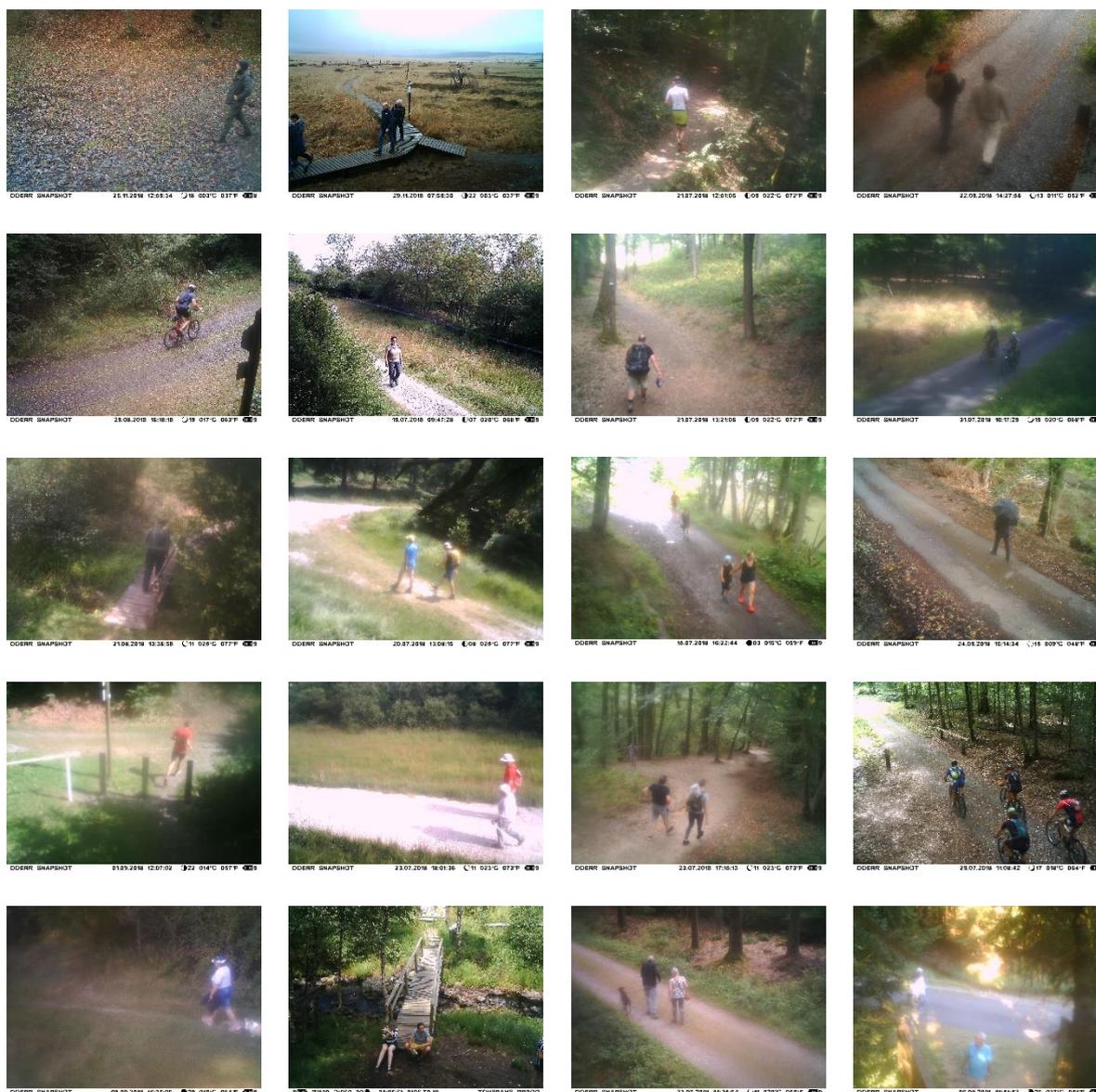


Figure 13: Exemple de photo prise par chaque caméra. Colonne 1: forêt d'Anlier: (a) Stand de Tir, (b) Fagne Jean Simon, (c) Pont, (d) Etang Fabrique, (e) Vallée ; Colonne 2 : Réserve naturelle des Hautes Fagnes : (f) Baraque Michel, (g) Botrange, (h) La Béole, (i) Polleur, (j) Pont Marie ; Colonne 3 : Parc naturel des Deux-Ourthe : (k) Barrage de Nisramont, (l) Cheslé, (m) Engreux, (n) Hérou, (o) Plateau des Tailles ; Colonne 4 : Forêt de Saint- Hubert : (p) Beyoli, (q) Bilaude, (r) Diglette Fourneau, (s) Masblette Pont Mauricy, (t) Priesse.

2) **Stockage des données**

Les images sont stockées sur un disque dur détenu par les membres de l'équipe. Elles sont placées dans des dossiers relatifs à l'emplacement de la caméra et dans des sous-dossiers correspondant à la période de capture. Enfin, les métadonnées sont aussi enregistrées dans le fichier (date, heure, minute, seconde). Une base de données SQL reçoit toutes ces informations et est exportable sous diverses formes par les personnes autorisées. Enfin, une application web a été créée par le Pr. Marc Dufrière afin de pouvoir étiqueter manuellement les photos en fonction de ce que l'opérateur y voit.

3) **Traitement des données**

Les deux premiers mois ont servis à apprendre les bases du langage Python (Python Software Foundation) ainsi que la théorie en ce qui concerne la détection et la classification d'objets sur des photos par ordinateur. Après avoir discuté avec des experts en analyse d'image de l'INBO (Instituut voor Natuur and Bozonderzoek, équivalent du DEMNA flamand), du bureau d'étude Aerospacelab et de la faculté des sciences appliquées de l'Université de Liège, il en est ressorti que créer un réseau neuronal convolutionnel afin d'analyser ces images n'est pas faisable dans le cadre d'un travail de fin d'études pour un futur bioingénieur. Dès lors, ils ont conseillé d'utiliser le modèle le plus actuel et adapté aux images rassemblées et de l'appliquer directement à l'étude. Il s'agit de l'architecture Mask R-CNN (He et al., 2017). Le script de l'architecture adaptée pour le dataset est disponible en Annexe 2, la description du modèle mentionné est disponible en Annexe 3.

III. Evaluation de la qualité du modèle

Afin de vérifier la qualité du modèle, un échantillonnage partiellement aléatoire a été réalisé. 160 photos par caméra ont été choisies (issues de la même période de récolte, choisie aléatoirement), ce qui fait un total de 3.200 photos, soit un sous-échantillon représentant 0,86 % de l'ensemble du dataset. Bien que Mask R-CNN ait été entraîné sur le set de données COCO (Figure 14) (Lin et al., 2014) qui comprend 1,5 million d'objets dont 250.000 personnes, il convient de vérifier si celui-ci fonctionne correctement sur les images de l'étude qui sont floutées et en extérieur. De plus, il est intéressant d'étudier l'efficacité de celui-ci sur des classes qui pourraient apporter des informations supplémentaires telles que les chiens, les sacs à dos, les vélos et les véhicules qui pourraient améliorer la compréhension du comportement des visiteurs. L'ensemble des 90 classes détectables par Mask R-CNN est disponible en Annexe 4.

Des matrices de confusion sont ensuite créées et comparées. Il faut différencier pour chaque classe : Les objets que le modèle détecte et classe correctement (vrai positif), ce qu'il croit détecter à tort (faux positif) et ce qu'il ne détecte pas (faux négatif). Enfin, ce que le modèle détecte, mais classe mal est donc un faux positif pour l'objet classifié et un faux négatif pour l'objet raté.

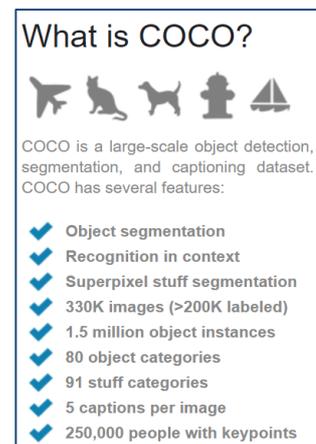


Figure 14: Set de données COCO (Common Objects in COntext), qui a servi à entraîner Mask R-CNN (Lin et al., 2014)

IV. Création de la base de données

En sortie de l'algorithme, on obtient d'une part les images traitées et d'autre part des fichiers texte avec la liste des objets détectés sur chaque photo. Grâce à leur identifiant unique composé d'une série de chiffres correspondant au numéro de la photo ainsi que de la date et du moment de la prise de vue (heure, minute, seconde), il est possible de relier chaque fichier texte aux métadonnées du cliché.

Une base de données correcte doit être la plus proche possible de la réalité et refléter au mieux le nombre d'objets qui sont passés devant les caméras. Sachant que les caméras sont configurées pour prendre deux photos à chaque événement, il faut tout d'abord éviter de compter chaque individu en double. Pour éviter cela, la moyenne est calculée entre chaque doublon. Ensuite, elle est arrondie à l'entier supérieur afin de maximiser les détections.

La redondance entre deux événements de détection peut aussi arriver. Si les mêmes personnes se trouvent pendant plus de dix secondes dans le champ de détection, elles seront alors comptées plusieurs fois. D'un autre côté, des personnes marchant à une certaine distance l'une de l'autre pourraient déclencher la photographie directement après le premier événement. Il ne s'agirait alors pas des mêmes individus, mais peut être de groupes organisés comme expliqué ci-dessus. Cette information comportementale est très importante en ce qui concerne la gestion des milieux et doit être conservée au maximum. Il faut donc trouver un équilibre entre la suppression de la redondance due aux personnes stagnant devant la caméra ou se déplaçant trop lentement devant celle-ci et les groupes étalés sur différentes séquences.

Pour diminuer l'impact de ce problème, trois journées ont été désignées au hasard pour chaque caméra. Une heure a été attribuée à chaque jour. Pour un total de 60 échantillons, les photos ont été parcourues visuellement jusqu'à trouver de la redondance, c'est-à-dire au moins quatre photos de ce qui est jugé être les mêmes personnes ou bien des individus faisant partie d'un même groupe. Il en ressort que dans 78 % des cas, les personnes constituant des groupes se trouvent au moins toutes sur une photo. Dès lors, lorsque deux photos sont trop proches dans le temps, seule la photo avec le plus de personnes détectées est conservée. De plus, l'analyse révèle que 80 % des visiteurs qui créent de la redondance ne stagnent pas devant la caméra, mais se déplacent simplement trop lentement.

Dernièrement, les données utilisées pour étudier la qualité du modèle ont été reprises. Comme on peut le voir dans le Tableau 2, au total, 1.299 individus ont été comptés sur les 3.200 photos. Il apparaît cependant que le pas de temps minimum de 10 secondes entre les photos n'est pas toujours bien respecté. Enlever les photos de personnes espacées de moins de 10 secondes décompte déjà 182 individus, soit 12 % de la population totale. Après avoir répété l'opération pour différents pas de temps, il est décidé de doubler le temps minimum entre deux photos. Un pas de temps de 20 secondes permet d'être sûr de ne pas surestimer le nombre de personnes fréquentant les différents massifs, quitte à perdre une partie de l'information concernant des groupes répartis sur plusieurs photos.

Matériel et méthode

Tableau 2: Nombre de personnes non comptées en fonction de la modification du pas de temps minimal entre deux événements de détection (n= 3.200 photos)

Pas de temps minimal	Nombre d'individus	Nb d'individus perdu	% d'individus perdu
<10 (bug)	1299		
10	1117	182	12%
11	1097	20	14%
12	1088	9	14%
13	1084	4	15%
14	1078	6	15%
15	1049	29	17%
16	1032	17	19%
17	1018	14	20%
18	1010	8	20%
19	999	11	21%
20	988	11	22%
21	981	7	23%
22	977	4	23%
23	974	3	23%
24	965	9	24%
25	962	3	24%
30	937	25	26%

La base de données a été créée et optimisée grâce à un script disponible en Annexe 2 et réalisé via Rstudio et le langage R (R Development Core Team, 2005).

V. Application de l'indice sur les données collectées par AGRETA

1) Comparaison avec la méthode de Lupp et al. (2016)

Afin de mettre ce prototype en application, il est d'abord important de le comparer aux méthodes existantes. La méthode la plus proche est celle utilisée dans l'étude de Lupp et al. (2016). Dans cette étude, les auteurs ont évalué la valeur récréative des deux forêts urbaines dans la région de Munich. Pour se faire, des interviews ont été réalisées, et des comptages ont été réalisés grâce à des pièges caméras. La période de prise de données ayant duré plus de 16 mois, la quantité de photos à analyser était énorme. Pour alléger ce travail, les auteurs ont décidé de sélectionner 90 jours sur l'ensemble. 20 jours ont été choisis sur la période avril-mai, d'autres ont été choisis au hasard et le reste selon les conditions météo. Sur ce sous-échantillon : une régression linéaire a été réalisée entre le nombre de photos par jour et le nombre de personnes issus de ces photos. Il en résulte un coefficient de détermination compris entre 76 et 96 %, ce qui traduit une forte corrélation entre le nombre de photos prises sur une journée et le nombre de personnes qui sont passées devant les caméras. L'extrapolation a donc pu être étendue à l'ensemble de la période.

En partant de ce protocole, quatre journées ont été sélectionnées aléatoirement par chaque caméra. Au total, 80 journées ont donc été sélectionnées puis analysées manuellement. Le nombre de photos et le nombre de personnes par journée ont été relevés. Les faux déclenchements à répétition liés à la luminosité ou à la météo ont été notifiés ainsi que tout autre événement qui pourrait nuire à la qualité de la régression.

2) Quelles informations peuvent être tirées de ce nouvel indice ?

Une fois la précision du modèle connue, les données récoltées dans le cadre du projet AGRETA servent d'exemple pour tester le procédé. Selon (Muhar et al., 2002), des informations à collecter sont le nombre de visites et leurs variations temporelles, l'activité des visiteurs ainsi que leur densité. Le nombre de visites totales est l'indicateur le plus important. Il permet par exemple de se rendre compte de façon simple de la pression anthropique sur un milieu, mais aussi de sa valeur (Bateman et al., 2006; Jones et al., 2010).

Ces informations peuvent être agglomérées sur l'ensemble des caméras, mais aussi détaillées pour chacune d'elles afin d'avoir une compréhension plus fine des variations entre caméras et des préférences pour certains lieux par rapport à d'autres.

Résultats

Les résultats sont présentés en trois parties. Tout d'abord, une exploration des données brutes est réalisée. Cette exploration prend en compte les variations temporelles à l'échelle de la région, des massifs et des caméras elles-mêmes ainsi que les relations qu'il peut y avoir entre les différentes caméras.

Dans la seconde partie, la même analyse est répétée en fonction du nombre de personnes comptées par l'algorithme. Cette partie des résultats constitue donc un exemple de l'application de l'indice sur le cas d'étude d'AGRETA. Afin de donner un aperçu des informations comportementales qu'il est possible d'obtenir, les différentes tailles de groupes sur une photo sont comparées avec différents facteurs qualitatifs.

Enfin, une comparaison est réalisée entre le nombre de photos obtenues et le nombre de personnes comptées. Un sous-échantillon de 80 journées a été étudié afin de comparer les données brutes à la méthode de subsampling utilisée par Lupp et al. (2016) (voir p.52). Finalement, le nombre de photos obtenues est comparé avec le nombre de personnes comptées par le modèle.

I. Etude du nombre de photos obtenues

1) Résultats et variations temporelles

Durant les sept premiers mois de relevés, les caméras ont capturé **356.201** photos. Les codes des massifs et des caméras sont disponibles dans le Tableau 3.

Tableau 3: Tableau de correspondance entre les noms des massifs et des caméras et leur code respectif.

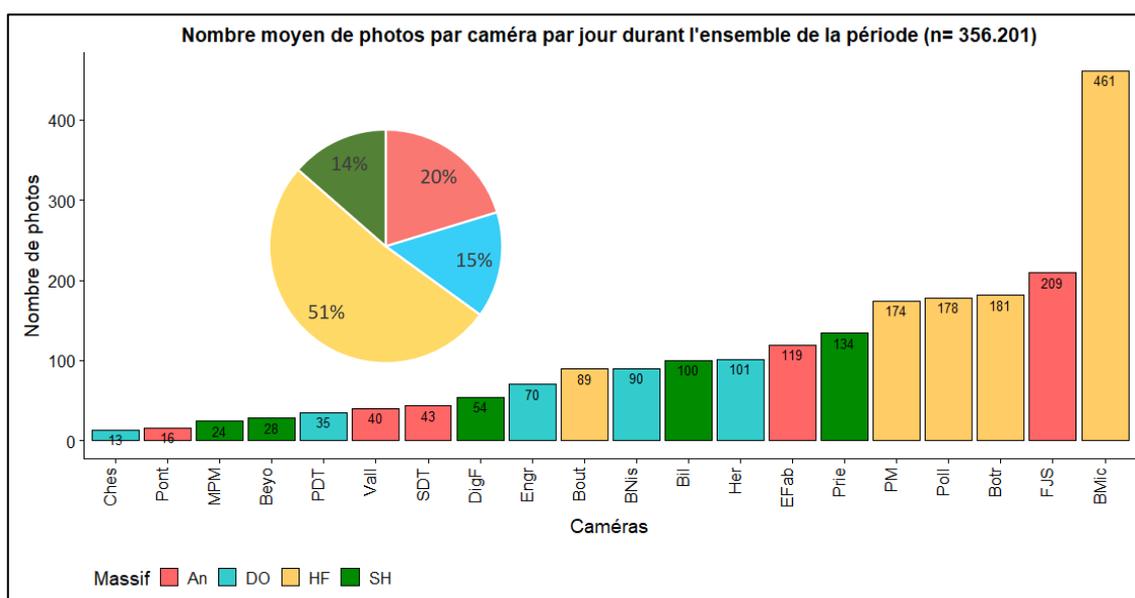
	Nom	Code	Nom	Code
Massif	Anlier	An	Deux Ourthes (PNDO)	DO
Caméra	Etang Fabrique	EFab	Barrage Nisramont	BNis
	Fagne Jean Simon	FJS	Chesle	Ches
	Pont	Pont	Engreux	Engr
	Stand De Tir	SDT	Herou	Her
	Vallee	Vall	Plateau Des Tailles	PDT
Massif	Hautes Fagnes	HF	Saint-Hubert	SH
Caméra	Baraque Michel	BMic	Beyoli	Beyo
	Botrange	Botr	Bilaude	Bil
	La Béole	Bout	Diglette Fourneau	DigF
	Polleur	Poll	Masblette Pont Mauricy	MPM
	Pont Marie	PM	Priesse	Prie

Selon les caméras, la période de prise de données a commencé entre le 11 juillet et le 8 août 2018. Le traitement des données dans le cadre de ce travail s'est arrêté entre le 22 janvier et le 5 février 2019. Le détail de la période de capture ainsi que le nombre de photos total par caméra est disponible dans le Tableau 4. Le Graphique 1 montre les moyennes par caméra et par jour pour l'ensemble de la période.

Résultats

Tableau 4: Répartition du nombre total de photos prises par caméra. Les indices «¹» et «²» devant les noms des caméras indiquent des circuits de randonnée communs. Les indices «^x» représentent des emplacements avec des risques accrus d'aller-retour de la part des visiteurs dû à des points de vue ou à des miradors d'observation de faune. Les indices «^a» indiquent des caméras dont la carte SD est pleine. Les indices «^b» indiquent des caméras éteintes.

Massif	Camera	Date mise en place	Date dernier relevé	Nombre de jours hors capture	Nombre de jours de capture	Nombre total de photos	Nombre moyen de photos par jour
An	EFab	21-08-18	05-02-19		168	19.936	119
An	FJS	21-08-18	05-02-19	61 ^a	107	22.348	209
An	Pont	21-08-18	05-02-19		168	2.762	16
An	SDT	21-08-18	05-02-19		168	7.242	43
An	Vall	21-08-18	05-02-19		168	6.774	40
Total massif					779	59.062	427
DO	BNis ²	12-07-18	29-01-19		201	18.003	90
DO	Ches	15-07-18	29-01-19		198	2.606	13
DO	Engr ²	12-07-18	29-01-19		201	13.986	70
DO	Her	12-07-18	29-01-19	43 ^b	158	16.010	101
DO	PDT	12-07-18	29-01-19		201	7.002	35
Total massif					959	57.607	309
HF	BMic ¹	11-07-18	12-02-19	80 ^{ab}	136	62.713	461
HF	Botr ¹	11-07-18	12-02-19	49 ^b	167	30.272	181
HF	Bout ¹	11-07-18	12-02-19		216	19.151	89
HF	PM ¹	11-07-18	12-02-19		216	37.674	174
HF	Poll	11-07-18	12-02-19	70 ^b	146	26.021	178
Total massif					930	175.831	1.083
SH	Beyo ^x	08-08-18	22-01-19		167	4.748	28
SH	Bil ^x	13-07-18	22-01-19		193	19.326	100
SH	DigF	08-08-18	22-01-19		167	9.096	54
SH	MPM	13-07-18	22-01-19		193	4.587	24
SH	Prie ^x	13-07-18	22-01-19		193	25.944	134
Total massif					913	63.701	286
Total général					3581	356.201	2.105



Graphique 1: Nombre moyen de photos prises par jour par caméra durant l'ensemble de la période

Résultats

Le Graphique 2 montre l'évolution du nombre de photos prises par semaine pour chaque caméra. 49 % des photos ont été capturées dans la réserve naturelle des Hautes Fagnes-Eiffel dont 18 % à la Baraque Michel. Le massif d'Anlier rassemble plus de photos (20 %) que les massifs de Saint-Hubert et du PNDO (14 et 15 %).

Dans le massif des Hautes Fagnes, entre le 21 septembre et le 31 octobre, la caméra de la Baraque Michel n'a pris aucune photo, car la carte SD était pleine. Entre le 19 et le 29 novembre, la caméra est tombée. Concernant le site de Polleur, la caméra étant éteinte, il n'y a plus de données à partir du 4 décembre. La caméra de Botrange semble aussi éteinte entre le 12 septembre et le 31 octobre. Il ne ressort pas de périodes particulièrement plus photographiées que d'autres. Certains pics d'affluence se sont déroulés en semaine notamment durant les semaines du 10 et 17 septembre et celle du 12 novembre.

Au sein du massif de Saint-Hubert, il existe une forte hétérogénéité entre les différentes caméras. La période la plus fréquentée est le milieu de l'été (dernière semaine de juillet). Les caméras de Diglette et Beyoli n'étaient pas encore placées à ce moment. La majorité des photos ont été prises à Priesse et à Bilaude. Particulièrement pour ces deux sites, on observe un regain net de déclenchement durant la fin de mois de septembre et le début du mois d'octobre.

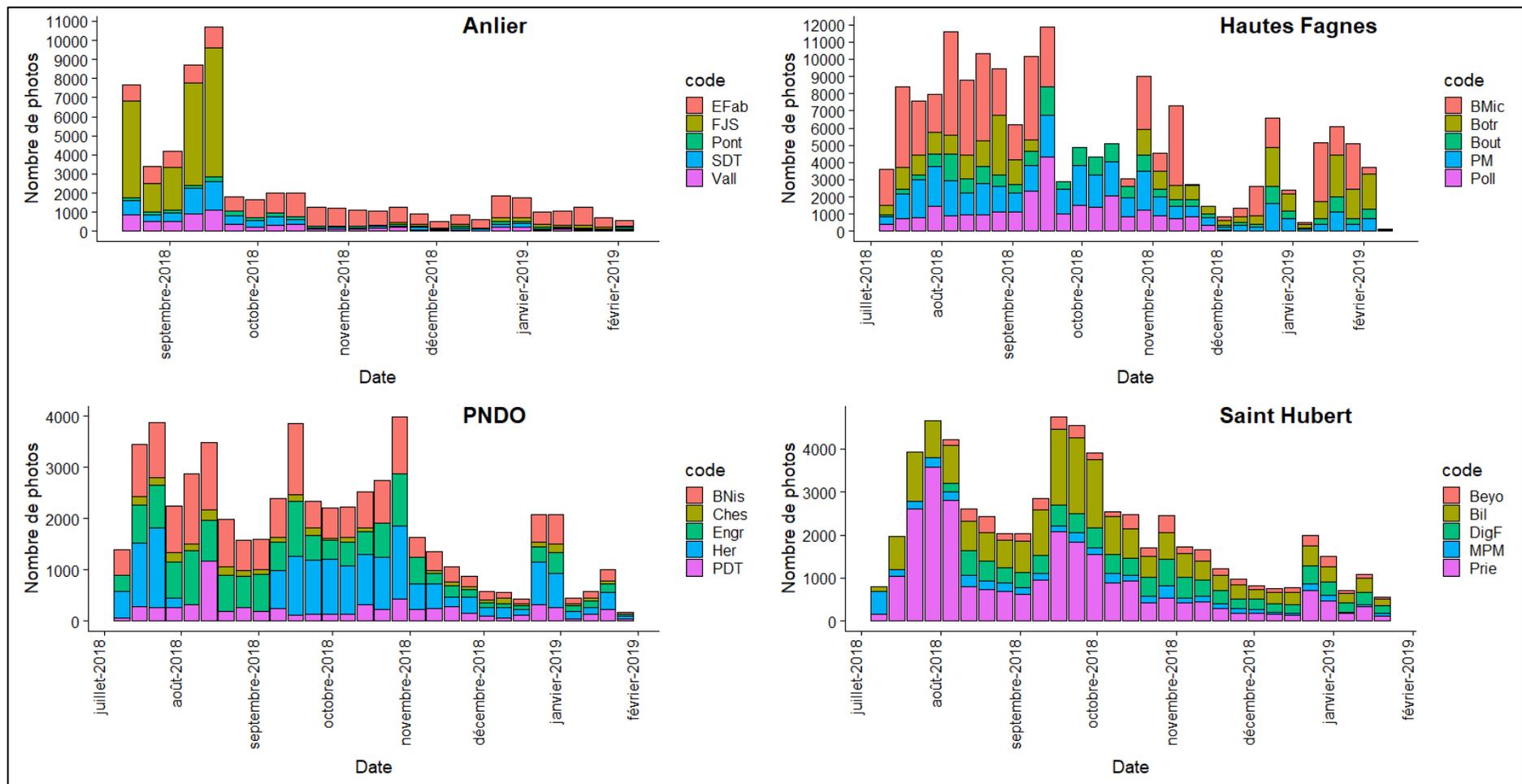
Concernant le Parc Naturel des Deux Ourthes, le nombre de photos à Engreux est fort similaire à celui d'Hérou et du Barrage de Nisramont. De plus, au fil du temps Hérou possède nettement plus de détections que Engreux, la caméra était cependant éteinte entre le 29 juillet et le 10 septembre. De manière générale, Hérou prend en moyenne 31 photos de plus que Engreux par jour et 11 de plus que le barrage de Nisramont.

Même si deux caméras du massif d'Anlier se trouvent dans le top sept des sites les plus photographiés, la grande majorité des photos ont été prises durant la fin du mois d'août et le début du mois de septembre. Ce qui correspond à la date de placement de ces caméras. Comme on le verra plus tard, de nombreuses photos prises sont des faux déclenchements dus à la végétation. De plus, la caméra Fagne-Jean-Simon n'a pas pris de photos entre le 21 septembre et le 21 novembre, car la mémoire de la carte SD était saturée, ce qui explique la diminution soudaine pour cette caméra entre la 3^{ème} et la 4^{ème} semaine de septembre. La caméra des Etangs de la Fabrique est restée constante tout au long de la période avec entre 500 et 2.000 personnes par mois tandis que Stand de Tir et Vallée ont vu leur pic atteint durant la fin de l'été et le début de l'automne. Dans l'ensemble, le massif d'Anlier est nettement moins photographié que les autres massifs.

De manière générale, il faut noter que les données pour le mois de juillet pour Saint-Hubert, PNDO et les Hautes Fagnes et le mois d'août pour Anlier sont incomplètes. On remarque aussi que le début du mois de décembre et le début du mois de janvier sont les moins photographiés. Un regain de détection est néanmoins observable durant les deux dernières semaines de l'année. Le mois d'août est généralement le plus photographié, au sein d'une distribution relativement normale avec les mois de juillet et septembre.

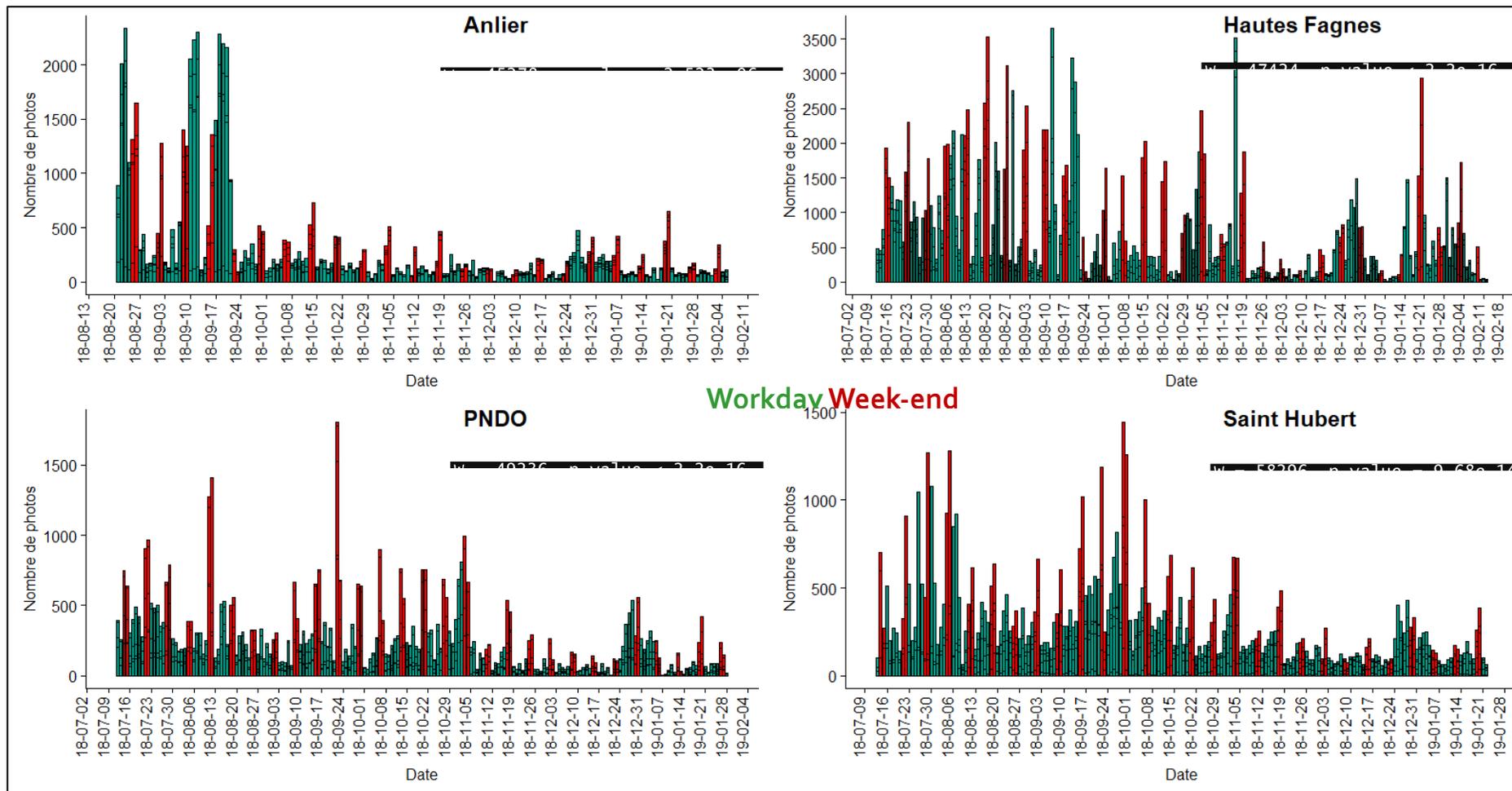
Une représentation de la quantité de déclenchements durant les week-ends par rapport à la semaine est disponible dans le Graphique 3. Puisque les distributions ne sont pas normales, il n'est pas conseillé d'utiliser un test paramétrique pour les comparer. Un test de comparaison des médianes a donc été réalisé (Wilcoxon, 1945; Mann et al., 1947). Il ressort de cette analyse que les fréquentations médianes durant les week-ends sont significativement supérieures que celles durant la semaine pour tous les massifs. Les caméras se déclenchent donc plus souvent durant les week-ends que durant la semaine.

Résultats



Graphique 2: Nombre de photos prises par caméra et par semaine sur chaque massif.

Résultats



Graphique 3: Nombre de photos prises par jour et par massif en fonction du jour de la semaine. Les jours de la semaine sont représentés en vert, les samedis et dimanches sont en rouge. Un test de Wilcoxon a été réalisé sur chaque distribution

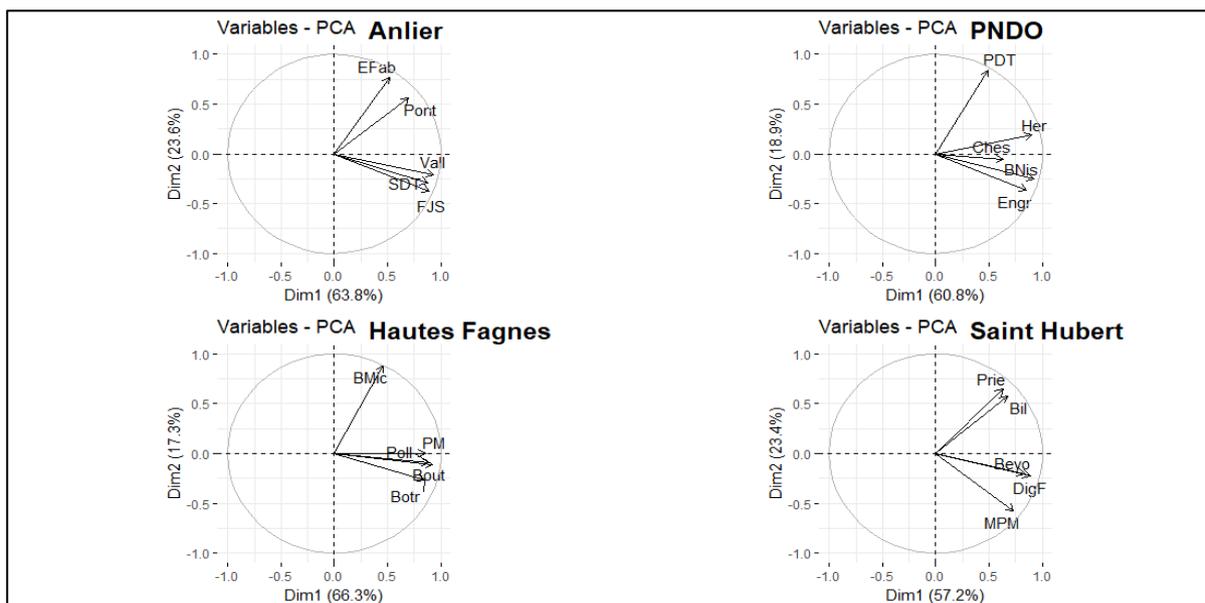
2) Relations entre caméras

Afin de déceler les éventuelles colinéarités entre les caméras, une matrice de corrélation a d'abord été créée (Annexe 5). Elle montre que la grande majorité des caméras réagissent de la même façon aux variations temporelles et semblent être corrélée à l'échelle régionale. Seuls Cheslé et le Plateau des Tailles (PNDO) sont considérés comme non corrélés avec plus d'un tiers du reste des caméras et sont donc relativement indépendants.

Le nombre de variables rendant difficile la visualisation des relations entre caméras, une analyse en composantes principales a été réalisée (Pearson, 1901; Hotelling, 1933) (Annexe 5). Le jeu de données comprend plusieurs valeurs manquantes. Puisque les caméras ont eu des problèmes techniques et qu'elles n'ont pas été placées en même temps, toutes les dates ne peuvent pas être prises en compte dans l'étude des corrélations entre caméras. Pour résoudre ce problème, les données manquantes ont été estimées par imputation, grâce au package «missMDA» (Josse et al., 2016) qui permet d'estimer le nombre de composantes principales optimales en imputant des valeurs selon la méthode de cross validation k-fold (Bro et al., 2008; Josse et al., 2012). La méthode a été testée pour un minimum d'une composante principale et maximum cinq.

De manière générale, il est difficile de déceler des tendances en voyant cette ACP. La totalité des caméras se retrouvent à droite de l'axe 1 qui explique 44 % de la variance. Trois caméras du massif d'Anlier ont participé majoritairement à la création de la 2^{ème} composante principale (qui explique 15,6 % de la variance). De plus, ce sont les trois caméras qui sont le mieux représentées sur le 1^{er} plan factoriel (Fagne Saint Jean, Vallée et Stand de Tir).

Pour y voir plus clair, le Graphique 4 se concentre sur chaque massif individuellement. Dans tous les cas, la première dimension explique plus de la moitié de la variance. La corrélation positive entre toutes les caméras subsiste à l'échelle des massifs. La différenciation sur le second axe est plus claire à cette échelle. Pour chaque massif, les points sont proches du cercle de corrélation, ce qui signifie que toutes les caméras sont bien représentées dans ce plan factoriel. Il est donc possible de tirer des conclusions.



Graphique 4: Analyse en composante principale de la relation entre les différentes caméras au fil du temps et à l'échelle de chaque massif – 1er plan factoriel

Résultats

Sur le deuxième axe, dans le massif d'Anlier, les caméras Vallée, Stand de Tir et Fagne Jean Simon s'opposent au Pont et aux Etangs de la Fabrique. En Hautes Fagnes, la caméra Baraque Michel s'éloigne clairement du reste du groupe et contribue à 78 % à cet axe. À Saint-Hubert, Priesse et Bilaude s'opposent à Beyoli, Diglette Fourneau et Masblette Pont Mauricy. Concernant le PNDO, les résultats sont plus nuancés. Engreux et le Barrage de Nisramont semblent s'opposer au plateau des Tailles et à Hérou dans une moindre mesure. La caméra de Cheslé est moins bien représentée sur ce plan factoriel.

II. Etude du nombre de personnes comptées

Cette partie décrit la qualité de la chaîne de traitement située entre le passage dans l'algorithme et le traitement des résultats. Étant donné que la qualité des informations qui sortent du modèle est liée aux données qui y sont entrées, il faut dans un premier temps étudier la qualité de cette relation.

Ensuite, l'indice ainsi établi est appliqué sur le jeu de donnée du projet AGRETA. Cet exemple permet de mettre en avant le type d'informations qu'il peut fournir. Comme discuté au point suivant, seules les informations liées aux personnes elles-mêmes (leur nombre, la variation temporelle de leur fréquentation, la taille des groupes qu'elles forment) sont étudiées. Le modèle non entraîné sur le dataset d'AGRETA n'est pas encore suffisamment fiable pour donner des informations quant à des caractéristiques plus précises telles que leurs moyens de déplacement, leur âge, leur sexe ou leur matériel.

1) Evaluation de l'efficacité du modèle

Après avoir étiqueté les 3.200 photos choisies de façon partiellement aléatoire (description p.31) et comparé avec les résultats du modèle, des matrices de confusion ont été réalisées pour chacune des 20 caméras. Les résultats complets sont disponibles dans l'Annexe 6 et sont résumés dans le Tableau 5. La qualité de la détection est définie par le ratio entre les vrais positifs et les faux négatifs. La qualité de la classification est définie par le ratio entre les vrais positifs et les faux positifs. En plus des personnes, les chiens, les sacs à dos, les vélos et les véhicules motorisés ont aussi été évalués après avoir remarqué leur présence récurrente et la qualité variable du modèle quant à ces classes.

On parle d'efficacité du modèle et pas de qualité de celui-ci, car le résultat dépend fortement de la qualité des photos qui lui sont données en entrée. Une critique de la configuration des images serait donc associée à la discussion de l'efficacité du modèle.

Après avoir notifié qu'une grande partie des vélos étaient classifiés en tant que moto et qu'aucune vraie moto n'était présent dans le sous-échantillon, les classes « vélos » et motos » ont été agrégées afin de créer une classe « deux roues ».

Résultats

Tableau 5: Evaluation de la qualité du modèle. La qualité de la détection est définie par le ratio entre les vrais positifs et les faux négatifs. La classification est définie par le ratio entre les vrais positifs et les faux positifs.

	Détection			
	Moyenne	Ecart-type	Intervalle de confiance (95%)	
			Inférieur	Supérieur
Personnes	89,0%	8,9%	84,8%	93,2%
Chiens	57,9%	40,6%	38,9%	77,0%
Sacs à dos	51,2%	22,0%	40,9%	61,5%
Véhicules	90,9%	14,5%	84,1%	97,7%
Vélos	67,2%	34,0%	51,3%	83,1%
Deux roues	77,7%	34,4%	61,6%	93,8%
	Classification			
	Moyenne	Ecart-type	Intervalle de confiance (95%)	
			Inférieur	Supérieur
Personnes	97,1%	2,2%	96,1%	98,2%
Chiens	87,1%	29,9%	73,1%	100,0%
Sacs à dos	81,6%	19,0%	72,7%	90,5%
Véhicules	45,8%	41,2%	26,5%	65,1%
Vélos	86,3%	27,3%	73,5%	99,0%
Deux roues	88,2%	25,5%	76,2%	100,0%

Concernant la détection des chiens et des sacs à dos, on remarque des moyennes trop faibles et des intervalles de confiance trop larges pour pouvoir tirer des informations robustes quant à leur présence (respectivement $57,9 \pm 19,0$ et $51,2 \pm 14,0$ %). De plus, les tests effectués sur les sacs à dos sont uniquement réalisés sur les personnes dos à la caméra ou de côté. Les personnes face à la caméra chez lesquelles on ne voit que les brides du sac à dos ne sont pas prises en compte. Concernant les chiens, lorsqu'ils sont présents, ils sont souvent classifiés comme d'autres animaux (69 %), des personnes (22 %) ou des objets (4 %)⁹.

Les résultats concernant la qualité de la détection et la classification pour les vélos ($67,2 \pm 15,9$ % et $86,3 \pm 12,8$ %) donnent les mêmes tendances que pour les chiens et les sacs à dos. Ils ne sont pas suffisamment précis pour être étudiés plus en profondeur lors de l'analyse. Cependant, sachant que la plupart des faux négatifs sont en fait des détections de motos alors qu'aucune moto ne se trouvaient dans l'échantillon, il est possible d'améliorer ces chiffres en fusionnant ces deux classes en une catégorie « deux roues ». On atteint alors un seuil de précision de $77,7 \pm 16,6$ % de détection et $88,2 \pm 12,0$ % de classification. Il arrive aussi que les deux classes soient comptabilisées pour le même objet (Figure 15).

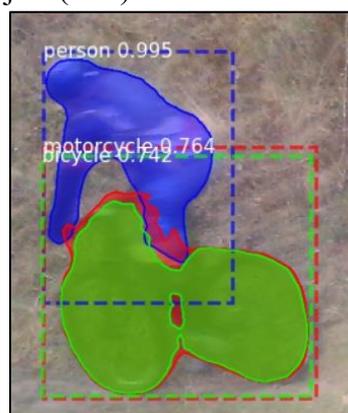


Figure 15: Problème de superposition de classes (vélo et moto)

⁹ N=102

Résultats

Concernant les véhicules, même si ceux-ci sont quasiment toujours détectés correctement ($90,9 \pm 6,8$ % du temps), le modèle présente un grand défaut à leur égard : la moyenne est basse et la variance élevée (respectivement 45,8% et 41,2%) en termes de classification. Cela est dû au taux très élevé de faux positif. L'arrière-plan dynamique (environnement extérieur) est souvent pris à tort pour un véhicule (Figure 16).

Les personnes sont détectées $89,0 \pm 4,2$ % du temps et classifiées correctement $97,1 \pm 1,0$ % du temps. Les 11,0 % de personnes non détectées sont dû à des causes inconnues dans 23,1 % des cas. Les 76,9 % de non-détectations restantes sont supposément connues et donc corrigéables (Graphique 5). La Figure 17 montre des causes possibles de non-détection.

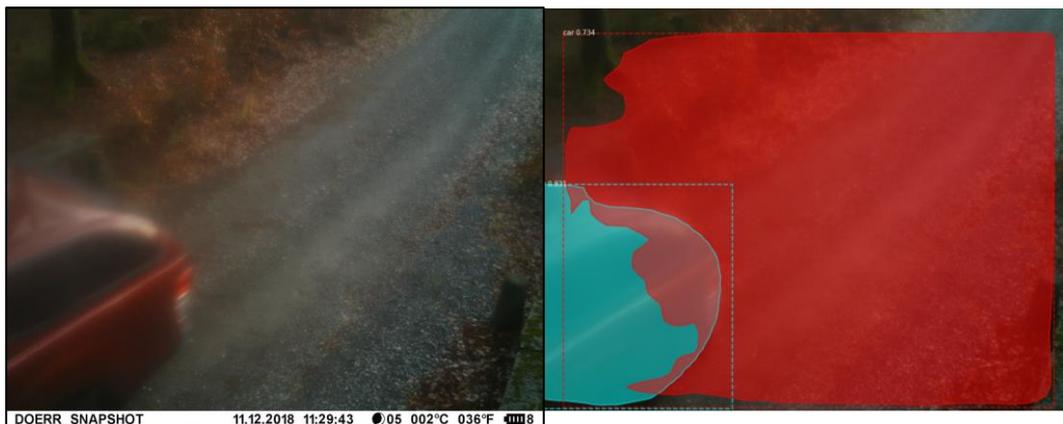
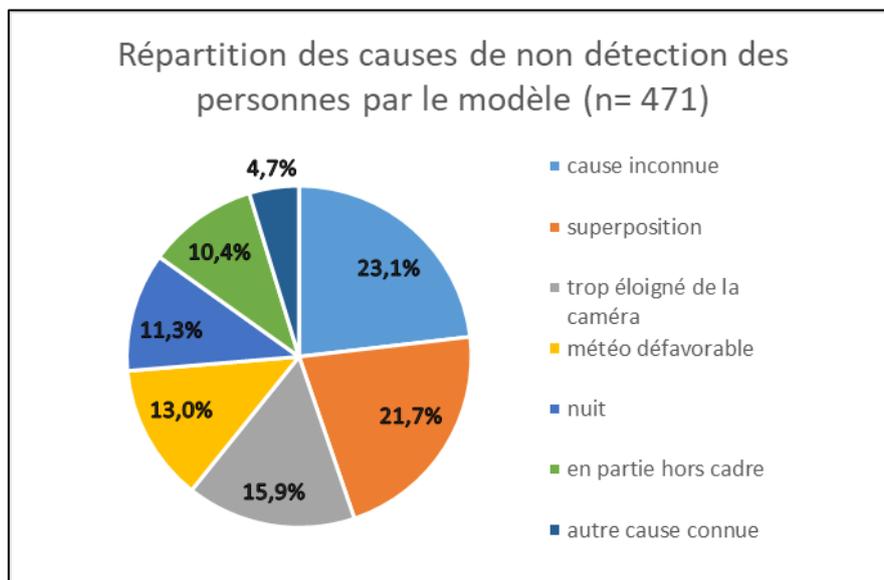


Figure 16: Exemple de faux positif dans le cas d'une voiture. La segmentation en bleu représente une vraie voiture tandis que celle en rouge est un faux positif



Graphique 5: Répartition des causes de non-détection des personnes par le modèle. Les autres causes connues peuvent être par exemple le port d'un parapluie ou la végétation dans le champ de la caméra.

Résultats



Figure 17: Causes possibles de mauvaise détection du modèle. De haut en bas et de gauche à droite : (a) conditions idéales, (b) superposition, (c) sujets trop éloignés, (d) buée due à la météo, (e) nuit, (f) sujet en partie hors cadre, (g) mauvaise exposition lumineuse, (h) auto-occlusion, (i) occlusion due au contexte

Résultats

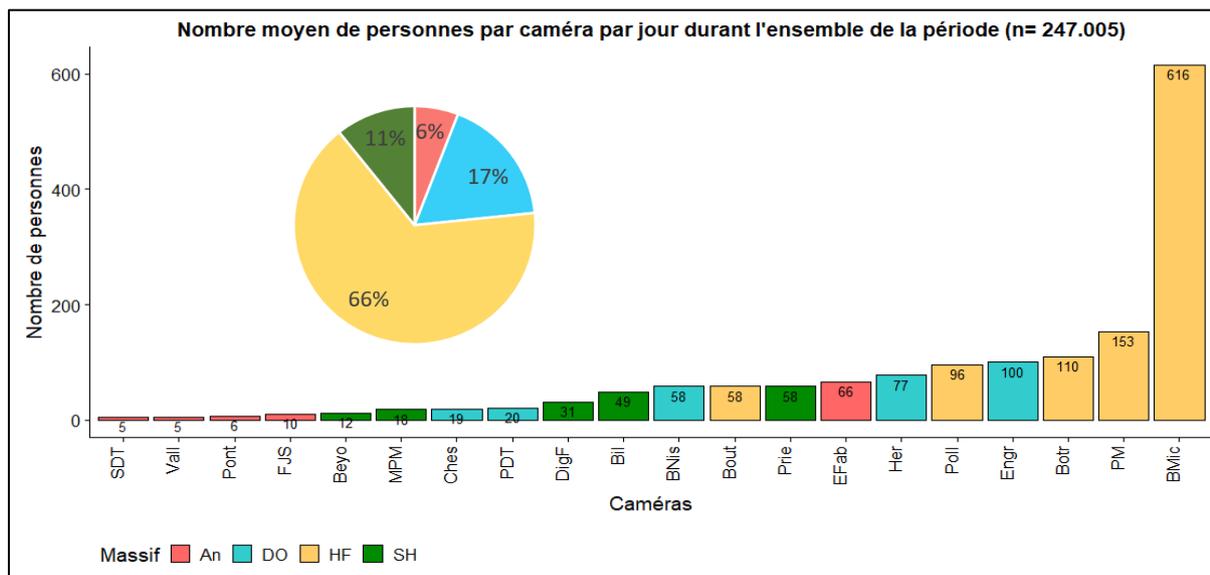
2) Résultats et variations temporelles

Après création et nettoyage de la base de données (intervalle de temps de 20 secondes entre deux photos), il reste 99.000 photos comprenant des personnes sur l'ensemble de la période pour un total de **247.005 personnes** détectées et classifiées par le modèle. Selon la précision de ce dernier, le total pourrait atteindre entre **274.000 et 289.000 personnes** compte tenu du ratio de détection et de classification. Il est aussi possible de décliner ce nombre par caméra (Tableau 7). La répartition des nombres moyens de personnes par jour et par caméra est montrée sur le Graphique 6. Le Graphique 7 détaille la fréquentation totale par caméra pour chaque massif durant l'ensemble de la période. Enfin, le Graphique 8 met en évidence les différences de fréquentation entre les week-ends et la semaine pour chaque massif.

Tableau 6: Répartition du nombre de personnes prises par caméra. Les indices « 1 » et « 2 » devant les noms des caméras indiquent des circuits de randonnée communs. Les indices « ^x » représentent des emplacements avec des risques accrus d'aller-retour de la part des visiteurs dû à des points de vue ou à des miradors d'observation de faune. Les indices « ^a » indiquent des caméras dont la carte SD est pleine. Les indices « ^b » indiquent des caméras éteintes.

Massif	Camera	Date mise en place	Date dernier relevé	Nb de jours hors capture	Nb de jours de capture	Nb total de personnes	Nb moyen de personnes par jour
An	EFab	21-08-18	31-01-19		163	10.826	66
An	FJS	21-08-18	31-01-19	61 ^a	102	1.018	10
An	Pont	21-08-18	31-01-19		163	1.030	6
An	SDT	21-08-18	31-01-19		163	872	5
An	Vall	21-08-18	31-01-19		163	827	5
Total massif					754	14.573	93
DO	BNis ^{x2}	12-07-18	29-01-19		201	11.703	58
DO	Ches	15-07-18	29-01-19		198	3780	19
DO	Engr ²	12-07-18	29-01-19		201	20.065	100
DO	Her	12-07-18	29-01-19	43 ^b	158	12.148	77
DO	PDT	12-07-18	29-01-19		201	4.018	20
Total massif					959	51.714	274
HF	BMic ¹	11-07-18	31-01-19	80 ^{ab}	124	76.364	616
HF	Botr ¹	11-07-18	31-01-19	49 ^b	155	16.999	110
HF	Bout ¹	11-07-18	31-01-19		204	11.932	58
HF	PM ¹	11-07-18	31-01-19		204	31.162	153
HF	Poll	11-07-18	31-01-19	70 ^b	134	12.828	96
Total massif					870	149.285	1.033
SH	Beyo ^x	08-08-18	22-01-19		167	1.989	12
SH	Bil ^x	13-07-18	22-01-19		193	9.536	49
SH	DigF	08-08-18	22-01-19		167	5.192	31
SH	MPM	13-07-18	22-01-19		193	3.438	18
SH	Prie ^x	13-07-18	22-01-19		193	11.278	58
Total massif					913	31.433	169
Total général					3.496	247.005	1.569

Résultats

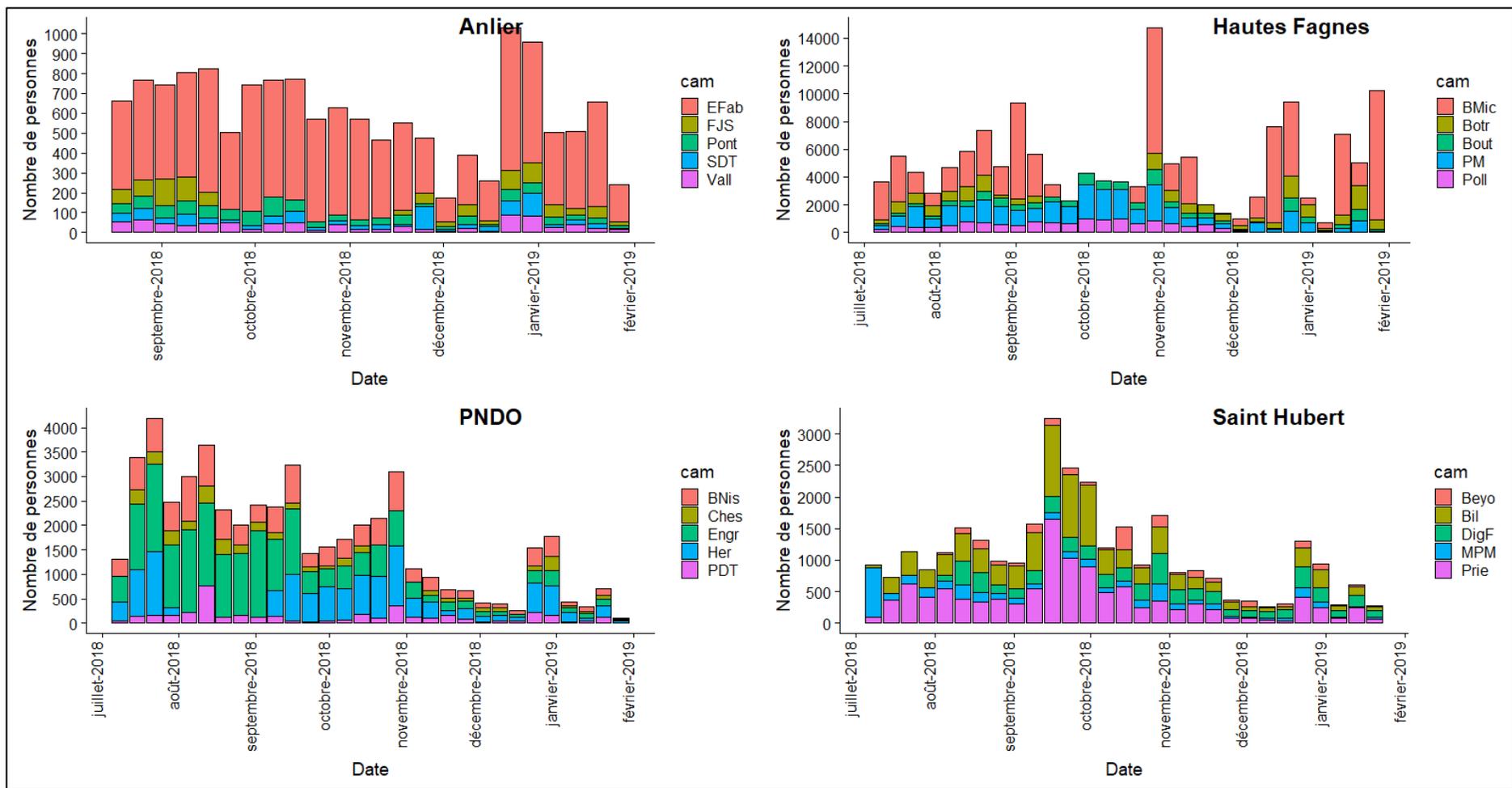


Graphique 6: Nombre moyen de personnes par caméra et par jour durant l'ensemble de la période

Le massif des Hautes Fagnes possède de loin le taux de visite le plus élevé avec 66 % des visiteurs dont 45 % de l'ensemble visitent la Baraque Michel. Le Pont Marie qui est le deuxième endroit le plus fréquenté ne représente que 10 % des visiteurs passés devant l'ensemble des caméras. Le pic de fréquentation pour le massif des Hautes Fagnes se situe entre fin septembre et début octobre. Ce pic n'est observable que pour les caméras de La Béole, Polleur et Pont Marie étant donné les problèmes techniques de la Baraque Michel et Botrange pour cette période. Certaines semaines, la Baraque Michel reçoit plusieurs milliers de visiteurs. Cependant, ces événements ne semblent pas avoir de lien précis avec la saisonnalité. La dernière semaine de janvier a accueilli plus de 7.000 visiteurs dans le massif (majoritairement dû à la Baraque Michel).

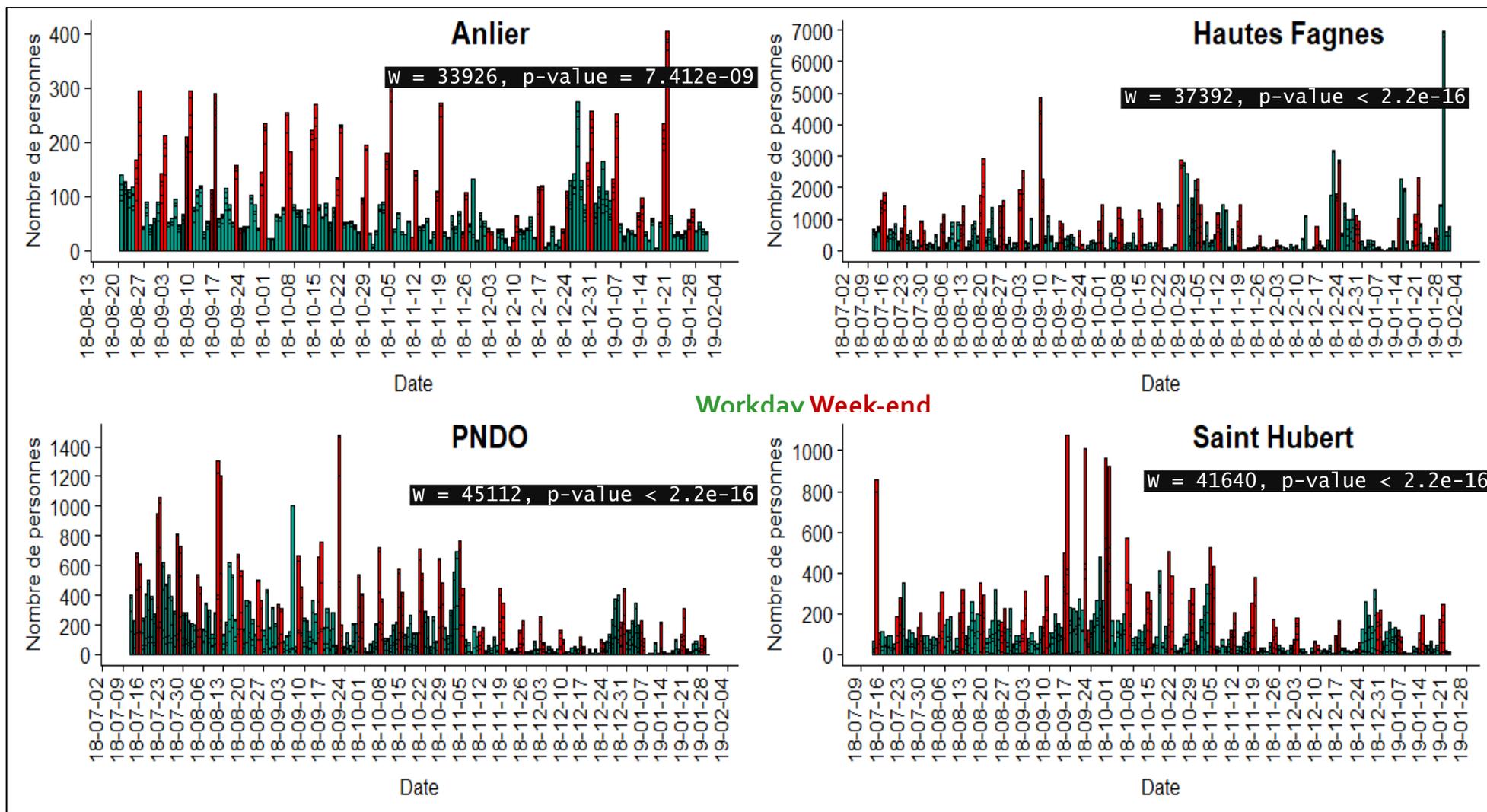
Concernant le Parc Naturel des Deux Ourthes, Cheslé et le Plateau des Tailles sont généralement les endroits les moins fréquentés. Leur maximum se situe la deuxième semaine d'août. C'est d'ailleurs la semaine la plus fréquentée pour l'ensemble des caméras de PNDO (sauf pour Hérou, puisque la caméra ne fonctionnait pas à ce moment). À elles deux, Cheslé et Hérou détectent en moyenne 39 personnes par jour contre 50 pour le Barrage de Nisramont. Hérou et Engreux sont les deux caméras les plus fréquentées de l'ensemble de la zone étudiée en dehors du massif des Hautes Fagnes. La fréquentation est maximale pour Engreux et Cheslé de mi-juillet à mi-août avant de décroître rapidement. Le Barrage de Nisramont continue à accueillir une quantité stable de visiteurs jusqu'au mois d'octobre.

Dans le massif d'Anlier, la caméra placée aux Etangs de la Fabrique se démarque nettement des autres. Premièrement, elle figure dans le top 7 des caméras avec la plus grande moyenne de détections quotidiennes alors que le massif d'Anlier ne regroupe que 6 % des visiteurs de l'ensemble de la zone d'étude. À elle seule, la caméra des Etangs de la Fabrique concentre 71 % de toutes les photos de personnes prises à Anlier. Cette importance semble persister au fil des mois. Comme cela sera décrit en détails plus tard, la caméra de Fagne Jean Simon est dans le top 2 des caméras prenant le plus de photos par jour avec 209 photos en moyenne. Pourtant, le Graphique 6 montre que seulement 10 personnes sont détectées par jour en moyenne par cette caméra.



Graphique 7: Variation du nombre de personnes par caméra et par semaine sur l'ensemble de la période

Résultats



Graphique 8: Nombre de personnes par jour et par massif sur l'ensemble de la période. Les barres rouges représentent les week-ends, les barres vertes sont les jours de la semaine

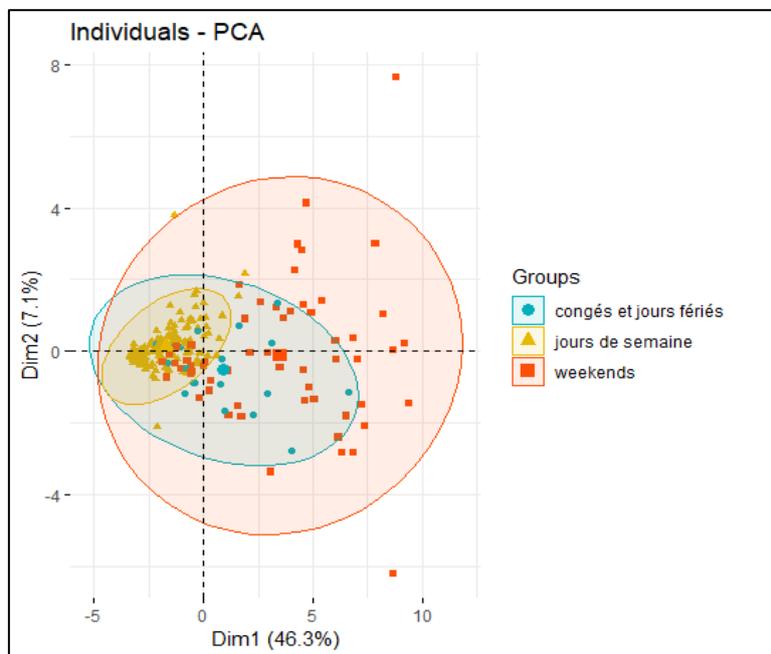
Résultats

La grande forêt de Saint-Hubert possède son pic de fréquentation à la moitié du mois de septembre. Cela est particulièrement marqué pour les caméras de Priesse et de Bilaude. Beyoli, Diglette Fourneau et Masblette Pont Mauricy reçoivent leur nombre maximum de visiteurs durant la deuxième moitié du mois d'octobre. Cependant, l'amplitude de ces visites est faible comparée aux deux premières caméras. Durant la première semaine de juillet, Masblette Pont Mauricy reçoit exceptionnellement plus de 600 visiteurs.

De manière générale, les pics de fréquentation touristiques varient entre mi-juillet et mi-octobre selon les massifs et les caméras. À partir de mi-novembre, le nombre de visites chute pour l'ensemble des caméras avant de réaugmenter subitement durant les deux dernières semaines de l'année. Ensuite, le faible taux de visites se maintient au moins jusqu'à la moitié du mois de janvier. Comme pour le nombre de photos, il y a significativement plus de détections de personnes durant les week-ends que pendant la semaine pour chacun des massifs.

3) Relations entre caméras

L'analyse des relations entre caméras suit la même logique que lors de l'étude du nombre de photos (p.40). La matrice de corrélation et l'ACP globale ont été placées en Annexe 7 car ils ne permettaient pas de distinguer des résultats suffisamment clairs. Cependant, placer les occurrences sur le 1^{er} plan factoriel en les distinguant en trois groupes, permet d'obtenir le Graphique 9. 46,3 % de la variance est expliquée par le 1^{er} axe. De plus, il semble que la variation de la fréquentation augmente en fonction de la période. En semaine (n=128), les variations sont faibles et les gens se comportent de façon régulière, durant les congés et les jours fériés (n=24), cette variabilité augmente et est maximale durant les week-ends (n=52).



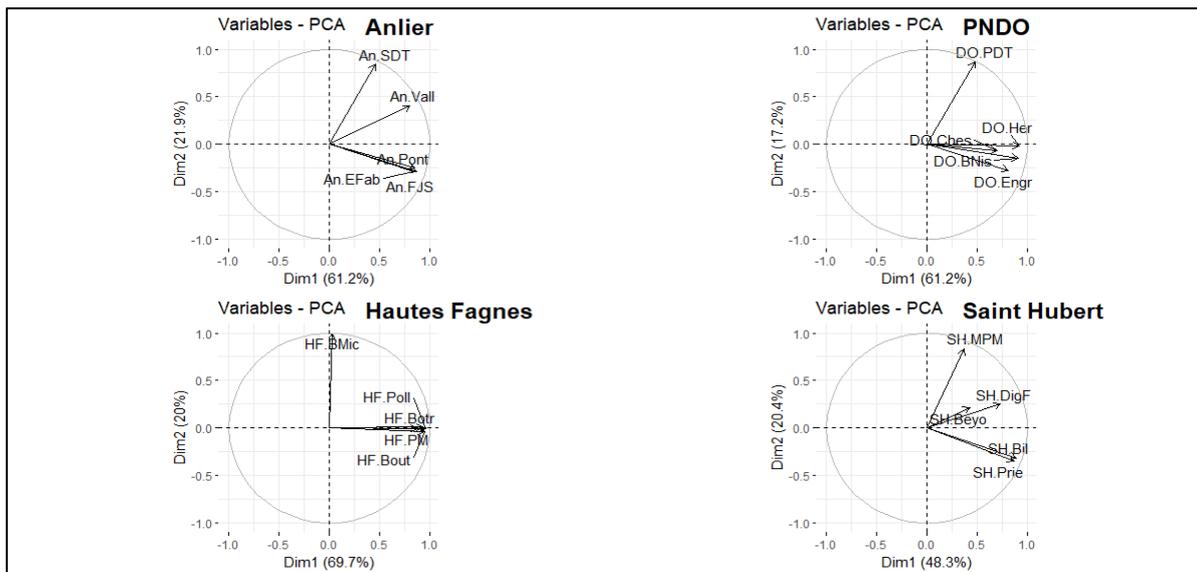
Graphique 9: 1er plan factoriel de l'ACP représentant les lignes de la matrice Caméras/Dates

Une ACP par massif est réalisée afin d'affiner les résultats (Graphique 10). Encore une fois, l'ensemble des caméras sont représentées vers les valeurs positives du 1^{er} axe. Il explique une grande partie de la variance dans chaque cas (entre 48 et 70 %). L'axe 2 permet de

Résultats

distinguer des groupes au sein de chaque massif. Dans le massif d'Anlier, les caméras Pont, Etangs de la Fabrique et Fagne Jean Simon semblent très liées et s'opposent aux caméras de la Vallée et du Stand de Tir.

Au PNDO, le Plateau des Tailles se distingue des autres caméras qui semblent très liées entre elles, même si Cheslé et Hérou ne sont pas vraiment impactées par le 2^{ème} axe et ne s'opposent pas directement au plateau des Tailles. La distinction dans le massif des Hautes Fagnes est la plus flagrante. La Baraque Michel est quasiment entièrement indépendante du reste des caméras qui sont extrêmement corrélées entre elles. Finalement, le massif de Saint-Hubert montre une corrélation nette entre les caméras de Bilaude et Priesse, qui s'opposent à Masblette Pont Mauricy et dans une moindre mesure à Diglette Fourneau. Beyoli est moins bien représenté dans ce 1^{er} plan factoriel.



Graphique 10: Analyse en composante principale. Evolution du nombre de personnes capturées par chaque caméra en fonction du temps et par massif.

4) Etude de la taille des groupes

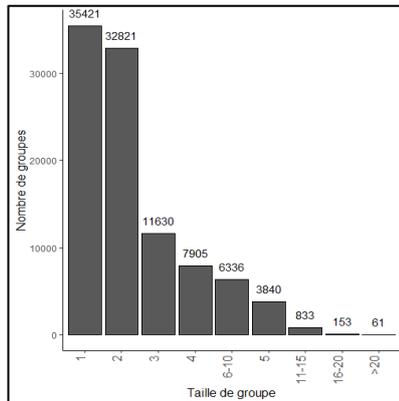
Etant donné que le modèle n'est pas encore assez précis pour étudier les moyens de locomotion ou le matériel emporté par les visiteurs, l'aspect qualitatif de l'indice est mis en avant grâce à l'étude de la taille des groupes qui passent devant les caméras. Comme expliqué p. 32, les personnes considérées comme faisant partie d'un même groupe après analyse visuelle se retrouvent sur la même photo dans 78 % des cas. Dans le cadre de ce document, les groupes sont donc considérés comme le nombre de personnes sur une même photo.

Le Graphique 11 décrit les quantités des différentes tailles de groupes formés sur l'ensemble de la zone d'étude. Plus de 55 % des personnes se déplacent seuls (14 %), en couple (27 %) ou à 3 (14 %). Une majorité de personnes font donc partie de la catégorie «groupe de 2».

Afin de tenter de mieux comprendre les modalités de leur formation, des analyses factorielles de correspondance (AFC) ont été réalisées. Ces AFC sont issues de matrices de contingence créées entre les différentes tailles de groupes et les autres modalités étudiées. Les valeurs dans chaque table sont relatives. En raison de la distribution en j inversé des données (« left skewed »), une transformation logarithmique ($\text{Log}(x+1)$) est testée pour chaque AFC. Si

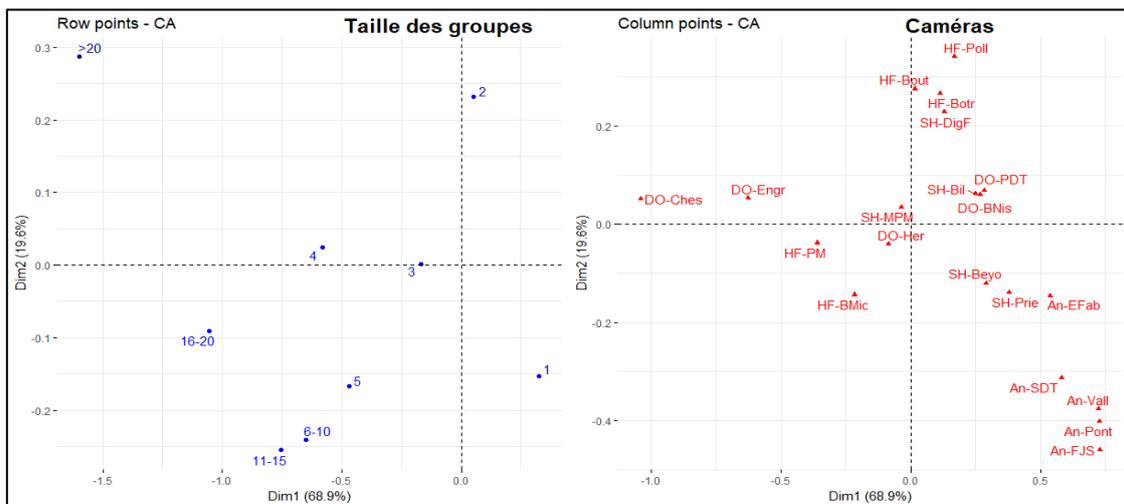
Résultats

aucune différence majeure n'est détectée, les données brutes sont conservées pour réaliser l'analyse. Il est très important de vérifier les tendances des biplots en allant parcourir les tables de données brutes.



Graphique 11: Répartition du nombre de groupes selon leur taille

Le Graphique 12 représente une AFC sur les données brutes de la matrice de contingence créée entre les différentes tailles de groupes et les caméras. La classe « plus de 20 personnes » a été traitée en temps qu'outlier et ne participe pas à l'étirement des axes (Bendixen, 2003). Les caméras sont réparties sur les quatre quadrants du graphique. Toutes les caméras du massif d'Anlier sont situées dans le deuxième quadrant, là où les personnes se déplacent seules selon le graphique de gauche. Cette tendance pourrait aussi se vérifier pour les caméras de Beyoli et Priesse (Saint-Hubert). Après vérification des données brutes, cette tendance est plutôt située sur Priesse et Bilaude. Les grands groupes sont situés dans le troisième quadrant. On y retrouve également les deux caméras les plus fréquentées de toute la zone d'étude : Baraque Michel et Pont Marie (Hautes Fagnes). Angulairement, il est tentant de relier les personnes se déplaçant en couple aux caméras de Polleur, La Béole, Botrange (Hautes Fagnes) et Diglette (Saint-Hubert). Néanmoins, leurs valeurs quant au premier axe sont faibles. Elles ont donc un faible impact. Le même constat peut être fait pour les caméras de Cheslé et d'Engreux. De manière



Graphique 12: Biplots représentant les correspondances entre les tailles des groupes et les caméras. À gauche, les différentes tailles de groupes sont représentées, à droite, les 20 caméras

Résultats

générale, il semblerait que les personnes seules et les couples tirent l'axe vers la droite, alors que les groupes à partir de 3 personnes le tirent vers la gauche.

Le procédé a été répété en fonction des jours de la semaine et des mois de l'année, mais ne sont pas très concluants. Les AFC sont consultables en Annexe 8.

III. Liens entre le nombre de photos et la fréquentation

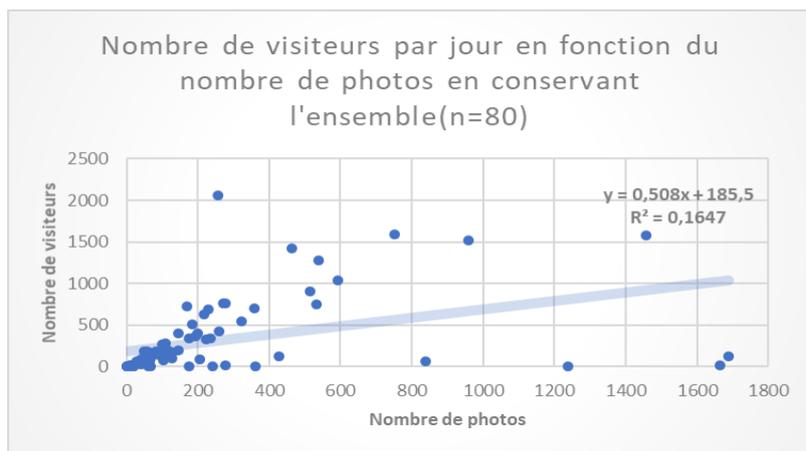
Cette partie vise à comparer le nombre de personnes comptabilisées en fonction du nombre de photos prises par chaque caméra. L'intérêt est de déceler les caméras mal configurées, sujettes aux faux déclenchements ou encore les caméras qui capturent plus de personnes par événement de détection que les autres. De plus, elle permet de mettre en lumière les biais éventuels liés aux extrapolations tels que ceux liés à la méthode de Lupp et al. (2016).

1) Comparaison avec la méthode de Lupp et al. (2016)

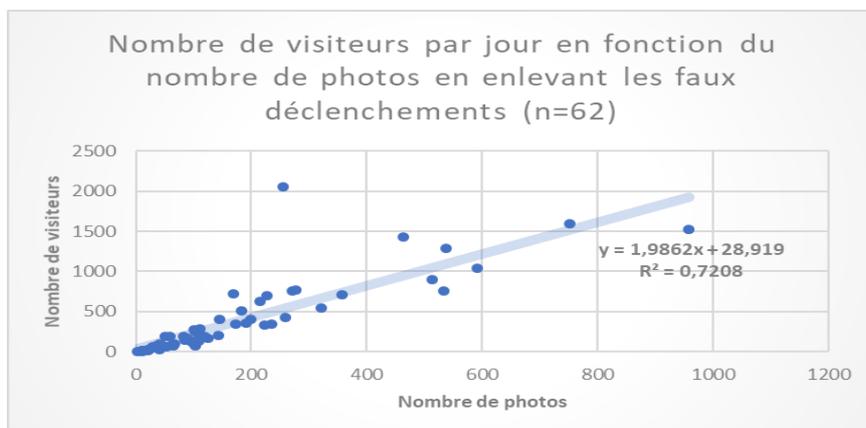
Les données ayant servi à créer les différentes régressions sont disponibles en Annexe 9. La méthode est expliquée p.34. Trois régressions linéaires ont été réalisées pour ce test. La première (Graphique 14) est issue de l'ensemble des données testées, c'est-à-dire les 80 journées échantillonnées aléatoirement (4 par caméra). Finalement, 18 journées présentent des problèmes de déclenchement et diminuent fortement la qualité de la régression qui n'est expliquée qu'à 16%. En les supprimant, on obtient une précision 4,5 fois meilleure pour atteindre 72 % (Graphique 15). Une journée exceptionnelle s'est déroulée dans le sous-échantillon, sur 250 photos, il y avait plus de 2.000 personnes. Enlever cette journée permet d'atteindre 88 % de corrélation (Graphique 13).

Finalement, le Tableau 7 montre les estimations faites par chaque régression ainsi que leur précision. L'ensemble est comparé aux résultats fournis par l'indice automatisé par l'algorithme.

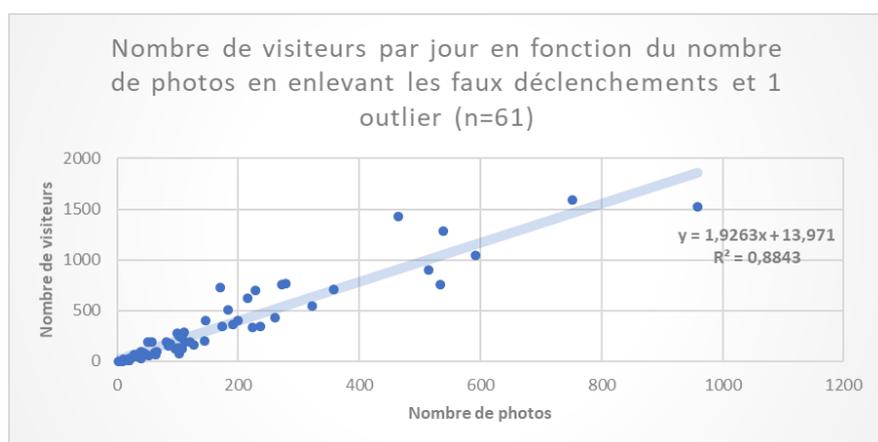
Résultats



Graphique 14: Régression linéaire évaluant la corrélation entre le nombre de visiteurs journaliers sur l'ensemble des caméras et le nombre de photos prises par jour



Graphique 15: Régression linéaire évaluant la corrélation entre le nombre de visiteurs journaliers sur l'ensemble des caméras et le nombre de photo en enlevant les photos faussement positives



Graphique 13: Régression linéaire évaluant la corrélation entre le nombre de visiteurs journaliers sur l'ensemble des caméras et le nombre de photo en enlevant les photos faussement positives et 1 outlier

Résultats

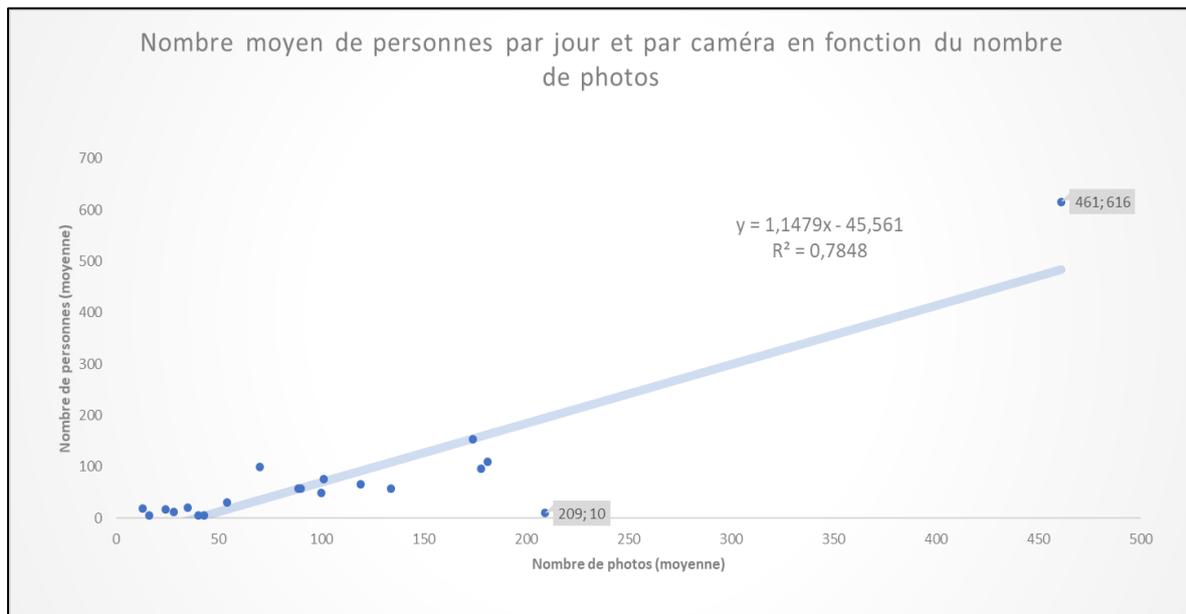
Tableau 7: Nombre de visiteurs total sur l'ensemble des caméras en fonction de la régression utilisées et par rapport à l'indice automatisé par l'algorithme

Pour 356.200 photos	R ²	Nombre de visiteurs estimés (arrondi)
Graphique 14	0,16	181.000
Graphique 15	0,72	708.000
Graphique 13	0,88	687.000
Mask R-CNN		247.000

Il apparait finalement qu'utiliser l'ensemble des données produit une sous-évaluation par rapport à l'estimation du modèle. Le rapport entre le nombre de personnes et la quantité de photos étant plus faible. En revanche, plus on supprime des événements indésirables, plus la qualité de la régression augmente. Cette augmentation de la qualité entraîne une très forte surestimation du nombre de visiteurs. Aucune régression ne se rapproche donc de l'estimation de Mask R-CNN.

2) Comparaison entre le nombre de photos et le nombre de personnes total

Le Graphique 16 reprend les informations moyennes fournies par les tableaux de résultats du nombre de photos (Tableau 4) et du nombre de personnes (Tableau 6). Même en conservant les caméras exceptionnelles, le coefficient de détermination reste à 79 %. En enlevant la caméra de

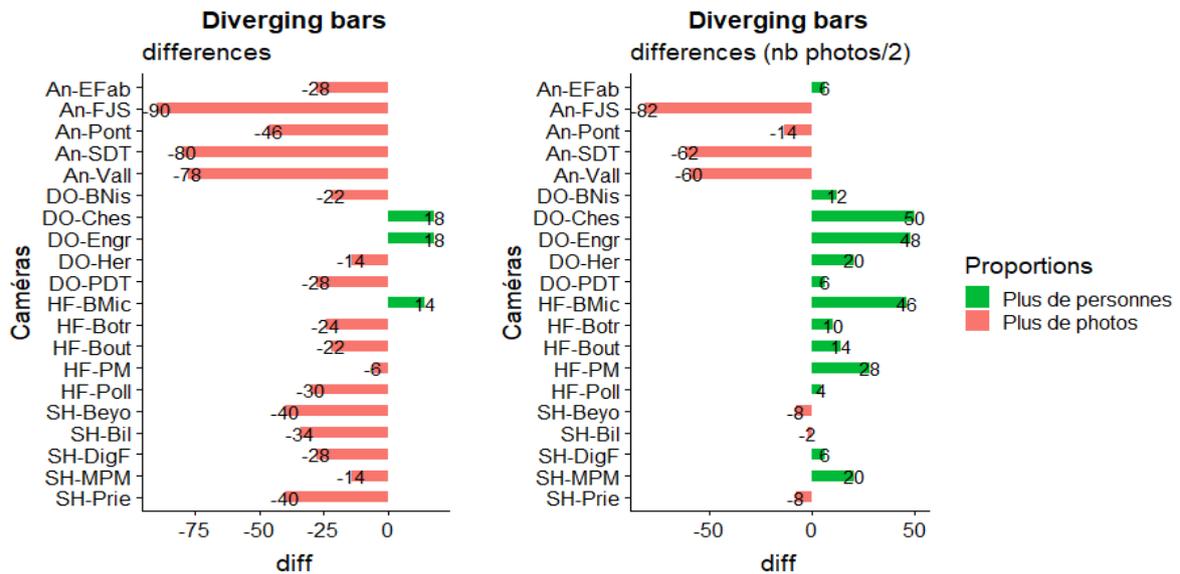


Graphique 16: Nombre moyen de personnes par jour et par caméra en fonction du nombre de photos. Chaque point représente une caméra

Fagne Jean Simon (209 photos par jour pour seulement 10 personnes), le R² atteindrait 90 %. Il faut noter que le nombre de personnes est calculé après diminution de la redondance dans la base de données.

Résultats

Le Graphique 17 décrit ces divergences entre le nombre de photos et le nombre de personnes. Pour construire ce graphique, les nombres moyens de photos et de personnes par caméras ont été transformés en proportions relatives (%photos + %personnes= 100% pour chaque caméra). Puisque chaque événement de détection prend deux photos, diviser le nombre de photos par deux permet de différencier les photos prises sans personnes des doublons.



Graphique 17: Graphique représentant les divergences entre le nombre moyen de photos prises et le nombre moyen de personnes comptées sur chaque caméra. À gauche, toutes les photos sont conservées. À droite, le nombre de photos est divisé par deux car chaque événement de détection correspond à deux photos.

Sur l'ensemble des caméras sauf trois, il y a plus de photos que des personnes comptées. Cette tendance est particulièrement marquée sur quatre caméras d'Anlier (toutes sauf les Etangs de la Fabrique). Cette disproportion est également marquée à Saint-Hubert. Concernant les trois caméras qui révèlent plus de personnes que de photos, deux se trouvent à PNDO (Cheslé et Engraux) et la dernière en Hautes Fagnes (Baraque Michel). Lorsque le nombre de photos est divisé par deux (2 photos par événement de détection), seules quatre caméras d'Anlier et dans une moindre mesure trois caméras de Saint-Hubert restent dans le négatif. Cela signifie que dans la majorité des cas il y a plus de photos que de personnes, mais moins d'évènements de détection.

Discussion

La discussion est séparée en trois parties. Tout d'abord, les résultats de fréquentation touristique du cas d'étude (AGRETA) sont discutés. Cette partie ne constitue pas le but premier du travail, mais permet de mettre en avant un éventail d'avantages et d'inconvénients liés à l'indice développé. Ensuite, la précision de l'indice est comparée à la méthode de Lupp et al. (2016). Finalement, la qualité des résultats est discutée. Dans cette dernière partie, plusieurs critiques et pistes d'amélioration sont détaillées. Cet ordre ne correspond pas à l'ordre de présentation des résultats. Il permet de mettre en avant les observations liées à la mise en pratique avant de discuter d'améliorations théoriques.

I. La fréquentation touristique en Ardenne

L'analyse de la fréquentation touristique en Ardenne sert d'exemple d'utilisation de l'indice dans un cas d'étude. Elle est donc relativement simple puisqu'elle ne fait pas partie de l'objectif principal du travail.

Concernant le comptage des personnes, certaines caméras se situent sur le même parcours. Le nombre total de personnes comptées est donc à nuancer, car les personnes peuvent être comptabilisées sur plusieurs caméras. En revanche, la fréquentation de chaque caméra individuellement est plus révélatrice. Les résultats sont décrits par massif avant de mettre en avant les tendances générales.

Réserve naturelle des Hautes-Fagnes Eifel

La réserve naturelle des Hautes-Fagnes Eifel présente la plus grande quantité de photos pour quatre de ses caméras, dont Polleur qui ne se situe pas sur le même circuit que les trois autres. En revanche, La Béole a obtenu moins de photos alors qu'il se situe sur le même circuit que les trois premières. Cela peut s'expliquer par le nombre de points d'accès élevé sur le massif qui fait que l'ensemble des promeneurs ne suivent pas tous le même circuit. Atteindre La Béole nécessite de marcher plus de huit kilomètres.

La Baraque Michel est toujours l'endroit le plus photographié. Quand elle ne l'est pas, c'est quasiment toujours à cause d'un problème technique. Elle est aussi l'endroit le plus visité de l'étude en accueillant presque 31 % des visiteurs. Le Pont Marie se trouvant sur le même chemin, mais accueille seulement 13 % des visiteurs. Ce qui signifie qu'un peu moins de la moitié des visiteurs des Hautes Fagnes suivent cette promenade. L'entièreté du massif accueille deux tiers des visiteurs photographiés sur l'ensemble de la zone d'étude. Cela confirme les résultats de la thèse de Colson (2009) qui définit le plateau des Hautes Fagnes et les vallées ardennaises comme les principaux pôles touristiques. À une échelle plus fine, les ACP montrent que les caméras des Hautes Fagnes sont très liées entre elles sauf la Baraque Michel qui est indépendante. On peut supposer que les touristes qui viennent dans les Hautes Fagnes veulent principalement voir les tourbières. La caméra de la Baraque Michel étant proche de la route et accessible via un caillebotis, elle se trouve vraisemblablement à un endroit préférentiel.

L'analyse de la taille des groupes semble révéler que les groupes entre 5 et 15 personnes sont le plus souvent observés à la Baraque Michel. Les groupes de 16 à 20 personnes semblent être associés au Pont Marie. En réalité, les données brutes montrent que la Baraque Michel

Discussion

accueille majoritairement toutes les tailles de groupe, sauf les groupes de 16 à 20 personnes qui se trouvent plus souvent à Pont Marie et à Engreux (PNDO) et les groupes de plus de 20 personnes.

La notoriété touristique du lieu pourrait effectivement attirer les visiteurs venant de loin et en groupe. Cependant, puisque les groupes ont été définis comme les personnes apparaissant sur la même photo, la position des caméras de la Baraque Michel et du Pont Marie sont biaisées, car elles peuvent capturer plus de personnes sur une même image. Le Graphique 28 montre également qu'il y a plus de personnes que des photos pour la Baraque Michel. Cela pourrait signifier que même si les groupes sont plus petits à cet endroit, ils sont plus réguliers.

La Grande Forêt de Saint-Hubert

À Saint-Hubert, Priesse et Bilaude accueillent la majorité des visiteurs. Ils s'écartent également des autres caméras dans l'ACP. Ces deux sites comportent des miradors d'observation très fréquentés lors du brame. L'attractivité principale de ce massif est due à sa densité de gibier et les possibilités d'observation. De plus, les résultats montrent que les pics de visite se situent entre fin septembre et début octobre, ce qui correspond à la période du brame du cerf.

Les sites de Masblette-Pont-Mauricy, Diglette et Beyoli semblent peu visités. Masblette-Pont-Mauricy possède un pic d'observation durant la deuxième semaine de juillet. Après vérification, il s'agit de personnes ayant placé leur table de pic-nic sous la caméra. Ils sont restés toute l'après-midi (de 14h à 17h30) devant la caméra. Même après diminution de la redondance, leur impact est encore fort. Diglette, Masblette-Pont-Mauricy et Beyoli se situent sur des routes. Leur groupement sur l'ACP ainsi que leur ratio nombre de personnes sur nombre de photos peut s'expliquer par le déclenchement répétitif des caméras par les véhicules (voitures, grumiers). Cette différence entre le nombre de personnes détectées et le nombre de photos est également visible sur le Graphique 28. Particulièrement sur ce massif, un nombre important de déclenchements peut aussi être dû au gibier.

Le Parc Naturel des Deux Ourthes

Comme pour les Hautes Fagnes, PNDO semble être principalement visité par des touristes (Colson, 2009). Engreux et le Barrage de Nisramont peuvent éventuellement être parcourus par les mêmes personnes lors d'une randonnée. Ces deux caméras sont d'ailleurs très liées en termes de nombre de photos. Lorsqu'on étudie le nombre de personnes, toutes les caméras sont colinéaires sauf le Plateau des Tailles qui réagit de manière indépendante aux variations. Cela peut s'expliquer par le fait que cette caméra est excentrée par rapport aux quatre autres.

Le site d'Engreux semble être le plus fréquenté et présente même plus de personnes que de photos. Ces proportions sont également observées pour le site de Cheslé qui est le moins fréquenté de PNDO. Cela signifie que ces sites attirent vraisemblablement les groupes plutôt que les personnes seules. Il ressort également de l'ACP que ces deux caméras sont plus susceptibles d'accueillir des groupes de quatre personnes. Cette tendance est vraie pour Cheslé. Par contre, Engreux a tendance à accueillir les groupes de plus grande taille dont la grande majorité des groupes de plus de 20 personnes. Encore une fois, l'angle de la caméra joue un grand rôle dans ce résultat. Le fait que Cheslé accueille si peu de visiteurs pourrait être dû au dénivelé important de cette partie du massif. Selon Colson (2009), les visiteurs préfèrent des chemins plats ou de légères pentes.

Le site d'Hérou possède de nombreux points d'accès et chemins qui se croisent, il est donc très imprécis de résumer ce site à une seule caméra. Le barrage de Nisramont présente un grand risque d'aller-retour en raison des personnes qui passent devant la caméra pour aller voir

Discussion

un point de vue et puis revenir à leur véhicule. De ce fait, la fréquentation du premier pourrait être sous-estimée et celle du second surestimée.

La Grande Forêt d'Anlier

Anlier est de loin le massif le moins fréquenté de la zone d'étude. Même si Fagne Jean Simon était la deuxième caméra ayant pris le plus de photos, elles se sont révélées être de faux positifs en raison de branches se déplaçant devant l'objectif. De manière plus générale, les caméras d'Anlier semblent récolter un grand nombre de faux positifs ou des photos que le modèle ne parvient pas à classifier. Cela peut être dû à plusieurs choses. Concernant Vallée et Stand de Tir, il semblerait que la caméra soit plus dans une zone ensoleillée, ce qui impacte fort la température de l'arrière-plan. De plus, la caméra est légèrement trop perpendiculaire au chemin. Cela combiné au fait que les sites soient fortement fréquentés par des cyclistes, ceux-ci passent peut-être trop vite par rapport au temps de latence. Concernant Pont, la caméra pourrait se déclencher lorsqu'il y a du mouvement derrière l'arbre à droite du cadre. En revanche, le modèle pourrait avoir du mal à classifier les personnes en raison de l'occlusion créée par les branches.

Le forêt d'Anlier possède peu d'infrastructures, elle est donc moins fréquentée par les touristes. Il pourrait s'agir d'une forêt plutôt fréquentée par les locaux. Les locaux fréquentent généralement la forêt pour des durées plus courtes et plus fréquemment pour des activités sportives (cyclisme, jogging) et pour promener leur chien (Rossi et al., 2015; Lupp et al., 2016). Cela correspond aux résultats de l'AFC, pour laquelle les différentes caméras d'Anlier semblent être liées aux personnes se déplaçant seules. Cette supposition se confirme à la vue des données brutes.

Général

A l'échelle régionale, différentes tendances ressortent également. Premièrement, le premier axe de toutes les ACP ne contient que des valeurs positives. Cela signifie que la fréquentation à l'échelle régionale est plutôt corrélée entre les différentes caméras, même si l'intensité des variations peut différer.

Les pics de fréquentation se situent majoritairement durant le milieu de l'été, durant les mois de juillet et d'août. La fréquentation diminue graduellement à partir de la mi-septembre jusqu'au mois de janvier (sauf pour Saint-Hubert). Un pic de fréquentation se déroule cependant durant la dernière semaine d'octobre pour tous les massifs ainsi que durant les deux dernières semaines de l'année. Il s'agit des vacances scolaires de la Toussaint et de Noël.

Même si seules les personnes ont été étudiées dans le cadre de ce travail, un nouvel avantage imprévu des pièges caméras par rapports aux écompteurs ressort. Lorsque des événements particuliers sont décelés dans les données de fréquentation (comme à Pont Marie et à Masblette-Pont-Mauricy), il est possible d'aller revoir les images à la date et au lieu donné afin de déterminer la raison de cet événement. Il est alors possible de mettre un nom sur ces événements exceptionnels.

Concernant la fréquentation en fonction de la date, il semblerait que les visiteurs adoptent un comportement régulier durant la semaine, le nombre de visiteurs dans la région est relativement constant. Il pourrait davantage s'agir des locaux s'adonnant à leurs activités habituelles. Durant les congés et les jours fériés, la variabilité augmente et devient maximale durant les week-ends. Il est possible que l'effet de saisonnalité soit plus impactant durant que les week-ends que durant la semaine. Dès lors, la période de l'année jouerait fortement sur les gens venant de plus loin pour profiter de certains espaces naturels.

II. Comparaison à la méthode de Lupp et al. (2016)

Au vu des résultats, il est donc possible d'avoir une bonne corrélation entre le nombre de photos et le nombre de visiteurs. Cependant, cela nécessite de nettoyer le jeu de données afin d'y enlever les déclenchements redondants ainsi que les outliers éventuels. La conservation de l'ensemble des photos entraîne une sous-estimation du nombre total de personnes. De plus, les faux déclenchements font grandement diminuer la qualité de la régression. Si on enlève ces faux déclenchements, une surestimation significative apparaît. Ce résultat est logique puisque l'ensemble des 356.200 photos prises contient beaucoup de déchets. Le but étant de pouvoir réaliser un comptage des personnes sans passer par le tri de ces données, il faudrait donc théoriquement se baser sur le premier graphique qui est très imprécis ($R^2 = 16\%$).

La méthode du subsampling ne permet donc de se servir efficacement de la méthode des caméra trap que dans certains cas de figure. La corrélation doit être assez élevée entre le nombre de photos et le nombre de visiteurs. Considérant l'autre aspect important du développement de cet indice qui est l'aspect qualitatif, le comptage d'attributs tels que des vélos, des sacs à dos ou des chiens, la corrélation entre le nombre de photos et leur occurrence serait vraisemblablement beaucoup moins élevées que pour le comptage des personnes.

Finalement, la corrélation entre l'ensemble des personnes comptées par le modèle et le nombre de photos prises par chaque caméra est bonne. Il faut cependant noter que le nombre de personnes comptées sur chaque caméra constitue un pas de temps d'au moins 20 secondes entre des photos de personnes.

III. Qualité des résultats

Avant toutes choses, il faut préciser que l'évaluation de la qualité du modèle par rapport au dataset ne dépend pas uniquement de celui-ci, mais de la chaîne de production de données en entier. La disposition des caméras et les difficultés liées à celles-ci impactent donc autant voire plus les résultats que l'algorithme lui-même. Une critique de ces deux composantes est donc nécessaire.

1) Configuration des caméras

La position des caméras est très hétérogène dans le cadre de l'étude d'AGRETA. Un positionnement plus régulier selon des règles fixes devrait pouvoir donner des résultats plus stables et plus comparables entre les différentes caméras. Ces règles sont décrites dans les paragraphes ci-dessous et doivent affecter d'une part la qualité des photos prises sur le terrain, mais aussi faciliter la détection et la classification des objets par l'algorithme.

Flouttage

Premièrement, travailler avec des images nettes devrait accentuer les différences entre pixels voisins et permettre à l'algorithme de mieux détecter les différents patterns. Généralement, les réseaux neuronaux réagissent mal au flou et au bruit, alors qu'ils résistent à la compression et aux contrastes (Dodge et al., 2016). Pour cela, réaliser une analyse d'impact sur la protection des données dans le cadre du RGPD pourrait permettre d'obtenir l'autorisation

Discussion

d'utiliser les caméras sans le floutage par bande adhésive¹⁰. Cela devrait aussi diminuer certaines causes de non-détection, en particulier celles liées aux conditions météo qui sont accentuées par l'humidité qui se loge entre la lentille de la caméra et la bande adhésive. Il est aussi possible que l'exposition lumineuse soit moins gênante sans cette bande.

Exposition lumineuse et végétation

L'exposition lumineuse crée des problèmes à deux niveaux : à la prise de la photo par la caméra et lors de la classification. Pour la prise des photos, les problèmes dus à l'environnement se posent dans beaucoup d'études qui utilisent des caméras traps. Towerton et al. (2008) ont récolté 98 % de faux positifs dans une étude en savane à cause du vent qui faisait bouger les herbes. Fairfax et al. (2014) ont aussi compté 45 % de faux positifs dus à l'environnement. Selon Wearn et al. (2017), le mouvement des plantes peut déclencher la caméra seulement après que celles-ci aient été chauffées par les rayons du soleil. C'est cette différence de température qui déclenche la photo. Concrètement, il existe trois façons de minimiser ce problème. Premièrement, acheter des caméras de meilleure qualité pourrait diminuer ces faux déclenchements, en supposant que les composants du capteur infrarouge soient moins sensibles. Il est aussi possible de diminuer la sensibilité de la caméra. Cependant, cela augmente le risque de faux négatif. Wearn et al. (2017) conseillent alors de placer la caméra dans un lieu ombragé, de façon à diminuer l'impact des rayons solaires sur la végétation mouvante.

Positionnement

Les caméras photographiant les personnes arrivant face à elles sont à proscrire (Baraque Michel, Pont, Pont Marie). En plus de favoriser l'effet de superposition des personnes qui rend la détection de celles-ci difficile par le modèle, cela place les personnes trop loin de la caméra lorsqu'elles s'éloignent (ou avant de s'approcher) et les rend indétectables. Cela favorise aussi la détection multiple de mêmes individus puisqu'ils restent alors plus de longtemps devant la caméra. De plus, la détection des sacs à dos dépend majoritairement de la position de la personne. Un sac à dos sur une personne venant face à la caméra aura très peu de chances d'être détecté puisque seules les brides du sac sont alors visibles.

Les meilleurs ratios de détection sont assignés aux caméras placées à angle d'environ 45° à 60° par rapport au chemin et pas trop en hauteur. Cela rend difficile l'identification de grands groupes sur une seule photo, car les personnes ne se trouvent pas toutes sur la même image. Néanmoins, il serait plus plausible de considérer les images redondantes comme des groupes espacés entre les événements de détections puisque les personnes prendraient moins de temps pour passer devant les caméras. Le chemin choisi doit aussi être simple, il faut éviter les croisements et les bifurcations.

Il faut aussi veiller à ne pas placer la caméra où des gens sont susceptibles de s'arrêter (point de vue, banc, souche, ...). Cela augmente le risque que les personnes stagnent devant la caméra et soient capturés des centaines voire des milliers de fois. Même s'il est possible grâce aux métadonnées de supprimer les prises consécutives si elles sont temporellement trop proches, il demeure très compliqué de différencier les personnes stagnant devant les caméras et les groupes éparpillés sur plusieurs événements de détection.

De manière générale, l'Annexe 10 rassemble les qualités et les défauts de chaque caméra. Le critère de l'angle de la caméra et du lieu (la simplicité du chemin) reste le plus important. Les caméras trop en hauteur peuvent gêner la classification des objets par le modèle. Il est

¹⁰ Dossier en cours à la date de dépôt de ce document (16 août 2019)

Discussion

préférable de les mettre plus en hauteur et de les protéger grâce à des cages ou des cadenas (Meek et al., 2013; Wearn et al., 2017).

2) **Qualité du modèle**

Etant donné la variabilité trop importante en termes de détection pour les sacs à dos, les chiens et les deux roues et en termes de classification pour les véhicules, il semble que les prendre en compte dans l'évaluation de l'indice soit à l'encontre de l'objectif de précision recherché. Les différents paramètres qui peuvent améliorer le processus sont détaillés ci-dessous.

Entraînement du modèle

En analysant le dataset utilisé pour entraîner le modèle (Lin et al., 2014), il est logique que les personnes soient les objets les mieux déterminés. 250.000 personnes ont été identifiées sur 66.800 photos dans ce dataset. En revanche, seulement 12.800 photos comprenant des véhicules (voiture, train, bus, camion), 5.800 comprenant des sacs à dos et 4.600 comprenant des chiens s'y trouvent. Concernant les vélos, les résultats sont plutôt bons comparés au nombre de photos sur lesquelles cette classe a été entraînée (3.400). Les motos ne sont présentes que sur 470 photos. Selon Russakovsky et al. (2015), il faut au moins 1.000 instances d'une classe pour que le modèle commence à être efficace. Compte tenu de l'arrière-plan très dynamique des photos d'AGRETA (en forêt et en tourbière) et le floutage de celles-ci, il semble logique que ces classes soient relativement mal reconnues.

Concernant le modèle en tant que tel, la façon la plus logique de l'améliorer serait d'entraîner les réseaux neuronaux qu'il contient à partir de zéro et avec les données propres à ce type d'étude et pour les objets qui sont pertinents (personnes, sacs à dos, chiens, vélos, bâtons de marche, ...). Cela devrait être possible compte tenu du nombre de données acquises par classe. Cependant, cela nécessite cependant une grande expertise en informatique et surtout du temps pour étiqueter ces images afin d'entraîner le modèle. L'alternative serait d'entraîner uniquement les couches supérieures du modèle. Ces couches sont les plus générales entre les différents dataset. Dans la plupart des cas, conserver ces couches et réentraîner uniquement les couches les plus complexes du réseau sera un gain de temps favorable (Yosinski et al., 2014).

Segmentation

Comme décrit plus haut, l'exposition lumineuse et le vent faisant bouger la végétation sont les principales causes de faux positif lors de l'utilisation de caméras. S'il n'est pas possible de régler le problème à la source, la segmentation des images devrait fournir des solutions supplémentaires.

Gomez et al. (2017) ont comparé un dataset d'images segmentées manuellement à un dataset d'images originales issues de caméra trap. Ils ont montré que pour huit réseaux neuronaux différents, les résultats de classification d'images sont significativement supérieurs pour le dataset segmenté. De plus, lors de l'entraînement d'un réseau neuronal, la segmentation permet d'éviter que le réseau n'assimile des objets à leur arrière-plan. Des études montrent par exemple que la discrimination entre le loup et le husky peut être biaisée par la neige (Ribeiro et al., 2016) ou la détection de chèvres par les collines vertes (Shane, 2018).

Les techniques de segmentation sont multiples. La plus simple consiste à reprendre l'ensemble des images d'un événement de détection et à calculer la moyenne des pixels entre toutes les images. De ce fait, les éléments qui ont bougé pourront être détectés (Weinstein, 2015). Cependant, Weinstein travaillait sur des timelapse. Hoebeke (2018) a remplacé la

Discussion

moyenne par la médiane. De cette façon, elle conserve uniquement des valeurs de pixels qui existent vraiment sur l'image. Cette technique permet de créer une image du second plan qui peut être soustraite des images d'origine de la séquence. De ce fait, seuls les individus qui se déplacent restent. Le risque de faux positif et de détection d'objets indésirables serait alors considérablement diminué.

Des méthodes plus élaborées existent (Chen et al., 2014; Gomez Villa et al., 2017). Yousif et al. (2019) ont récemment publié le logiciel « Animal Scanner » pour séparer les images contenant des animaux de celles contenant des humains et des faux positifs. Ce logiciel dépasse les autres méthodes en termes de segmentation, mais aussi en termes de détection d'animaux et d'humains.

Superposition interclasse

Les vélos et les motos présentent une grande superposition interclasse. Fusionner les deux classes en une métaclasse est sûrement la solution la plus simple actuellement pour supprimer cette superposition (Liu, 2008). De plus, aucune moto n'a été détectée visuellement dans le dataset. Cette métaclasse « deux roues » pourrait donc correspondre aux vélos. Il est aussi possible d'isoler les classes qui se superposent du reste du problème. De ce fait, un algorithme spécifique peut résoudre ce problème plus simple avant de le réintégrer dans la classification générale (Cormen et al., 1990). Bao-Liang Lu et al. (1999) décrivent les principales variantes de cette méthode.

Dans ce travail, les 90 classes connues par Mask R-CNN ont été prises en compte. Une autre idée serait de limiter le modèle à la détection des objets d'intérêt et pas l'ensemble de ceux qu'il a appris. En effet, au-delà des zones d'ombre, les objets fixes sur l'image que le modèle connaît, tel que les bancs et les panneaux sont détectés de façon quasiment systématique sur les photos et augmentent inutilement la quantité de données.

Ce problème de superposition de classes est aussi très important dans le cas du chien. Etant un mammifère, l'ordinateur peut le confondre très facilement avec toutes les espèces de gibier quadrupède de la forêt. Les premiers à avoir tenté de classifier automatiquement les animaux sauvages sont Chen et al., (2014). Ils ont également rendu publique le dataset créé pour leur étude qui contient 20 espèces d'animaux d'Amérique du Nord. 14.300 photos pour l'entraînement du modèle et 9.500 pour la vérification. Cette approche a été améliorée par Gomez et al. (2017) et tous deux se sont servi de techniques de segmentation.

IV. Autres pistes d'optimisation de la méthode

Météo

Il est prouvé que la météo influence la fréquentation des espaces naturels. Par exemple, les cyclistes semblent être plus sensibles au mauvais temps que les personnes faisant leur jogging (Janowsky et al., 2003). Il est possible d'obtenir les relevés météorologiques directement de l'institut national (en Belgique, l'Institut Royal Météorologique) et de les coupler à la base de données conçue. Une autre alternative est possible. Puisque la caméra encode la température lors de chaque capture, il est possible d'extraire les données de température affichées sur la photo. Pour se faire, le package PyTesseract¹¹ peut être utilisé. Il s'agit d'un outil de reconnaissance de caractère qui peut donc identifier la température indiquée sur la photo et l'encoder dans un fichier texte.

Réidentification

Une façon intéressante de régler le problème de la distinction entre les personnes en groupe et les personnes stagnant devant la caméra serait d'intégrer à l'algorithme des mécanismes de réidentification des personnes (Xiao et al., 2017). De plus, cela pourrait permettre d'identifier les personnes faisant demi-tour ou passant devant plusieurs caméras. Il serait alors possible d'étudier les trajets préférentiels et les durées de promenade.

Adaptation des capteurs

La façon la plus efficace d'éviter les faux déclenchements pourrait être de se servir des techniques utilisées par les écompteurs pour déclencher les caméras. Puisque le but est de photographier des personnes et non des animaux, les capteurs infrarouges peuvent être troqué pour des détecteurs plus simples d'utilisation. Comme les personnes marchent sur les chemins et sont moins discrètes que les animaux, des dalles piézoélectriques ou les tubes pneumatiques pourraient être utilisées (Melville et al., 2004). Une pression sur ces dalles pourrait déclencher les caméras, ce qui devrait permettre d'éviter une grande majorité de faux déclenchements. Ces dalles peuvent être enterrées dans le sol et donc être discrètes. De plus, placer deux rangées de dalles pourrait donner une information sur le sens de déplacement des visiteurs sans devoir analyser les photos (Millhäusler et al., 2016). Il faut toutefois vérifier l'efficacité de cette méthode sur sol gelé (Ross, 2005).

Chiffrement

Concernant le chiffrement des images, il faut trouver le bon équilibre entre le degré de protection et donc la complexité de l'algorithme (exemple : algorithme AES), ou la rapidité du chiffrement de l'image (exemple : algorithme blowfish) (Thakur et al., 2011). Une grande partie des scripts disponibles sur GitHub¹² sont codés grâce à la librairie PyCrypto (Litzenberger, n.d.).

¹¹ <https://pypi.org/project/pytesseract/>

¹² <https://github.com/>

Conclusion

L'étude menée vise à mettre en place une nouvelle technique d'évaluation de la fréquentation touristique dans les milieux naturels. Améliorer ce type d'évaluation est crucial afin d'optimiser la gestion des milieux naturels et de les préserver des pressions anthropiques tout en améliorant leur attractivité. Les principaux manquements à combler sont la prise d'informations à long terme, la compréhension fine du comportement des visiteurs et une précision accrue des comptages. Le prototype actuel présente une bonne précision concernant le comptage des personnes sur des photos issues de caméras trap. Cependant, la configuration de ces caméras peut être grandement améliorée et l'architecture de traitement des images doit être davantage adaptée. L'étude montre qu'il est possible de rattraper le retard généré par le manque d'indicateurs pertinents concernant les services écosystémiques récréatifs. Grâce aux nouvelles technologies, les défauts majeurs liés à la prise de photos par caméra traps sont minimisés. L'indice développé peut déjà servir à rendre mieux compte de la valeur des services écosystémiques récréatifs sur le long terme. Cette mise en avant des activités touristiques ne doit pas se faire au détriment des services de production et de régulation. Elle doit permettre de mieux prendre en compte la multifonctionnalité des espaces naturels et les interrelations entre les services. Le placement des pièges caméras doit être homogénéisé afin de minimiser les biais entre les sites. De plus, l'algorithme utilisé doit être adapté à ce nouveau type de dataset. Les photos prises en extérieur sont complexes à analyser en raison de leur arrière-plan dynamique. Les principales tâches qui doivent désormais être entreprises sont la classification manuelle des photos afin de pouvoir entraîner les dernière couches du réseau neuronal ainsi que la mise en place d'un protocole de segmentation adéquat, afin de supprimer les images vides et de concentrer la classification sur les objets en mouvement. Enfin, puisque le but est de photographier des humains et non des animaux, les capteurs infrarouges des pièges caméras pourraient être remplacés par des capteurs plus adaptés, tels que des dalles à pression.

Références bibliographiques

- Abedi N., Bhaskar A. & Chung E., 2014. Tracking spatio-temporal movement of human in terms of space utilization using Media-Access-Control address data. *Applied Geography* **51**, 72–81.
- Ahas R., Aasa A., Roose A., Mark Ü. & Silm S., 2008. Evaluating passive mobile positioning data for tourism surveys: An Estonian case study. *Tourism Management* **29**(3), 469–486.
- Alegre J. & Garau J., 2010. Tourist satisfaction and dissatisfaction. *Annals of Tourism Research* **37**(1), 52–73.
- Alexandre L., 2017. *La guerre des intelligences*, 250.
- Arnberger A., 2006. Recreation use of urban forests: An inter-area comparison. *Urban Forestry & Urban Greening* **4**(3–4), 135–144.
- Arnberger A., Haider W. & Brandenburg C., 2005. Evaluating Visitor-Monitoring Techniques: A Comparison of Counting and Video Observation Data. *Environmental Management* **36**(2), 317–327.
- Arnberger A. & Hinterberger B., 2003. Visitor monitoring methods for managing public use pressures in the Danube Floodplains National Park, Austria. *Journal for Nature Conservation* **11**(4), 260–267.
- Bambi G. & Iacobelli S., 2017. Study and Monitoring of Itinerant Tourism along the Francigena Route, by Camera Trapping System. *Almatourism* (6), 144–165.
- Bao-Liang Lu & Ito M., 1999. Task decomposition and module combination based on class relations: a modular neural network for pattern classification. *IEEE Trans. Neural Netw.* **10**(5), 1244–1256.
- Bartczak A., 2015. The role of social and environmental attitudes in non-market valuation. *Forest Policy and Economics* **50**, 357–365.
- Bartczak A., Englin J. & Pang A., 2012. When are Forest Visits Valued the Most? An Analysis of the Seasonal Demand for Forest Recreation in Poland. *Environ Resource Econ* **52**(2), 249–264.
- Bateman I.J., Day B.H., Georgiou S. & Lake I., 2006. The aggregation of environmental benefit values: Welfare measures, distance decay and total WTP. *Ecological Economics* **60**(2), 450–460.
- Baumgärtner S. & Quaas M., 2010. What is sustainability economics? *Ecological Economics* **69**(3), 445–450.
- Bebbington J., Brown J. & Frame B., 2007. Accounting technologies and sustainability assessment models. *Ecological Economics* **61**(2–3), 224–236.
- Bell S., Tyrväinen L., Sievänen T., Pröbstl U. & Simpson M., 2007a. Outdoor Recreation and Nature Tourism: A European Perspective. *Living Rev. Landscape Res.* **1**.
- Bendixen M., 2003. A Practical Guide to the Use of Correspondence Analysis in Marketing Research. *Marketing Bulletin* 15.

Références bibliographiques

- Bowler D.E., Buyung-Ali L.M., Knight T.M. & Pullin A.S., 2010. A systematic review of evidence for the added benefits to health of exposure to natural environments. *BMC Public Health* **10**, 456.
- Breyne J., 2017. *Tourisme, biodiversité et bien-être*.
- Breyne J., Vanbelinghen N. & Abildtrup J., 2018. *L'écotourisme en Ardenne, Les actions et les attentes des opérateurs touristiques par rapport aux espaces naturels, Interreg V - AGRETA*.
- Bro R., Kjeldahl K., Smilde A.K. & Kiers H.A.L., 2008. Cross-validation of component models: A critical look at current methods. *Anal Bioanal Chem* **390**(5), 1241–1251.
- Brockington D., 2011. Ecosystem services and fictitious commodities. *Envir. Conserv.* **38**(4), 367–369.
- Buckley R., 2009. *Ecotourism: Principles and Practices.*, CABI, Wallingford, UK.
- Budowski G., 1976. Tourism and Environmental Conservation: Conflict, Coexistence, or Symbiosis? *Envir. Conserv.* **3**(1), 27–31.
- Campbell M., 2010. Seeing is believing: Using digital cameras to monitor trail use in Riding Mountain national park. In: *Ecosystem Based Management: Beyond Boundaries*. Presented at the Proceedings of the Sixth International Conference of Science and the Management of Protected Areas Wolfville, Nova Scotia, 313–319.
- Campbell M.J., 2006. Monitoring trail use with digital still cameras: Strengths, limitations and proposed resolutions. In: Siegrist, D., Clivaz, C., Hunziker, M., Iten, S. eds. . Presented at the Exploring the nature of management. Proceedings of third international conference on monitoring and management of visitor flows in recreational and protected areas. Rapperswil, Switzerland: University of Applied Sciences., 13–17.
- Carnol M. & Verheyen K., 2010. Les services écosystémiques dans les forêts mélangées et pures : perception des utilisateurs et connaissances scientifiques. *Forêt Wallonne* (106), 49–59.
- Ceballos-Lascuráin H., 1996. *Tourism, Ecotourism, and Protected Areas: The State of Nature-based Tourism Around the World and Guidelines for Its Development*, IUCN, 301.
- Cessford G. & Muhar A., 2003. Monitoring options for visitor numbers in national parks and natural areas. *Journal for Nature Conservation* **11**(4), 240–250.
- Chen G., Han T.X., He Z., Kays R. & Forrester T., 2014. Deep convolutional neural network based species recognition for wild animal monitoring. In: *2014 IEEE International Conference on Image Processing (ICIP)*. Presented at the 2014 IEEE International Conference on Image Processing (ICIP) IEEE, Paris, France, 858–862.
- Cherem G.J. & Driver B.L., 1983. Visitor employed photography: a technique to measure common perceptions of natural environments. *Journal of Leisure Research* **15**, 65–83.
- Civitello D.J., Cohen J., Fatima H., Halstead N.T., Liriano J., McMahon T.A., Ortega C.N., Sauer E.L., Sehgal T., Young S. & Rohr J.R., 2015. Biodiversity inhibits parasites: Broad evidence for the dilution effect. *PNAS* 201506279.
- Cobb J.B. & Griffin D.R., 1976. *Process Theology: An Introductory Exposition*, Westminster John Knox Press, 200.
- Colson V., 2009. La fonction récréative des massifs forestiers wallons : analyse et évaluation dans le cadre d'une politique forestière intégrée.

Références bibliographiques

- Commissariat Général au Tourisme - SPW (CGT), 2017. La Wallonie touristique en chiffres - Données 2017, Jambes.
- Conlon K.E., 2014. Investigating the Relationship between Trail Conditions and Visitor Behavior Using the Camera Trap Method (Graduate thesis).
- Cormen T.H., Leiserso C.E. & Riverst L., 1990. *Introduction to Algorithms*, MIT Press, Cambridge, MA, USA.
- Costanza R., d'Arge R., Groot R. de, Farber S., Grasso M., Hannon B., Limburg K., Naeem S., O'Neill R.V., Paruelo J., Raskin R.G., Sutton P. & Belt M. van den, 1997. The value of the world's ecosystem services and natural capital. *Nature* **387**(6630), 253.
- Cronin S., 2010. Camera trap talk. *Photographic Society*.
- Cutler T.L. & Swann D.E., 1999. Using Remote Photography in Wildlife Ecology: A Review. *Wildlife Society Bulletin (1973-2006)* **27**(3), 571–581.
- Daniel T.C., Muhar A., Arnberger A., Aznar O., Boyd J.W., Chan K.M.A., Costanza R., Elmquist T., Flint C.G., Gobster P.H., Gret-Regamey A., Lave R., Muhar S., Penker M., Ribe R.G., Schauppenlehner T., Sikor T., Soloviy I., Spierenburg M., Taczanowska K., Tam J. & von der Dunk A., 2012. Contributions of cultural services to the ecosystem services agenda. *Proceedings of the National Academy of Sciences* **109**(23), 8812–8819.
- de Groot R.S., Wilson M.A. & Boumans R.M.J., 2002. A typology for the classification, description and valuation of ecosystem functions, goods and services. *Ecological Economics* **41**(3), 393–408.
- Debusschere M., Lusyne P., Dewitte P., Baeyens Y., Meersman F.D., Seynaeve G., Demunter C., Reis F. & Reuter H.I., 2017. Big data et statistiques : un recensement tous les quarts d'heure... 22.
- Del Casino V.J. & Hanna S.P., 2000. Representations and identities in tourism map spaces. *Progress in Human Geography* **24**(1), 23–46.
- Dodge S. & Karam L., 2016. Understanding how image quality affects deep neural networks. In: *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*. Presented at the 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX), 1–6.
- Dumont B., Roovers P. & Gulinck H., 2005. Estimation of off-track visits in a nature reserve: a case study in central Belgium. *Landscape and Urban Planning* **71**(2), 311–321.
- Eagles P.F.J., 2002. Trends in Park Tourism: Economics, Finance and Management. *Journal of Sustainable Tourism* **10**(2), 132–153.
- Eagles P.F.J., 2014. Research priorities in park tourism. *Journal of Sustainable Tourism* **22**(4), 528–549.
- Esu B.B. & Ebitu E., 2010. Promoting and Emerging Tourism Destination. *Global Journal of Management and Business Research* **10**(1), 21–28.
- European Commission, 2014. Feasibility Study on the Use of Mobile Positioning Data for Tourism Statistics, Eurostat.
- European Environment Agency (EEA), 1999. Environmental indicators: Typology and overview (Technical report No. 25), European Environment Agency, Copenhagen.

Références bibliographiques

- Fairfax R.J., Dowling R.M. & Neldner V.J., 2014. The use of infrared sensors and digital cameras for documenting visitor use patterns: a case study from D'Aguilar National Park, south-east Queensland, Australia. *Current Issues in Tourism* **17**(1), 72–83.
- Farber S.C., Costanza R. & Wilson M.A., 2002. Economic and ecological concepts for valuing ecosystem services. *Ecological Economics* **41**(3), 375–392.
- Forest Research, 2010. Benefits of green infrastructure, Forest Research, Farnham.
- Froment A., 1968. L'ancienne économie rurale de l'Ardenne et son incidence sur la végétation des Hautes Fagnes. *BSGLg* **1**(4), 23–39.
- Gasparatos A. & Scolobig A., 2012. Choosing the most appropriate sustainability assessment tool. *Ecological Economics* **80**, 1–7.
- GEIE Destination Ardenne, 2017. *Conférence de Presse : Projet AGRETA*.
- GEIE Destination Ardenne, 2018. *Destination Ardenne: Les chiffres clé*.
- Girshick R., 2015. Fast R-CNN. Presented at the IEEE International Conference on Computer Vision (ICCV).
- Girshick R., Donahue J., Darrell T. & Malik J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. Presented at the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- Gomez Villa A., Salazar A. & Vargas F., 2017. Towards automatic wild animal monitoring: Identification of animal species in camera-trap images using very deep convolutional neural networks. *Ecological Informatics* **41**, 24–32.
- gouvernement wallon, 2004. Déclaration de politique régionale 2004-2009, Jambes.
- Gretzel U., 2011. Intelligent systems in tourism. *Annals of Tourism Research* **38**(3), 757–779.
- Griffiths M. & Van Schaik C.P., 1993. Camera-trapping: a new tool for the study of elusive rain forest animals. *Tropical Biodiversity* **1**, 131–135.
- Guillén J., García-Olivares A., Ojeda E., Osorio A., Chic O. & González R., 2008. Long-Term Quantification of Beach Users Using Video Monitoring. *Journal of Coastal Research* **246**, 1612–1619.
- Gursoy D., McCleary K.W. & Lepsito L.R., 2007. Propensity to Complain: Affects of Personality and Behavioral Factors. *Journal of Hospitality & Tourism Research* **31**(3), 358–286.
- Haines-Young R. & Potschin M.B., 2018. Common International Classification of Ecosystem Services (CICES) V5.1 and Guidance on the Application of the Revised Structure 53.
- Hammond A.L. & World Resources Institute (Eds.), 1995. *Environmental indicators: a systematic approach to measuring and reporting on environmental policy performance in the context of sustainable development*, World Resources Institute, Washington, D.C., 43.
- Hartig T., Evans G.W., Jamner L.D., Davis D.S. & Gärling T., 2003. Tracking restoration in natural and urban field settings. *Journal of Environmental Psychology* **23**(2), 109–123.
- Hartmann R., 1988. Combining field methods in tourism research. *Annals of Tourism Research* **15**(1), 88–105.
- He K., Gkioxari G., Dollár P. & Girshick R., 2017. Mask R-CNN. Presented at the The IEEE International Conference on Computer Vision (ICCV), 2961–2969.

Références bibliographiques

- Heyman E., 2012. Analysing recreational values and management effects in an urban forest with the visitor-employed photography method. *Urban Forestry & Urban Greening* **11**(3), 267–277.
- Hoebeke L., 2018. Automated recognition of people and identification of animal species in camera trap images (Master's degree thesis).
- Hornback K.E. & Eagles P.F.J., 1999. *Guidelines for public use measurement and reporting at parks and protected areas*, IUCN, Gland, Switzerland, 86.
- Hotelling H., 1933. Analysis of a complex of statistical variables into principal components. *J Educ Psychol* **24**, 417–441, 498–520.
- Inventaire Permanent des Ressources Forestières de Wallonie, 2008. La forêt Wallonne en quelques chiffres au Cycle 1. <http://iprfw.spw.wallonie.be/cy1-chiffres.php>, (13/05/2019).
- IPBES, 2019. Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science - Policy Platform on Biodiversity and Ecosystem Services.
- Janowsky D.V. & Becker G., 2003. Characteristics and needs of different user groups in the urban forest of Stuttgart. *Journal for Nature Conservation* **11**(4), 251–259.
- Jones A., Wright J., Bateman I.J. & Schaafsma M., 2010. Estimating Arrival Numbers for Informal Recreation: A Geographical Approach and Case Study of British Woodlands.
- Josse J. & Husson F., 2012. Selecting the number of components in principal component analysis using cross-validation approximations. *Computational Statistics & Data Analysis* **56**(6), 1869–1879.
- Josse J. & Husson F., 2016. missMDA: A Package for Handling Missing Values in Multivariate Data Analysis. *Journal of Statistical Software* **70**(1), 1–31.
- Kammler M. & Schernewski G., 2004. Spatial and temporal analysis of beach tourism using webcam and aerial photographs. In: *Managing the Baltic Sea*. Presented at the Coastline Reports 2G. Schernewski & N. Löser, Schwerin, Germany, 121–128.
- Kareiva P., Watts S., McDonald R. & Boucher T., 2007. Domesticated Nature: Shaping Landscapes and Ecosystems for Human Welfare. *Science* **316**(5833), 1866–1869.
- Karmanov D. & Hamel R., 2008. Assessing the restorative potential of contemporary urban environment(s): Beyond the nature versus urban dichotomy. *Landscape and Urban Planning* **86**(2), 115–125.
- Kays R., Kranstauber B., Jansen P., Carbone C., Rowcliffe M., Fountain T. & Tilak S., 2009. Camera traps as sensor networks for monitoring animal communities. In: *2009 IEEE 34th Conference on Local Computer Networks*. Presented at the 2009 IEEE 34th Conference on Local Computer Networks (LCN 2009)IEEE, Zurich, Switzerland, 811–818.
- Kellner L. & Egger R., 2016. Tracking Tourist Spatial-Temporal Behavior in Urban Places, A Methodological Overview and GPS Case Study. In: Inversini, A., Schegg, R. eds. *Information and Communication Technologies in Tourism 2016*. Springer International Publishing, Cham, 481–494.
- Keul A. & Kühberger A., 1997. Tracking the Salzburg tourist. *Annals of Tourism Research* (24), 1008–1012.

Références bibliographiques

- Krizhevsky A., Sutskever I. & Hinton G.E., 2012. ImageNet classification with deep convolutional neural networks. *In: Communications of the ACM*. Presented at the Proceedings of the 25th International Conference on Neural Information Processing Systems Lake Tahoe, NV, 1097–1105.
- Kumar M. & Kumar P., 2008. Valuation of the ecosystem services: A psycho-cultural perspective. *Ecological Economics* **64**(4), 808–819.
- LeCun Y., Bengio Y. & Hinton G., 2015. Deep learning. *Nature* **521**(7553), 436–444.
- LeCun Y. & et al., 1990. Handwritten digit recognition with a back-propagation network. Presented at the Proc. Advances in Neural Information Processing Systems, 396–404.
- Li F.F. & Karpathy A., 2015. *CS231n Stanford lecture notes on CNN*.
- Li W., 2004. Environmental management indicators for ecotourism in China's nature reserves: A case study in Tianmushan Nature Reserve. *Tourism Management* **6**.
- Lin T.-Y., Maire M., Belongie S., Bourdev L., Girshick R., Hays J., Perona P., Ramanan D., Zitnick C.L. & Dollár P., 2014. Microsoft COCO: Common Objects in Context. Presented at the European Conference of Computer Vision (ECCV).
- Litzenberger D.C., n.d. *PyCrypto—The Python Cryptography Toolkit*.
- Liu C.-L., 2008. Partial discriminative training for classification of overlapping classes in document analysis. *IJDAR* **11**(2), 53.
- Long J., Shelhamer E. & Darrell T., 2015. Fully Convolutional Networks for Semantic Segmentation. *CVPR* **10**.
- Lupp G., Förster B., Kantelberg V., Markmann T., Naumann J., Honert C., Koch M. & Pauleit S., 2016. Assessing the Recreation Value of Urban Woodland Using the Ecosystem Service Approach in Two Forests in the Munich Metropolitan Region. *Sustainability* **8**(11), 1156.
- Lynch J., Vogt C., Sindrity S. & Nelson C., 2002. Measuring and Monitoring Trail Use: A Nationwide Survey of State and Federal Trail Managers, Department of Park, Recreation and Tourism Resources Michigan State University, Michigan, USA.
- Maillefert M. & Petit O., 2017. Vers une démarche intégrée d'évaluation et de représentation des services écosystémiques : perspective interdisciplinaire et enjeux en milieu urbain. *Environnement Urbain / Urban Environment* (Volume 11).
- Mann H.B. & Whitney D.R., 1947. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* **18**(1), 50–60.
- Marcelino P., 2018. Transfer learning from pre-trained models. *Towards Data Science*. <https://towardsdatascience.com/transfer-learning-from-pre-trained-models-f2393f124751>, (29/03/2019).
- Mayer M., Müller M., Woltering M., Arnegger J. & Job H., 2010. Mayer M, Mueller M, Woltering M, Arnegger J, Job H. The economic impact of tourism in six German national parks. *Landscape and Urban Planning* **97**, 73–82.
- Meek P. & Fleming P. (Eds.), 2014. *Camera Trapping*, CSIRO Publishing.
- Meek P.D., Ballard G.-A. & Fleming P.J.S., 2013. A permanent security post for camera trapping. *Australian Mammalogy* **35**(1), 123.

Références bibliographiques

- Melville S. & Ruohonen J., 2004. The development of a remote-download system for visitor counting. Presented at the Proceedings of the Second Conference on Monitoring and Management of Visitor Flows in Recreational and Protected Areas. Sievañnen, J. Erkkonen, J. Jokimañki, J. Saarinen, S. Tuulentie, and E. Virtanen, Rovaniemi, Finland, 37–43.
- Mikolov T., Kombrink S., Burget L., Cernocky J. & Khudanpur S., 2011. Extensions of recurrent neural network language model. *In: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Presented at the ICASSP 2011 - 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)IEEE, Prague, Czech Republic, 5528–5531.
- Millennium Ecosystem Assessment (Program) (Ed.), 2005. *Ecosystems and human well-being: synthesis*, Island Press, Washington, DC, 137.
- Millhäusler A., Anderwald P., Haeni M. & Haller R.M., 2016. Publicity, economics and weather – Changes in visitor numbers to a European National Park over 8 years. *Journal of Outdoor Recreation and Tourism* **16**, 50–57.
- Minsky M. & Papert S.A., 1969. *Perceptrons: An Introduction to Computational Geometry*, The MIT Press, Massachussets, USA, 9.
- Mirjalili V., 2017. How does the region proposal network (RPN) in Faster R-CNN work? *Quora*.
- Mooney H., Larigauderie A., Cesario M., Elmquist T., Hoegh-Guldberg O., Lavorel S., Mace G.M., Palmer M., Scholes R. & Yahara T., 2009. Biodiversity, climate change, and ecosystem services. *Current Opinion in Environmental Sustainability* **1**(1), 46–54.
- Moore S., 2012. Current and future issues in natural area tourism with a special focus on visitor monitoring. *In: Fredman, P., Stenseke, M., Liljendahl, H., Mossing, A., Laven, D. eds. Outdoor Recreation in Change – Current Knowledge and Future Challenges*. Presented at the The 6th International Conference on Monitoring and Management of Visitors in Recreational and Protected Areas Stockholm, Sweden, 20–21.
- Muhar A., Arnberger A. & Brandenburg C., 2002. Methods for Visitor Monitoring in Recreational and Protected Areas: An Overview. Presented at the Monitoring and Management of Visitor Flows in Recreational and Protected Areas, 6.
- Murphy P.E., 1992. Urban Tourism and Visitor Behavior. *American Behavioral Scientist* **36**(2), 200–211.
- Neal J.D. & Gursoy D., 2008. A Multifaceted Analysis of Tourism Satisfaction. *Journal of Travel Research* **47**(1), 53–62.
- Norton L.R., Inwood H., Crowe A. & Baker A., 2012. Trialling a method to quantify the ‘cultural services’ of the English landscape using Countryside Survey data. *Land Use Policy* **7**.
- O’Connor S., Campbell R., Cortez H. & Knowles T., 2009. Whale Watching Worldwide: Tourism Numbers, Expenditures and Expanding Economic Benefits, a Special Report from the International Fund for Animal Welfare Prepared by Economists at Large, International Fund for Animal Welfare, Yarmouth, MA.
- OECD, 1993. OECD core set of indicators for environmental performance reviews, Organization for Economic Cooperation and Development, Paris, France.

Références bibliographiques

- Olsson M.P.O., Widén P. & Larkin J.L., 2008. Effectiveness of a highway overpass to promote landscape connectivity and movement of moose and roe deer in Sweden. *Landscape and Urban Planning* **85**(2), 133–139.
- Pearson K., 1901. On lines and planes of closest fit to systems of points in space. *Philos Mag* **2**, 559–572.
- Peterson L.M. & Thomas J.A., 1998. Performance of TrailMaster infrared sensors in monitoring captive coyotes. *Wildlife Society bulletin*.
- Pettebone D., Newman P. & Lawson S.R., 2010. Estimating visitor use at attraction sites and trailheads in Yosemite National Park using automated visitor counters. *Landscape and Urban Planning* **97**(4), 229–238.
- Pintér L., Hardi P., Martinuzzi A. & Hall J., 2012. Bellagio STAMP: Principles for sustainability assessment and measurement. *Ecological Indicators* **17**, 20–28.
- Plunus J. & Heinen D., 2012. Les impacts socio-économiques du projet LIFE Hautes-Fagnes : étude préliminaire et exposition du cas concret pour les entreprises locales. *Les amis de la Fagne* **288**(1), 13–15.
- Portail Biodiversité Wallonie (SPW), 2014. Le méta-projet de restauration des tourbières de Haute-Ardenne. *biodiversite.wallonie.be*. <http://biodiversite.wallonie.be/fr/meta-projet-life-de-restauration-des-tourbieres-de-haute-ardenne.html?IDC=5778>, (13/05/2019).
- Python Software Foundation, n.d. *Python Language Reference, version 3.7*.
- R Development Core Team, 2005. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ren S., He K., Girshick R. & Sun J., 2016. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. Presented at the IEEE Transaction on Pattern Analysis and Machine Learning.
- Ribeiro M.T., Singh S. & Guestrin C., 2016. “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*. Presented at the the 22nd ACM SIGKDD International Conference ACM Press, San Francisco, California, USA, 1135–1144.
- Rice C.G., Kucera T.E. & Barrett R.H., 1995. Trailmaster® Camera System. *Wildlife Society Bulletin* **23**(1), 110–113.
- Richards D.R. & Tunçer B., 2018. Using image recognition to automate assessment of cultural ecosystem services from social media photographs. *Ecosystem Services* **31**, 318–325.
- Roberts N.J., 2011. Investigation into survey techniques of large mammals: surveyor competence and camera-trapping vs. transect-sampling. *Bioscience Horizons* **4**(1), 40–49.
- Rodríguez J., Beard J., Bennett E., Cumming G., Cork S., Agard J., Dobson A. & Peterson G., 2006. Trade-offs across Space, Time, and Ecosystem Services. *Ecology and Society* **11**(1).
- Rohrer B., 2016. How do Convolutional Neural Networks work? *End-to-End Machine Learning*.

Références bibliographiques

- Rosenblatt F., 1958. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* **65**(6), 386–408.
- Ross J., 2005. *Visitor counters in parks: management practice for counter calibration*, Wellington.
- Rossi S.D., Byrne J.A. & Pickering C.M., 2015. The role of distance in peri-urban national park use: Who visits them and how far do they travel? *Applied Geography* **63**, 77–88.
- Rowcliffe J.M., Carbone C., Jansen P.A., Kays R. & Kranstauber B., 2011. Quantifying the sensitivity of camera traps: an adapted distance sampling approach. *Methods in Ecology and Evolution* **2**(5), 464–476.
- Rundle S., 2002. Monitoring Low Volume Walker Use of a Remote Mountain Range: a Case Study of the Arthur Range, Tasmania, Australia. Presented at the Monitoring and Management of Visitor Flows in Recreational and Protected AreasA. Arnberger, C. Brandenburg, A. Muhar, 53–58.
- Rupf R., Wernli M. & Filli F., 2006. Visitor counting with acoustic slab sensors in the Swiss National Park. In: Siegrist, D., Clivaz, C., Hunziker, M., Iten, S. eds. . Presented at the Exploring the nature of management. Proceedings of the third international conference on monitoring and management of visitor flows in recreational and protected areasUniversity of Applied Sciences, Rapperswil. Rapperswil, Switzerland.
- Russakovsky O., Deng J., Su H., Krause J., Satheesh S., Ma S., Huang Z., Karpathy A., Khosla A., Bernstein M., Berg A.C. & Fei-Fei L., 2015. ImageNet Large Scale Visual Recognition Challenge. *IJCV* **115**(3), 211–252.
- Schneiders A., Van Daele T., Van Landuyt W. & Van Reeth W., 2012. Biodiversity and ecosystem services: Complementary approaches for ecosystem management? *Ecological Indicators* **21**, 123–133.
- Séquin E.S., Jaeger M.M., Brussard P.F. & Barrett R.H., 2003. Wariness of coyotes to camera traps relative to social status and territory boundaries. *Can. J. Zool.* **81**(12), 2015–2025.
- Shane J., 2018. Do neural nets dream of electric sheep? <https://aiweirdness.com/post/171451900302/do-neural-nets-dream-of-electric-sheep>, (07/08/2019).
- Shiras G., 1935. Hunting Wild Life with Camera and Flashlight: A Record of Sixty-five Years' Visits to the Woods and Waters of North America. *National geographic society*.
- Shoval N. & Isaacson M., 2007. Tracking tourists in the digital age. *Annals of Tourism Research* **34**(1), 141–159.
- Stabler M.J., 1997. *Tourism Sustainability: Principles to Practice*, New York.
- Stynes D.J., 2005. Economic significance of recreational uses of national parks and other public lands. *Social Science Research Review* **5**(1), 36.
- Swann D.E., Hass C.C., Dalton D.C. & Wolf S.A., 2004. Infrared-triggered cameras for detecting wildlife: an evaluation and review. *Wildlife Society Bulletin* **32**(2), 357–365.
- Swann D.E., Kawanishi K. & Palmer J., 2010. Evaluating Types and Features of Camera Traps in Ecological Studies: A Guide for Researchers. In: O'Connell, A.F., Nichols, J.D., Karanth, U.K. eds. *Camera Traps in Animal Ecology: Methods and Analyses*. Springer, Tokyo, Dordrecht, London, Heidelberg, New York, 27–43.

Références bibliographiques

- Swanson A., Kosmala M., Lintott C., Simpson R., Smith A. & Packer C., 2015. Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. *Scientific Data* **2**, 150026.
- TEEB, 2010. The Economics of Ecosystem Services and Biodiversity, Earthscan, London.
- Thakur J. & Kumar N., 2011. DES, AES and Blowfish: Symmetric Key Cryptography Algorithms Simulation Based Performance Analysis **1**(1), 7.
- Towerton A.L., Penman T.D., Blake M.E., Deane A.T., Kavanagh R.P. & Dickman C.R., 2008. The potential for remote cameras to monitor visitation by birds and predators at Malleefowl mounds. *Ecological Management & Restoration* **9**(1), 64–67.
- Trollet F., Huynen M.-C., Vermeulen C. & Hambuckers A., 2014. Use of camera traps for wildlife studies. A review. *Biotechnol. Agron. Soc. Environ.* **9**.
- Tsang S.-H., 2018. Review: FCN — Fully Convolutional Network (Semantic Segmentation). *Towards Data Science*.
- Uijlings J.R.R., van de Sande K.E.A., Gevers T. & Smeulders A.W.M., 2013. Selective Search for Object Recognition. *Int J Comput Vis* **104**(2), 154–171.
- Vallet J., 2009. Gradient d'urbanisation et communautés végétales d'espaces boisés
/>Approche à plusieurs échelles dans trois agglomérations du Massif armoricain.
- van Loon-Steensma J.M. & Vellinga P., 2013. Trade-offs between biodiversity and flood protection services of coastal salt marshes. *Current Opinion in Environmental Sustainability* **5**(3–4), 320–326.
- Ventura A., 2015. Écotourisme : une brève introduction. *ELOHI* (7), 3–15.
- Vine S.J., Crowther M.S., Lapidge S.J., Dickman C.R., Mooney N., Piggott M.P. & English A.W., 2009. Comparison of methods to detect rare and cryptic species: a case study using the red fox (*Vulpes vulpes*). *Wildl. Res.* **36**(5), 436.
- Vo Q.T., Kuenzer C., Vo Q.M., Moder F. & Oppelt N., 2012. Review of valuation methods for mangrove ecosystem services. *Ecological Indicators* **23**, 431–446.
- Voulodimos A., Doulamis N., Doulamis A. & Protopapadakis E., 2018. Deep Learning for Computer Vision: A Brief Review. *Computational Intelligence and Neuroscience*.
- Wal-ES, 2016. Les services écosystémiques en Wallonie.
- Watson A.E., Cole D.N., Turner D.L. & Reynolds P.S., 2000. Wilderness recreation use estimation: a handbook of methods and systems (No. RMRS-GTR-56), U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Ft. Collins, CO.
- Wearn O.R. & Glover-Kapfer P., 2017. Camera-trapping for conservation: a guide to best-practices. *WWF Conservation Technology Series* **1**(1).
- Weathers W.W. & Seymour R.S., 1998. Behaviour and Time-Activity Budgets of Malleefowl *Leipoa ocellata* in South Australia. *Emu - Austral Ornithology* **98**(4), 288–296.
- Weaver D.B., 2001. Ecotourism as Mass Tourism: Contradiction or Reality?, Cornell University.
- Wegge P., Pokheral C.P. & Jnawali S.R., 2004. Effects of trapping effort and trap shyness on estimates of tiger abundance from camera trap studies. *Animal Conservation forum* **7**(3), 251–256.

Références bibliographiques

- Weinstein B.G., 2015. MotionMeerkat: integrating motion video detection and ecological monitoring. *Methods in Ecology and Evolution* **6**(3), 357–362.
- Werbos J., 1974. Beyond regression: New tools for prediction and analysis in the behavioral sciences.
- Whitelaw P.A., King B.E.M. & Tolkach D., 2014. Protected areas, conservation and tourism – financing the sustainable dream. *Journal of Sustainable Tourism* **22**(4), 584–603.
- Wilcoxon F., 1945. Individual comparisons by ranking methods. *Biometrics Bulletin* **1**(6), 80–83.
- Wimpey J.F. & Marion J.L., 2010. The influence of use, environmental and managerial factors on the width of recreational trails. *Journal of Environmental Management* **91**(10), 2028–2037.
- Wolf I.D., Hagenloh G. & Croft D.B., 2012. Visitor monitoring along roads and hiking trails: How to determine usage levels in tourist sites. *Tourism Management* **33**(1), 16–28.
- World Tourism Organization (UNWTO) (Ed.), 2018. *Faits saillants OMT du tourisme, édition 2018*, World Tourism Organization (UNWTO).
- Xiao T., Li S., Wang B., Lin L. & Wang X., 2017. Joint Detection and Identification Feature Learning for Person Search. Presented at the Conference on Computer Vision and Pattern...IEEE Conference on Computer Vision and Pattern Recognition (CVPR)Honolulu, HI, USA, 3415–3424.
- Yosinski J., Clune J., Bengio Y. & Lipson H., 2014. How transferable are features in deep neural networks? *Advances in Neural Information Processing Systems* **27** 3320–3328.
- Yousif H., Yuan J., Kays R. & He Z., 2019. Animal Scanner: Software for classifying humans, animals, and empty frames in camera trap images. *Ecology and Evolution* **9**(4), 1578–1589.
- Yu X., Wang J., Kays R., Jansen P.A., Wang T. & Huang T., 2013. Automated identification of animal species in camera trap images. *EURASIP Journal on Image and Video Processing* **2013**(1).

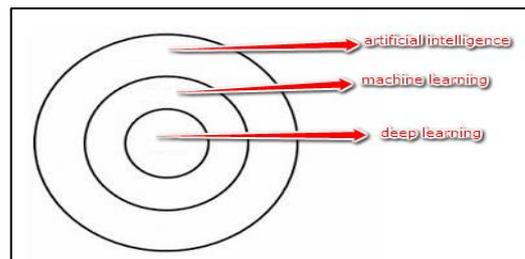
Annexes

Annexe 1 : Histoire et fonctionnement des réseaux neuronaux convolutionnels

1) Intelligence artificielle et Automatisation

Jusqu'à récemment, la démocratisation des techniques d'analyse d'images par ordinateur était principalement freinée par la quantité de données nécessitant une puissance informatique qui n'était pas disponible pour le grand public (Debusschere et al., 2017). L'essor des technologies liées au Big Data de cette dernière décennie a permis de s'y réintéresser de manière plus efficace. Aujourd'hui il est désormais possible d'automatiser la détection et la classification de sujets sur des images prises par camera trap. Cette délivrance permet de se libérer de la contrainte liée au temps que cela nécessite lorsque les images sont analysées manuellement comme dans la plupart des études actuelles (Conlon, 2014; Bambi et al., 2017). Cependant, faute d'avoir trouvé la technique adaptée à l'ensemble des problèmes de ce type, de nombreux scientifiques se relaient dans l'amélioration de cette méthode (Yu et al., 2013).

Le terme « intelligence artificielle » est vaste. Selon Alexandre (2017), il va des algorithmes programmés pour réaliser des tâches précises qui existent depuis l'invention de l'ordinateur à des concepts futuristes tels qu'une conscientisation de la machine. Depuis 2012, le deep learning se développe. L'ordinateur dépasse l'Homme dans plusieurs tâches en créant sa propre vision du monde grâce à un réseau de neurones artificiels.



Selon LeCun et al. (2015), les termes « deep learning » et « machine learning » sont souvent confondus. Le premier est en fait un sous ensemble du second. Le machine learning, désigne le fait que l'ordinateur puisse faire des prédictions en fonction des données insérées dans l'algorithme. Dans les algorithmes classiques, on insère des paramètres pour obtenir une réponse en fonction des règles encodées. Dans le machine learning, on insère les paramètres et la réponse, l'algorithme en déduit alors les règles qu'il pourrait réappliquer par la suite. Le deep learning ne gère plus une base de données avec des attributs bien spécifiques en entrée du modèle, mais des données qui vont être insérées dans des réseaux de neurones artificiels, inspirés des réseaux de neurones biologiques. L'algorithme se sert de ces données brutes pour créer lui-même de nouveaux attributs qui lui conviennent mieux pour s'entraîner et créer des résultats probants. Il s'agit de couches neuronales intermédiaires qui seront brièvement décrites au prochain point.

a. Réseaux neuronaux artificiels

Ce système nerveux théorique porte le nom de « perceptron », du mot « percevoir », il s'agit d'un réseau comportant une seule couche de neurones (Rosenblatt, 1958). Chaque neurone correspond à une valeur d'entrée. Ensuite amélioré par Werbos (1974) en partie grâce aux critiques de Minsky et Papert (1969), entre en scène le perceptron multicouche, ajoutant des

couches de neurones supplémentaires et cachées dont le rôle est de transformer l'information en entrée en attributs plus compréhensibles par l'ordinateur (Figure 19). Plus il y a de couches cachées, plus le modèle est complexe et « profond » (du terme « deep » learning). Chaque neurone d'une couche est relié à l'ensemble des neurones de la couche précédente. La couche en sortie contient les classes du modèle. Les données en entrée doivent donc contribuer à classer l'ensemble de l'information dans une catégorie.

L'autre intérêt du travail de Werbos (1974) est donné par le concept d'algorithme de rétropropagation de l'erreur pour améliorer l'entraînement du modèle. Les neurones ayant le plus contribué à l'erreur du modèle verront leur poids grandement diminué à la prochaine itération alors que ceux qui fournissent des résultats plus justes seront plus importants (Figure 19).

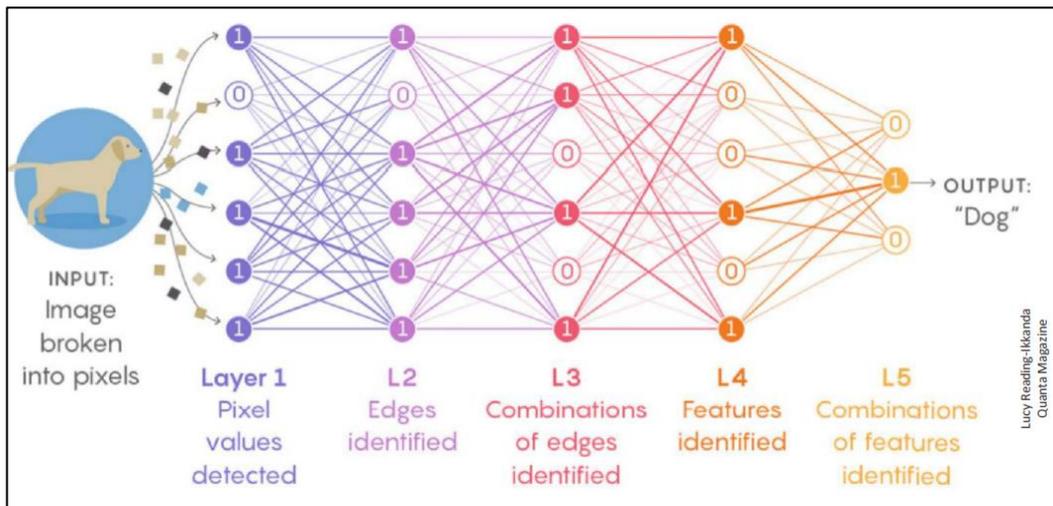


Figure 19: Exemple de perceptron multicouche (source: web)

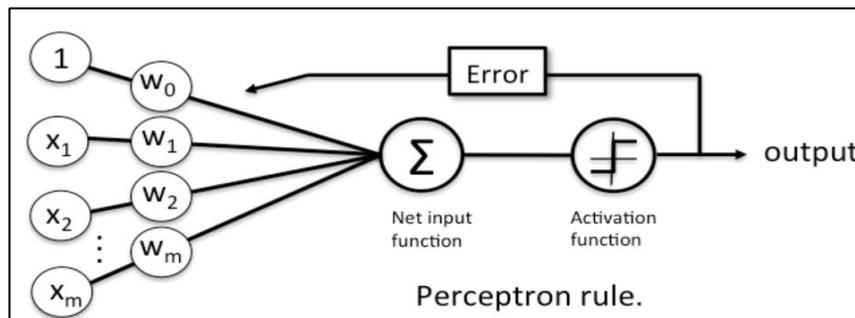


Figure 19: Principe de fonctionnement du perceptron et de la rétropropagation de l'erreur (source :web)

Grâce à la rétropropagation, les poids des relations sont donc modifiés à chaque itération dans le but de minimiser l'erreur du modèle en sortie. Cette modification des poids ne se fait pas au hasard, mais généralement selon le principe de descente du gradient, qui ne sera pas abordé ici.

b. Réseaux neuronaux convolutionnels (CNN)

Les scientifiques se sont finalement rendu compte que pour étudier des données sous forme d'images, de vidéos, de sons ou de textes, les résultats étaient trop dépendants de l'architecture du perceptron multicouche (LeCun et al., 2015). Les réseaux neuronaux convolutionnels ont alors été créés. Ils permettent de faciliter l'entraînement du modèle et d'obtenir des résultats plus généralisables principalement grâce au principe de convolution

Annexes

(Voulodimos et al., 2018). Ils sont majoritairement utilisées depuis que le modèle AlexNet (Krizhevsky et al., 2012) a gagné le « *ImageNet Large Scale Visual Recognition Competition* » challenge en 2012 (Russakovsky et al., 2015).

Contrairement à l'œil humain pour lequel reconnaître et classifier des images est facile, l'ordinateur voit seulement une matrice de valeurs numériques. Pour expliquer le fonctionnement des réseaux neuronaux convolutionnels, un cas simple est tiré de (Rohrer, 2016). Il s'agit d'une classification d'image entre deux catégories : la lettre « X » et la lettre

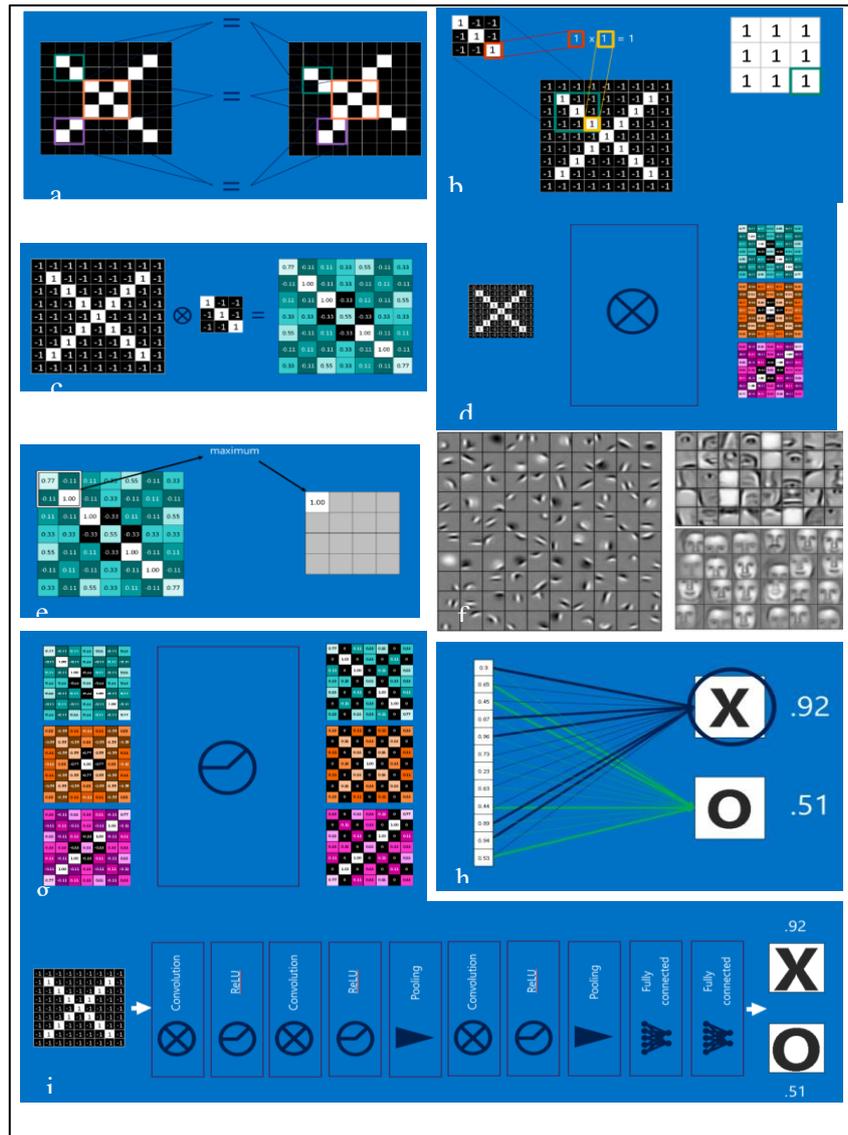


Figure 20: Schémas représentant les différentes étapes constituant un réseau neuronal convolutionnel (Rohrer, 2016). De haut en bas et de gauche à droite : (a) concordance entre parties d'image, (b) passage d'un filtre sur l'image, (c) carte de correspondance entre le filtre et l'image, (d) répétition de l'opération pour les différents filtres, (e) max pooling, (f) exemples de patterns à détecter avec une complexité croissante, (g) fonction ReLU, (h) fully

Annexes

« O ». L'image mesure 9x9 pixels et est en noir et blanc. Les données sont considérées par l'ordinateur comme une matrice 9x9 avec des valeurs binaires¹³.

Dans le cas des CNN et à la différence des réseaux neuronaux classiques, les neurones d'une couche ne sont pas connectés à l'ensemble des neurones de la couche précédente. Le CNN part du principe qu'il est plus simple de trouver des correspondances entre des fragments d'image qu'entre des images entières. Sur la figure 6(a), trois parties de l'image sont donc mises en évidence. Chaque pixel représente un neurone.

Couche de convolution

Lors de l'entraînement du réseau, l'opérateur lui a donné une multitude d'exemples de « X » et de « O », en donnant au modèle la réponse qu'il devait atteindre. Au fur et à mesure, le modèle a retenu les parties d'image caractéristiques de ces lettres, celles qui étaient le plus redondantes et discriminantes. Lors de l'utilisation du modèle, la phase de convolution consiste à faire voyager ces parties caractéristiques sur l'image à analyser. Sur la figure 6(b), le fragment retenu par le modèle lors de l'entraînement (appelé « filtre ») fait 3 pixels de côté. Lorsqu'il voyage sur l'image, le produit de la superposition de chaque pixel deux à deux est calculé (blanc sur blanc : $1 \times 1 = 1$; noir sur noir : $-1 \times -1 = 1$). Ensuite, la moyenne des neuf pixels est calculée (maximum 1 si les neuf pixels correspondent, minimum -1 s'il n'y a aucune correspondance). Finalement, une carte de correspondance entre le filtre et l'image à analyser est créée (Figure 6(c)). L'opération est réalisée autant de fois que le modèle à conserver de filtres lors de l'entraînement (Figure 6(d)). Dès lors, l'image perd des pixels en hauteur et en largeur (puisque 9 pixels n'en deviennent qu'un), mais gagne énormément en profondeur, par superposition des différentes cartes de correspondances avec les filtres.

Couche de pooling

Généralement après la couche convolutionnelle (réalisant ces filtrages), le réseau passe par une couche de pooling (Figure 6(e)). Lors de cette étape, les pixels des images filtrées sont agglomérés afin de réduire les dimensions de l'image. La taille de l'agglomérat (ici, 2x2 pixels n'ont deviendront qu'un seul) et le type d'agglomération sont aussi dépendant du design du modèle (par exemple, l'agrégation par la valeur maximale de l'agglomérat est le « max pooling »). En plus de contrer l'augmentation de la quantité pixels générée par la couche convolutionnelle, ce sous-échantillonnage permet aussi d'éviter le surajustement du modèle qui serait à des effets de position du sujet dans l'image (Voulodimos et al., 2018). Cela crée donc une invariance compte tenu de l'échelle et de la distorsion entre les images (LeCun et al., 1990).

Niveaux de complexité du CNN

Puisque la complexité du modèle dépend du nombre de couches, les filtres qui passent sur l'image à analyser en début de réseau sont très simple pour l'ordinateur et complexes pour l'œil humain. Il s'agit de bord, de coin, et de tout autre variation soudaine entre des valeurs de pixels voisins. Plus les couches se succèdent, plus les filtres retenus pour le modèle sont complexes pour l'ordinateur et simples pour l'œil humain. Si le but était de détecter des personnes, le modèle chercherait alors des visages en fin de réseau (Figure 6(f)).

¹³ Pour une image couleur en 8 bits, on aurait une matrice 9x9x3 car les couleurs se répartissent selon 3 canaux : rouge, vert et bleu. L'intensité de chaque couleur pour chaque pixel se répartirait entre 0 et 255 (2^8 bits).

Fonction d'activation

Après chaque convolution, les résultats passent dans une fonction d'activation appelée Rectification Linear Unit (ReLU) qui permet simplement de transformer les valeurs négatives en zéros afin d'éviter les problèmes mathématiques liés à des valeurs infinies ou extrêmement petites (Figure 6(h)).

Les couches entièrement connectées

Finalement, l'information produite passe par un dernier type de couche au sein du réseau. Il s'agit des couches entièrement connectées, qui correspondent à celles des perceptrons multicouches. Pour se faire, toutes les matrices de pixels créées et transformées tout au long du réseau sont assemblées et aplaties en un long vecteur. Dès la première couche de convolution, chaque pixel appartient donc à un filtre provenant soit d'une partie d'un « X », soit d'une partie d'un « O ». Tout au long du réseau, ce niveau de complexité a augmenté tout en gardant en mémoire à quelle classe le pixel appartenait. Dès lors, un vote pondéré se fait entre l'ensemble des pixels du vecteur pour l'une ou l'autre classe (Figure 6, (h)).

c. Différentes méthodes d'apprentissage

Selon les données à analyser, la façon d'utiliser ces modèles change (Figure 21). En général des sets de données de taille considérable existent et regroupent des images très variées. Si les données à analyser sont du même type que ces datasets, il convient alors de se baser sur des modèles qui ont été entraînés dessus afin de ne pas refaire le travail plusieurs fois. Deux cas de figure peuvent alors se présenter : si la quantité de données à analyser est faible, utiliser le modèle tel quel permettra sans doute de fournir les meilleurs résultats. C'est ce qu'on appelle « transfer learning ». Si la quantité d'images est grande, il peut être plus efficace de « fine-tuner » le modèle. Cela consiste à réentraîner uniquement les dernières couches du réseau. Pour rappel, les premières couches détectent des caractéristiques très générales, communes à grand nombre d'images comme les bords et les courbes. En revanche, les dernières couches sont bien plus spécifiques et distinguent les textures et les éléments propres à chaque image (Figure 6 (f)). De plus, utiliser cette méthode pour un dataset réduit conduit généralement à un surajustement du modèle par rapport aux données d'entraînement, surtout si les dernières couches du modèle sont entièrement connectées (Russakovsky et al., 2015).

Annexes

Si le dataset à étudier présente des données très spécifiques (par exemple des données médicales ou de la classification de hiéroglyphes) il veut mieux entraîner le modèle à partir de zéro afin que celui-ci soit spécialisé dans ce type d'analyses.

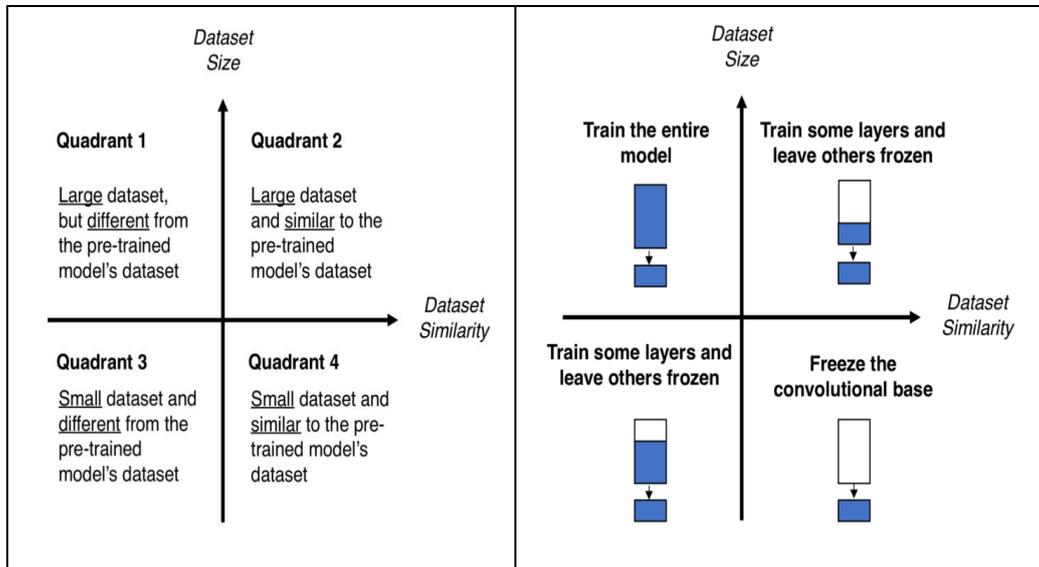


Figure 21: Quelle méthode d'apprentissage choisir en fonction du dataset disponible? Tiré de Marcelino (2018)

Annexe 2 : Protocole d'utilisation de l'indice et pistes d'améliorations

1) Acquisition des photos à analyser et préparation des outils

- Avant de commencer à suivre ce protocole, les données photographiques doivent déjà être en votre possession et classées selon votre convenance, le plus simple étant de la classer par lieux et par dates de relevé. La qualité des photos dépend des remarques faites dans le corps du document principal.
- Une base de données SQL doit être préparée afin de pouvoir recevoir les métadonnées des photos (cette partie n'est pas décrite ici).
- Si vous travaillez sur un serveur distant
 - Être connecté au VPN grâce à Virtual Network¹⁴ si le travail se fait sur un serveur de l'ULiège
 - Avoir téléchargé un logiciel de Virtual Network Computing (VNC) tel que VNC Viewer¹⁵ afin de pouvoir naviguer sur le serveur distant

¹⁴ https://www.ulg.ac.be/cms/c_54680/fr/service-vpn

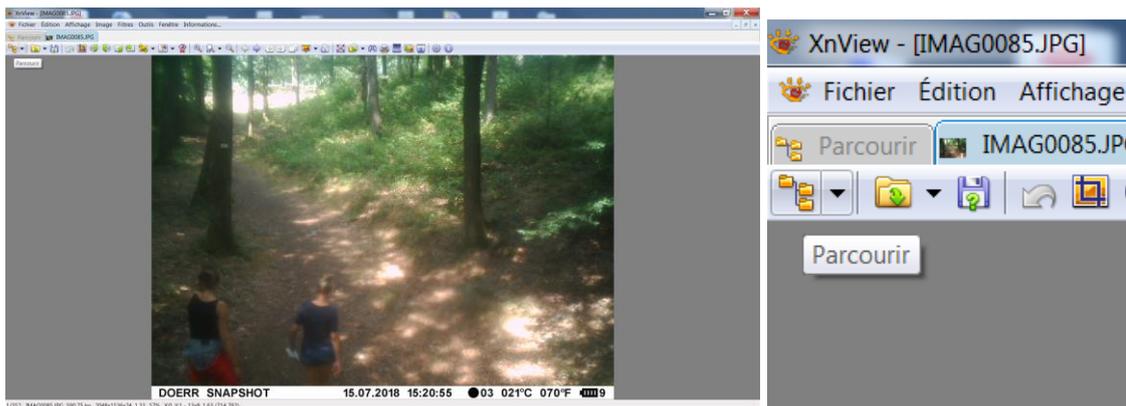
¹⁵ <https://www.realvnc.com/fr/connect/download/viewer/>

Annexes

- Avoir téléchargé un client FTP tel que WinSCP¹⁶ ou FileZila pour pouvoir uploader les fichiers sur le serveur distant
- Avoir téléchargé le répertoire du modèle sur internet (ici, Mask R-CNN¹⁷) ainsi que les poids résultant de l'entraînement du modèle sur le dataset COCO (voir TFE)¹⁸
- Avoir demandé à l'informaticien d'installer les packages nécessaires au fonctionnement du modèle sur le serveur distant (disponible sous le nom de « requirements.txt » dans le dossier du modèle)

Prétraitement et organisation des fichiers (par M. Dufrière)

- Renommer les fichiers pour récupérer nom, date et heure de prise de vue dans le nom
- Copier les images dans un dossier avec un nom court (sans accent ni blanc) par exemple (Barrage_Nisramont, Chesle, ...)
- Ouvrir une des photos avec le logiciel XNView¹⁹ (gratuit)
- Cliquer sur la liste des fichiers en haut à droite pour obtenir la liste de toutes les images



- Sélectionner toutes les images (CTRL+A)
- Cliquer Menu Outils / Renommer par lot et écrire la séquence suivante pour renommer les fichiers => <Filename>_<Date Taken [Y-m-d_H-M-S]>

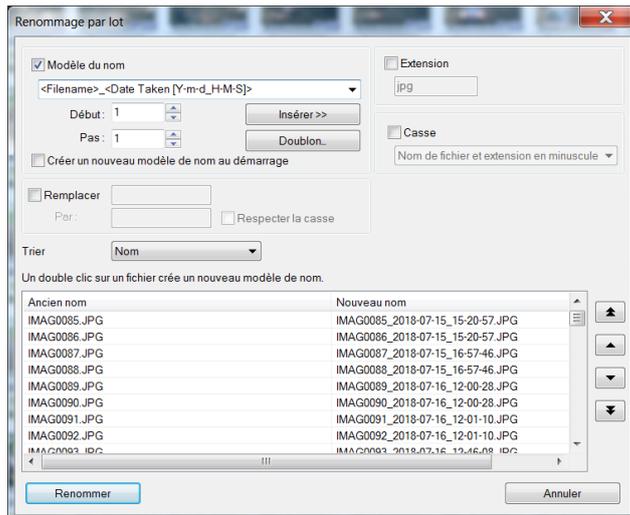
¹⁶ <https://winscp.net/eng/docs/lang:fr>

¹⁷ https://github.com/matterport/Mask_RCNN

¹⁸ https://github.com/matterport/Mask_RCNN/releases/download/v2.0/mask_rcnn_coco.h5

¹⁹ <https://www.xnview.com/fr/>

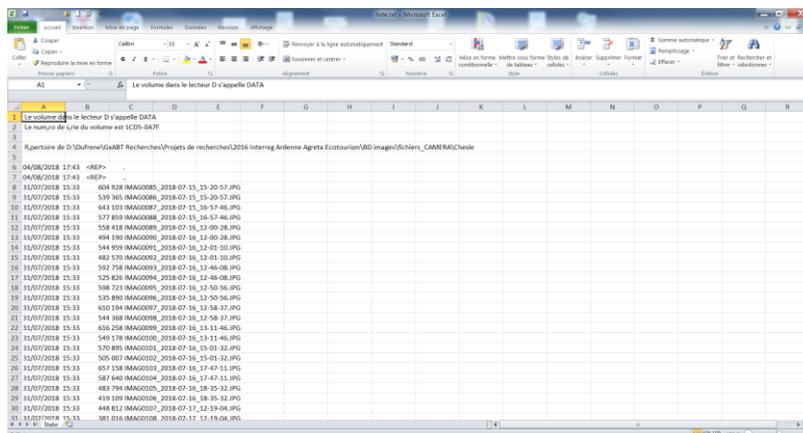
Annexes



- Cliquer sur renommer

Créer la liste des fichiers d'images pour l'importation dans la BD MySQL (par M. Dufrene)

- Cliquer sur le nom du dossier contenant les images en appuyant sur la touche Maj/Shift de manière à pouvoir « ouvrir une fenêtre de commande ici ».
- Dans la fenêtre de commandes, taper `dir >liste.txt`
- Ouvrir avec XLS le fichier liste.txt.



L'objectif est d'avoir un fichier XLS avec 12 colonnes

Ordre	Nom de la variable	Type	Description	Contenu au départ
1	ID_images	int(11)	Code univoque de l'image dans la BD	0 par défaut, autoincrémentée
2	Directory	varchar(250)	Nom du dossier où se trouvent les images	Chesle, Barrage_Nisramont, ...
3	Filename	varchar(50)	Nom de l'image	IMAG0427_2018-07-29_09-56-34.JPG
4	Date	date	Date en format DATE pour MYSQL	2018-07-28 - PAS le format 2018/07/28 par défaut XLS

Annexes

5	Annee	int(11)	Annee en nombre	2018
6	Mois	int(11)	Mois	7
7	Jour	int(11)	Jour	28
8	Heure	int(11)	Heure	10
9	Minute	int(11)	Minute	34
10	Seconde	int(11)	Seconde	10
11	Description_ID	int(11)	Code de la description associée qui sera complétée avec l'interface web	0
12	Validation	int(11)	(0/1) Validation de la description	0

- On utilise alors les formules xls utilisant la fonction stxt() pour compléter les variables :

ID	Directory	Filename	Date	Année	Mois	Jour	Heure	Minute	Seconde	Description_ID	Validation
0	Chesle	IMAG0085_2018-07-15_15-20-57.JPG	2018-07-15	2018	07	15	15	20	57	0	0
10	Chesle	IMAG0086_2018-07-15_15-20-57.JPG	2018-07-15	2018	07	15	16	57	46	0	0
11	Chesle	IMAG0087_2018-07-15_16-57-46.JPG	2018-07-15	2018	07	15	16	57	46	0	0
12	Chesle	IMAG0088_2018-07-15_16-57-46.JPG	2018-07-15	2018	07	15	16	57	46	0	0

- On peut utiliser un modèle fait une fois pour toute en xls (modele_liste_fichier_camera.xls) pour recopier automatiquement les formules pour les différentes listes.
- On recopie alors le contenu dans un fichier xls qu'on sauve en format csv (séparateur ;)

ID	Directory	Filename	Date	Année	Mois	Jour	Heure	Minute	Seconde	Description_ID	Validation
0	Chesle	IMAG0085_2018-07-15_15-20-57.JPG	2018-07-15	2018	07	15	15	20	57	0	0
1	Chesle	IMAG0086_2018-07-15_15-20-57.JPG	2018-07-15	2018	07	15	15	20	57	0	0
2	Chesle	IMAG0087_2018-07-15_16-57-46.JPG	2018-07-15	2018	07	15	16	57	46	0	0
3	Chesle	IMAG0088_2018-07-15_16-57-46.JPG	2018-07-15	2018	07	15	16	57	46	0	0
4	Chesle	IMAG0089_2018-07-16_12-00-28.JPG	2018-07-16	2018	07	16	12	00	28	0	0
5	Chesle	IMAG0090_2018-07-16_12-00-28.JPG	2018-07-16	2018	07	16	12	00	28	0	0
6	Chesle	IMAG0091_2018-07-16_12-01-10.JPG	2018-07-16	2018	07	16	12	01	10	0	0
7	Chesle	IMAG0092_2018-07-16_12-01-10.JPG	2018-07-16	2018	07	16	12	01	10	0	0
8	Chesle	IMAG0093_2018-07-16_12-46-08.JPG	2018-07-16	2018	07	16	12	46	08	0	0
9	Chesle	IMAG0094_2018-07-16_12-46-08.JPG	2018-07-16	2018	07	16	12	46	08	0	0
10	Chesle	IMAG0095_2018-07-16_12-50-56.JPG	2018-07-16	2018	07	16	12	50	56	0	0
11	Chesle	IMAG0096_2018-07-16_12-50-56.JPG	2018-07-16	2018	07	16	12	50	56	0	0
12	Chesle	IMAG0097_2018-07-16_12-58-37.JPG	2018-07-16	2018	07	16	12	58	37	0	0
13	Chesle	IMAG0098_2018-07-16_12-58-37.JPG	2018-07-16	2018	07	16	12	58	37	0	0
14	Chesle	IMAG0099_2018-07-16_13-11-46.JPG	2018-07-16	2018	07	16	13	11	46	0	0
15	Chesle	IMAG0100_2018-07-16_13-11-46.JPG	2018-07-16	2018	07	16	13	11	46	0	0
16	Chesle	IMAG0101_2018-07-16_15-01-32.JPG	2018-07-16	2018	07	16	15	01	32	0	0
17	Chesle	IMAG0102_2018-07-16_15-01-32.JPG	2018-07-16	2018	07	16	15	01	32	0	0
18	Chesle	IMAG0103_2018-07-16_17-47-11.JPG	2018-07-16	2018	07	16	17	47	11	0	0
19	Chesle	IMAG0104_2018-07-16_17-47-11.JPG	2018-07-16	2018	07	16	17	47	11	0	0
20	Chesle	IMAG0105_2018-07-16_18-35-32.JPG	2018-07-16	2018	07	16	18	35	32	0	0
21	Chesle	IMAG0106_2018-07-16_18-35-32.JPG	2018-07-16	2018	07	16	18	35	32	0	0
22	Chesle	IMAG0107_2018-07-17_12-19-04.JPG	2018-07-17	2018	07	17	12	19	04	0	0
23	Chesle	IMAG0108_2018-07-17_12-19-04.JPG	2018-07-17	2018	07	17	12	19	04	0	0
24	Chesle	IMAG0109_2018-07-17_12-42-20.JPG	2018-07-17	2018	07	17	12	42	20	0	0
25	Chesle	IMAG0110_2018-07-17_12-42-20.JPG	2018-07-17	2018	07	17	12	42	20	0	0
26	Chesle	IMAG0111_2018-07-17_12-42-47.JPG	2018-07-17	2018	07	17	12	42	47	0	0
27	Chesle	IMAG0112_2018-07-17_12-42-47.JPG	2018-07-17	2018	07	17	12	42	47	0	0
28	Chesle	IMAG0113_2018-07-17_12-42-47.JPG	2018-07-17	2018	07	17	12	42	47	0	0

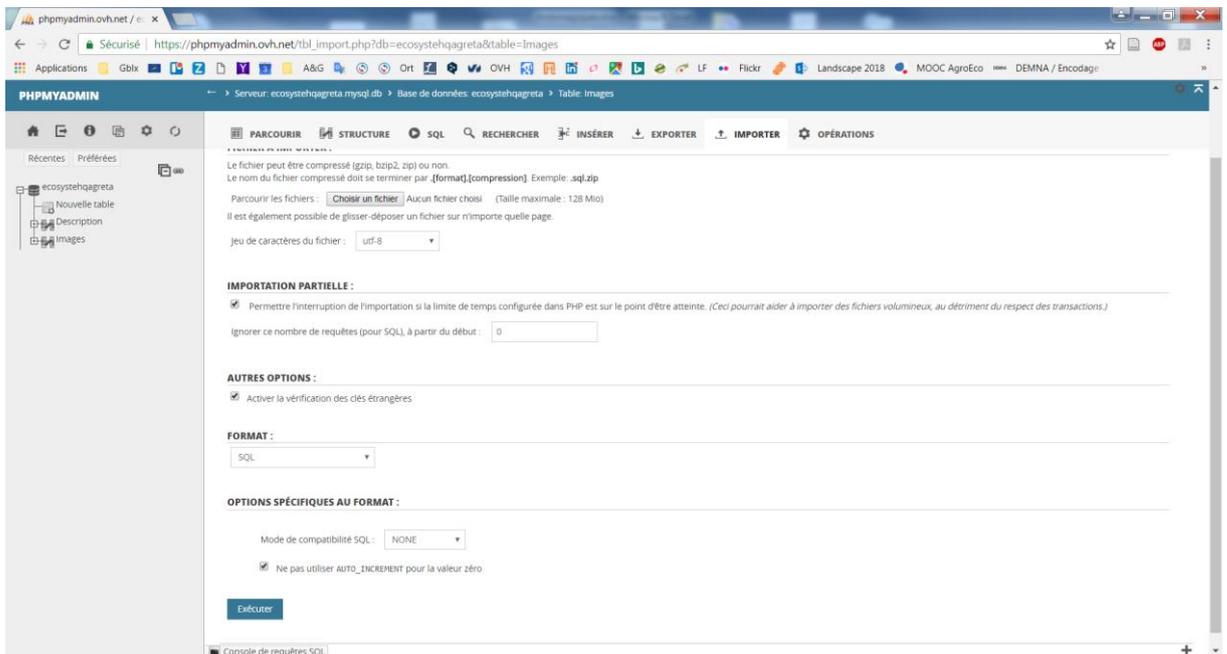
Annexes

Ce fichier est en principe prêt à être importé dans la BD MySQL. Pour éviter de nombreuses manipulations, on peut combiner des listes de différents dossiers dans un seul dossier.

Attention : TOUTES les colonnes doivent être remplies !!!

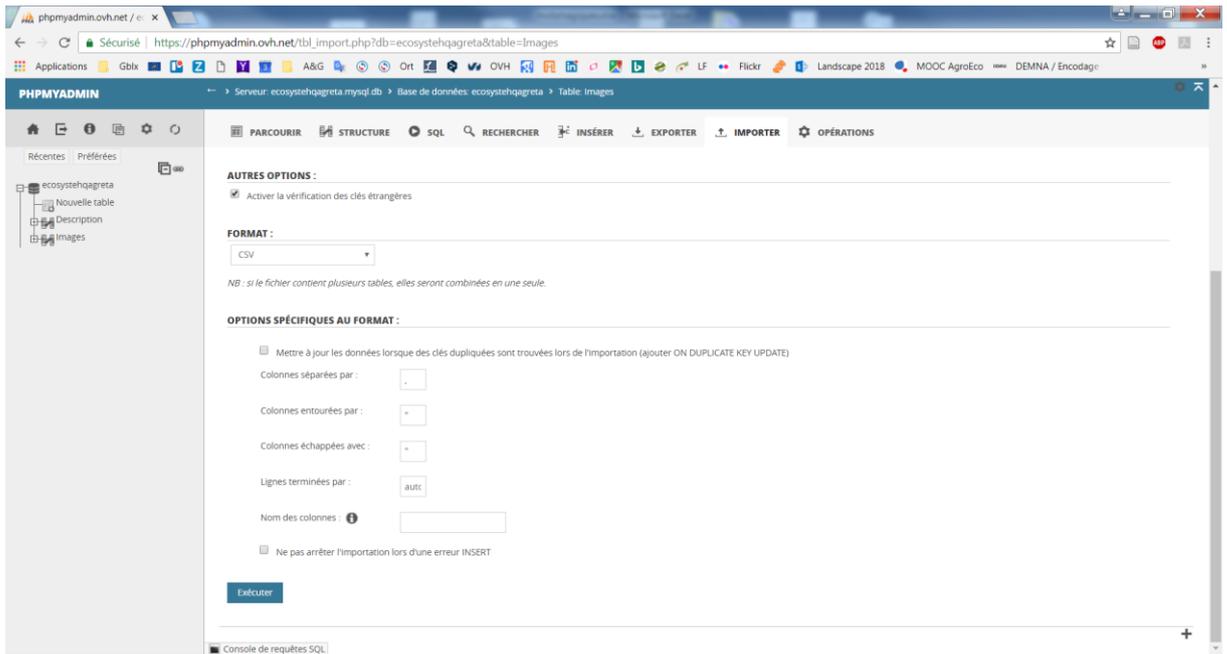
Importation dans la base de données SQL (par M. Dufrière)

- Importer la liste des images dans la BD MySQL préalablement créée.
- Ouvrir la BD MySQL
- Cliquer sur la table qui contient les données des images
- Cliquer sur « Importer »



- Choisir le fichier préparé (par exemple Chesle-Liste.csv)

Annexes

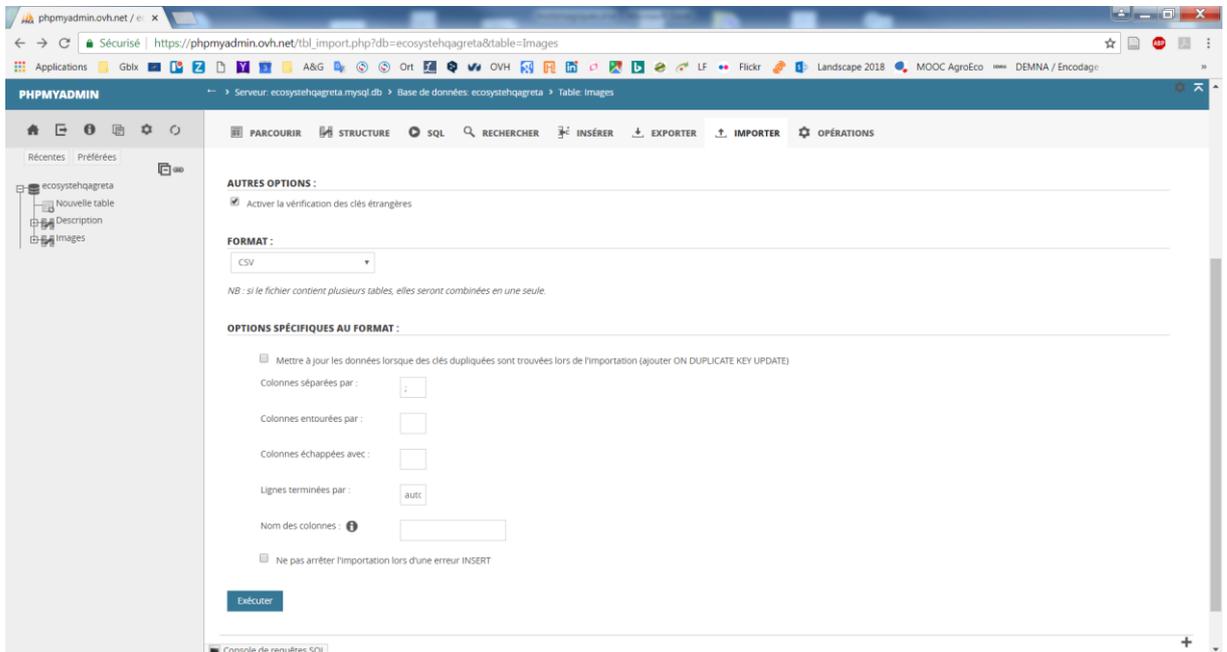


- Corriger les items :

« colonnes séparées par » => ;

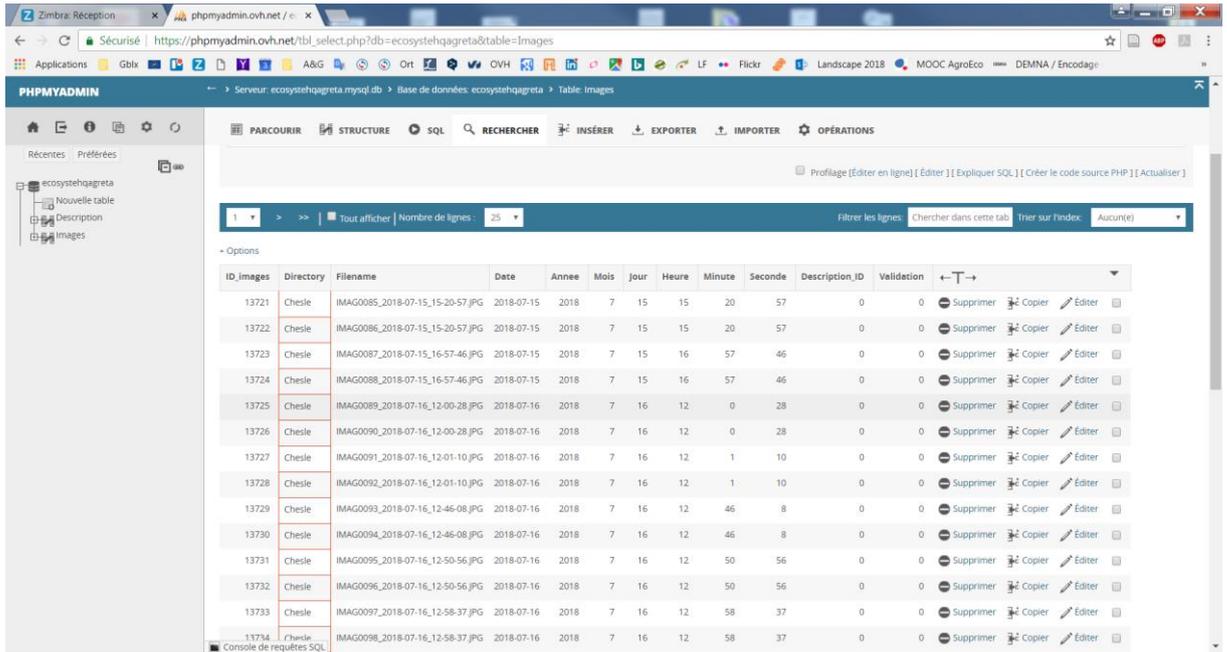
« colonnes entourées par » => rien

« colonnes échappées par » => rien



- Exécuter l'import. Normalement, les fiches doivent être importées. Vérifier ensuite que le contenu de l'import est OK. Sinon, il faut éliminer les fiches importées (via une requête SQL), corriger et recommencer ...

Annexes



ID_images	Directory	Filename	Date	Annee	Mois	Jour	Heure	Minute	Seconde	Description_ID	Validation	Actions
13721	Chesle	IMAG0085_2018-07-15_15-20-57.JPG	2018-07-15	2018	7	15	15	20	57	0	0	Supprimer Copier Editer
13722	Chesle	IMAG0086_2018-07-15_15-20-57.JPG	2018-07-15	2018	7	15	15	20	57	0	0	Supprimer Copier Editer
13723	Chesle	IMAG0087_2018-07-15_16-57-46.JPG	2018-07-15	2018	7	15	16	57	46	0	0	Supprimer Copier Editer
13724	Chesle	IMAG0088_2018-07-15_16-57-46.JPG	2018-07-15	2018	7	15	16	57	46	0	0	Supprimer Copier Editer
13725	Chesle	IMAG0089_2018-07-16_12-00-28.JPG	2018-07-16	2018	7	16	12	0	28	0	0	Supprimer Copier Editer
13726	Chesle	IMAG0090_2018-07-16_12-00-28.JPG	2018-07-16	2018	7	16	12	0	28	0	0	Supprimer Copier Editer
13727	Chesle	IMAG0091_2018-07-16_12-01-10.JPG	2018-07-16	2018	7	16	12	1	10	0	0	Supprimer Copier Editer
13728	Chesle	IMAG0092_2018-07-16_12-01-10.JPG	2018-07-16	2018	7	16	12	1	10	0	0	Supprimer Copier Editer
13729	Chesle	IMAG0093_2018-07-16_12-46-08.JPG	2018-07-16	2018	7	16	12	46	8	0	0	Supprimer Copier Editer
13730	Chesle	IMAG0094_2018-07-16_12-46-08.JPG	2018-07-16	2018	7	16	12	46	8	0	0	Supprimer Copier Editer
13731	Chesle	IMAG0095_2018-07-16_12-50-56.JPG	2018-07-16	2018	7	16	12	50	56	0	0	Supprimer Copier Editer
13732	Chesle	IMAG0096_2018-07-16_12-50-56.JPG	2018-07-16	2018	7	16	12	50	56	0	0	Supprimer Copier Editer
13733	Chesle	IMAG0097_2018-07-16_12-58-37.JPG	2018-07-16	2018	7	16	12	58	37	0	0	Supprimer Copier Editer
13734	Chesle	IMAG0098_2018-07-16_12-58-37.JPG	2018-07-16	2018	7	16	12	58	37	0	0	Supprimer Copier Editer

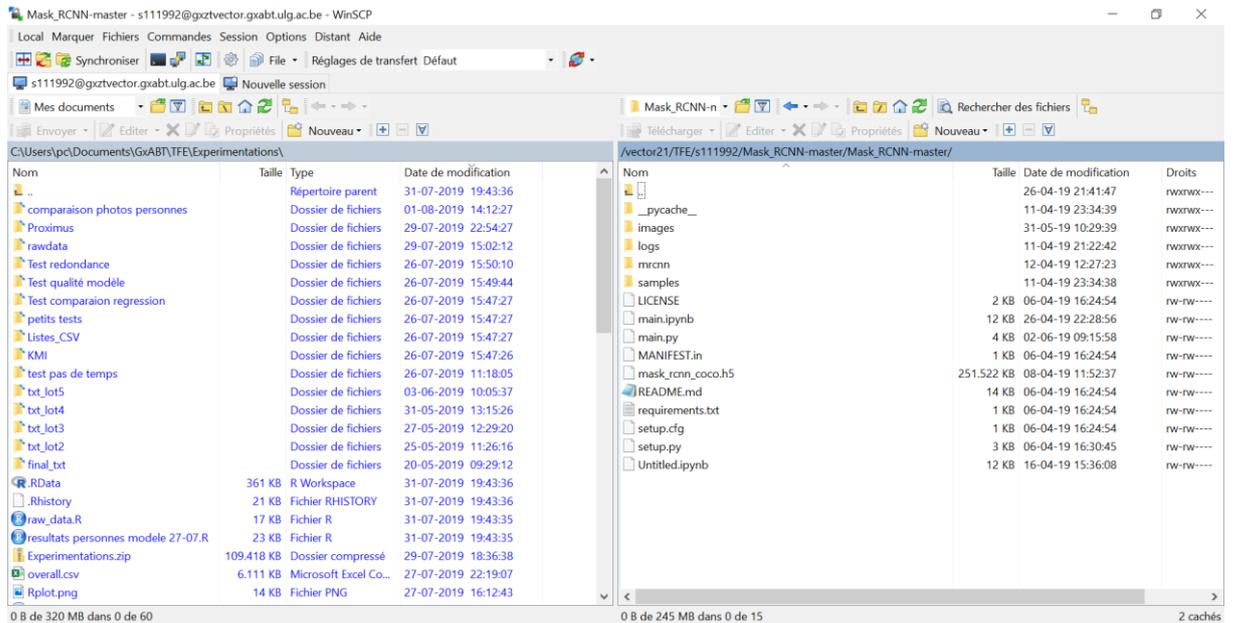
- La variable date est correcte (séparée par des « - » et non des « / »), description_ID = 0 et Validation = 0.

Upload des photos (exemple de ce qui a été fait dans ce travail)

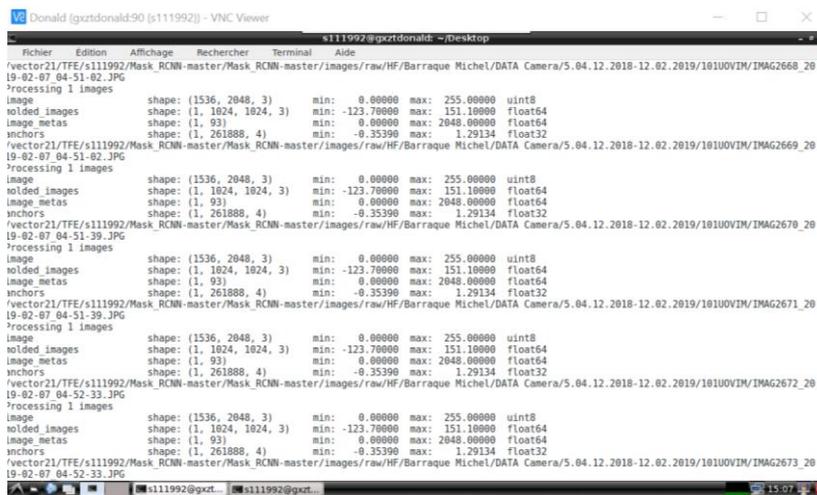
- Après avoir obtenu l'accès à la plateforme CAMI de l'ULiège (contact : Pr. N. Gengler) et être connecté au VPN de l'ULiège, se connecter au serveur distant via le client FTP utilisé.
 - S'assurer que la version de Python installée soit au moins Python 3.4
- Uploader le dossier contenant le modèle à utiliser (ici, Mask R-CNN²⁰) ainsi que les fichiers à traiter. Cela peut prendre un temps conséquent en fonction du nombre de fichiers : PREVOIR PLUSIEURS JOURS.

²⁰ https://github.com/matterport/Mask_RCNN

Annexes



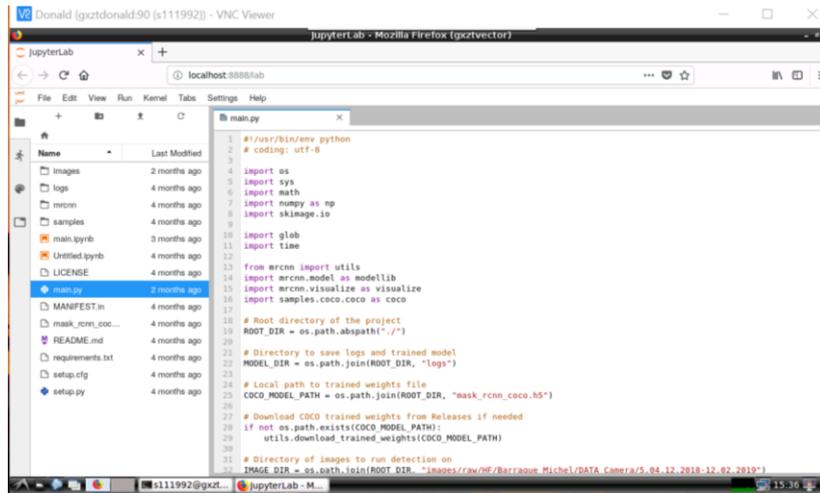
- Se connecter au logiciel VNC afin de pouvoir travailler sur le serveur distant (identifiants du serveur fournis par le Pr. Gengler ou l'informaticien Bernard Segers)



Le serveur distant fonctionne en lignes de commande Linux il est nécessaire de se familiariser un minimum avec cet OS.

- Pour naviguer dans les fichiers -> se servir du client FTP
- Pour modifier les scripts et les tester sous Linux => installer Jupyter notebook via la commande « python3 -m pip install jupyter ». Ensuite, taper « jupyter lab » dans l'invite de commandes.

Annexes



- Une fois l'accès établi, et les dossiers importés, l'architecture doit ressembler à :
- Mask R-CNN
 - __pycache__
 - **images**
 - raw
 - processed
 - txt
 - logs (inutile)
 - **mrcnn** (le modèle)
 - samples (inutile)
 - **main.py** (le script à lancer)
 - **mask_rcnn_coco.h5** (les poids du modèle)
 - (autres)
- Placer les photos non traitées dans le répertoire « raw ». Ne pas hésiter à faire des sous-dossiers à l'intérieur afin de séparer les lots à traiter. En sortie, les images traitées iront dans le dossier « processed » et les fichiers textes avec les classes assignées aux objets détectés iront dans le dossier « txt »

Alternative principale

Dans le cadre d'un consortium universitaire entre la Région Wallonne et le FNRS, il est possible de passer par des clusters distants plus puissants car certains sont équipés de GPU, contrairement au cluster du CAMI. Il s'agit du Consortium des Équipements de Calcul Intensif21 (CECI). Ces clusters sont accessibles à tous les membres des universités wallonnes et devraient permettre de minimiser grandement les temps de traitement. Cependant, le logiciel qui permet la planification des travaux entre tous les chercheurs (Slurm) nécessite un certain temps de prise en main22.

²¹ <http://www.ceci-hpc.be/>

²² <https://support.ceci-hpc.be/doc/contents/QuickStart/SubmittingJobs/SlurmTutorial.html>

Autres alternatives

- Selon le type de fichiers à traiter, il est possible d'acheter directement une carte graphique équipée d'un GPU pour son ordinateur personnel. Selon David Colignon (CECI), pour des images JPEG faisant entre 300 et 600 ko, une carte graphique basique (entre 500 et 1000 euros) permettra déjà d'obtenir des résultats plus rapides qu'avec un cluster en CPU.
- L'unité des Ressources forestières du professeur P. Lejeune (GxABT) possède aussi des cluster GPU.

2) Le script Python

Lancer le traitement des images

Quand tous les fichiers (répertoire du modèle, poids du modèle, photos dans le dossier « raw ») sont bien placés sur le serveur distant et que vous êtes sur le VNC, dans l'invite de commande linux :

- Se placer à l'endroit du dossier de travail grâce à la commande « **cd ... (Pathway)** ».
 - Pour revenir en arrière : **cd ..**
 - Liste des fichiers dans le dossier : **ls**
- Vérifier que vous êtes sur la bonne version de python en tapant « **python** »
 - voir les versions de python installées : **ls /usr/bin/python***
 - Changer de version : **alias python='/usr/bin/python3.6'**
- Taper « **jupyter lab** » pour avoir accès au script à lancer et y modifier les pathways. Le script à modifier se nomme « **main.py** » et se trouve dans le dossier « Mask R-CNN ».
- Modifier les pathways pour correspondre à votre architecture (voir point 0). Sauver (CTRL+S)
- Le modèle présente un problème et n'enregistre pas les images. Pour le corriger :
 - Ouvrir « **visualize.py** » dans « **mrcnn** » avec jupyter lab, ligne 169, ajouter avec l'indentation
 - **if path_out is not None:**
plt.savefig(path_out)
- Fermer jupyter lab. Dans l'invite de commande, faire CTRL+C.
- Lancer le script : taper « **python main.py** »
- S'il y a des erreurs ... Suivre les instructions
 - Package manquant : **(-m) pip install SomePackage --user**
 - Mettre à jour un package : **(-m) pip install -upgrade SomePackage --user**

Le traitement peut prendre plus d'une semaine... (environ 3 secondes par photos sur le CAMI), préférer des petits lots pour éviter de monopoliser la mémoire vive. Répéter opération en changeant les pathways.

Annexes

Description du script

Le script en Python n'est pas directement l'architecture Mask R-CNN mais plutôt une interface qui permet de s'en servir en Transfert Learning. Le script appelé « main.py » se déroule comme suit :

A modifier

```
#!/usr/bin/env python
# coding: utf-8

#Packages
import os #Interactions avec l'OS
import sys #idem
import math #fonctions mathématiques
import numpy as np # permet la gestion des tableaux et matrices
import skimage.io # collection d'algorithmes pour traiter des images
import glob # permet d'aller chercher des fichiers dans une arborescence selon critères
import time # calcul du temps de travail

# Fonctions issues du modèle Mask R-CNN
from mrcnn import utils
import mrcnn.model as modellib
import mrcnn.visualize as visualize

# Root directory of the project
ROOT_DIR = os.path.abspath(".")

# Directory to save logs and trained model
MODEL_DIR = os.path.join(ROOT_DIR, "logs")

# Local path to trained weights file
COCO_MODEL_PATH = os.path.join(ROOT_DIR, "mask_rcnn_coco.h5")

# Download COCO trained weights from Releases if needed
if not os.path.exists(COCO_MODEL_PATH):
    utils.download_trained_weights(COCO_MODEL_PATH)

# Directory of images to run detection on (A CHANGER)
IMAGE_DIR = os.path.join(ROOT_DIR, "images/raw/HF/")
IMAGE_OUTPUT = os.path.join(ROOT_DIR, "images/processed_lot5")
TXT_OUTPUT = os.path.join(ROOT_DIR, "images/txt_lot5")
```

Annexes

```
class InferenceConfig(coco.CocoConfig):
    # Set batch size to 1 since we'll be running inference on
    # one image at a time. Batch size = GPU_COUNT * IMAGES_PER_GPU
    GPU_COUNT = 0
    IMAGES_PER_GPU = 0

config = InferenceConfig()
config.display()

### Create Model and Load Trained Weights

# Create model object in inference mode.
model = modellib.MaskRCNN(mode="inference", model_dir=MODEL_DIR, config=config)

# Load weights trained on MS-COCO
model.load_weights(COCO_MODEL_PATH, by_name=True)

#Class names
class_names = ['BG', 'person', 'bicycle', 'car', 'motorcycle', 'airplane',
               'bus', 'train', 'truck', 'boat', 'traffic light',
               'fire hydrant', 'stop sign', 'parking meter', 'bench', 'bird',
               'cat', 'dog', 'horse', 'sheep', 'cow', 'elephant', 'bear',
               'zebra', 'giraffe', 'backpack', 'umbrella', 'handbag', 'tie',
               'suitcase', 'frisbee', 'skis', 'snowboard', 'sports ball',
               'kite', 'baseball bat', 'baseball glove', 'skateboard',
               'surfboard', 'tennis racket', 'bottle', 'wine glass', 'cup',
               'fork', 'knife', 'spoon', 'bowl', 'banana', 'apple',
               'sandwich', 'orange', 'broccoli', 'carrot', 'hot dog', 'pizza',
               'donut', 'cake', 'chair', 'couch', 'potted plant', 'bed',
               'dining table', 'toilet', 'tv', 'laptop', 'mouse', 'remote',
               'keyboard', 'cell phone', 'microwave', 'oven', 'toaster',
               'sink', 'refrigerator', 'book', 'clock', 'vase', 'scissors',
               'teddy bear', 'hair drier', 'toothbrush']

# Run Object Detection

image_data = dict() # new empty dictionary

# Load an image from the images folder (A CHANGER EN FONCTION DU PATHWAY)
# à la fin de la commande : *.JPG signifie que l'ordinateur doit sélectionner tous les fichiers finissant par .JPG
# **/ et "recursive = TRUE" signifient qu'il doit aussi chercher dans tous les sous-dossiers.
```

Annexes

```
file = glob.glob("/vector21/TFE/s111992/Mask_RCNN-master/Mask_RCNN-
master/images/raw/HF/**/*.*.JPG",recursive=True)

for file_names in file:
    print (file_names)
    image = skimage.io.imread(os.path.join(IMAGE_DIR, file_names))

    # Run detection

    detected_objects = model.detect([image], verbose=1)
    image_data[file_names] = detected_objects # ajoute au dico
    file_name =os.path.basename(file_names)

    #Visualize results
    r = detected_objects[0]

    # ECRITURE DU FICHIER TXT- Même filename que pour l'image (donc unique) mais avec l'extension .txt
    with open("/vector21/TFE/s111992/Mask_RCNN-master/Mask_RCNN-
master/images/txt_lot5"+'/'+'file_name.split('.')[0]+'+'.txt','w') as txt_file:
        for class_id,score in zip(r['class_ids'],r['scores']):
            txt_file.write(str(class_id)+'\t'+str(score)+'\n')

    #ENREGISTREMENT DE L'IMAGE TRAITEE DANS LE DOSSIER PROCESSED
    visualize.display_instances(image, r['rois'], r['masks'], r['class_ids'], class_names,
IMAGE_OUTPUT+'/'+'file_name, r['scores'])

#END#
```

3) Création base de données

Récupérer les données

- Ouvrir le client FTP et se connecter au serveur distant
- Récupérer les fichiers texte créés
- Ouvrir le base de données SQL
- Exporter la table aux format.CSV (ce sont les métadonnées de toutes les photos)

Script R création base de données (exemple sur le massif de Saint-Hubert)

```
#####
##### CREATION BDD ST HUBERT #####
#####
rm(list = ls())
#####
# Packages
```

Annexes

```
#####  
library(dplyr)  
library(stringr)  
library(lubridate)  
  
# Ouverture liste images  
#####  
setwd("C:\\Users\\pc\\Documents\\GxABT\\TFE\\Experimentations")  
  
bdd_images <- read.csv(file="Images.csv",header=TRUE)  
bdd_images <- unique(bdd_images)  
  
#####  
#Gestion des fichiers txt  
#####  
#Nouveau WD pour travailler sur les txt  
setwd("C:/Users/pc/Documents/GxABT/TFE/Experimentations/final_txt/txt")  
  
#création liste fichiers textes pour la boucle  
list_files <- list.files()  
  
#Les photos sans aucunes détection créent des fichiers textes vides. Il faut créer une ligne de 0  
# pour que R comprenne qu'il faut les traiter comme les autres  
#-----  
#info à écrire dans .txt vides  
append_to_empty_txt <- t(c(0,0))  
  
for (i in list_files) {  
  print(i)  
  if (file.info(i)$size == 0) {write.table (append_to_empty_txt, file = i,sep = "\t",  
                                           row.names = FALSE, col.names= FALSE)  
  } else  
  next  
}  
  
# Fusion des fichiers txt  
#-----  
rm(full_txt) # sinon, les data s'additionnent  
filename = c() #empty vector to store 1 filename per observation  
  
for (i in list_files){  
  
  # if the merged dataset doesn't exist, create it  
  if (!exists("full_txt")){  
    full_txt <- read.table(i, header=FALSE, sep="\t")
```

Annexes

```
tmp_lines <- nrow(full_txt) # amount of lines in my solo txt file
tmp_filename <- rep(i,tmp_lines) # ajout du filename x nb de lignes du fichier

filename <- c(filename, tmp_filename)

# if the merged dataset does exist, append to it
}else{
tmp_dataset <- read.table(i, header=FALSE, sep="\t")

tmp_lines <- nrow(tmp_dataset) # amount of lines in my solo txt file
tmp_filename <- rep(i,tmp_lines) # ajout du filename x nb de lignes du fichier

full_txt <- rbind(full_txt, tmp_dataset)
rm(tmp_dataset)

# Create a vector with 1 filename per observation
filename <- append(filename, tmp_filename)

#supprimer les attributs temporaires
rm(tmp_filename)
rm(tmp_lines)
rm(i)
}
}

# fusion infos txt et leur filename
full_txt <- cbind (filename, full_txt)

# ajout des attributs de chaque classe
# Classes disponibles en annexe
classes <- read.table("C:/Users/pc/Documents/GxABT/TFE/Experimentations/mask_rcnn_classes.txt",
                    header=FALSE, sep="\t")
full_txt <- left_join(full_txt, classes, by= c("V1", "V1"))
#-----

# JOINTURE FULL DATABASE

#-----

# Il faut d'abord enlever le .txt et le .jpg des filename des deux tables (package stringr)
full_txt$filename <- str_sub(full_txt$filename, end = -5)
bdd_images$Filename <- str_sub(bdd_images$Filename, end = -5)

#jointure
data <- inner_join(bdd_images, full_txt, by= c("Filename" = "filename"))
```

Annexes

```
# Pourquoi après ne jointure interne il y a plus de lignes que "full_txt" ?
# certaines métadonnées on été encodées deux fois dans la base de données SQL
#SOLUTION DEJA PRESENTE PLUS BAS: on supprime la redondance

#nettoyage de la base de données
data <- data[,-c(1,11,12)] # colonnes inutiles

# Noms colonnes
names(data)[1] <- "Emp_cam"
names(data)[12] <- "class"
names(data)[11] <- "Top_1"
names(data)[10] <- "raw_class"

# Oubli: ajouter colonne "Massif"
Massif <- as.character(c(0,0))

for (i in 1:nrow(data)){

  if (data$Emp_cam[i] == "Stand_De_Tir" | data$Emp_cam[i] == "Fagne_Jean_Simon" | data$Emp_cam[i] == "Etang_Fabrique"
      | data$Emp_cam[i] == "Pont" | data$Emp_cam[i] == "Vallee" ){
    Massif[i] <- "Anlier"
  } else if (data$Emp_cam[i] == "Baraque_Michel" | data$Emp_cam[i] == "Botrange" | data$Emp_cam[i] == "Bout"
             | data$Emp_cam[i] == "Polleur" | data$Emp_cam[i] == "Pont_Marie" ) {
    Massif[i] <- "HF"
  } else if (data$Emp_cam[i] == "Barrage_Nisramont" | data$Emp_cam[i] == "Chesle" | data$Emp_cam[i] == "Engreux"
             | data$Emp_cam[i] == "Herou" | data$Emp_cam[i] == "Plateau_Des_Tailles" ) {
    Massif[i] <- "PNDO"
  } else if (data$Emp_cam[i] == "Beyoli" | data$Emp_cam[i] == "Diglette_Fourneau" | data$Emp_cam[i] == "Bilaude"
             | data$Emp_cam[i] == "Masblette_Pont_Mauricy" | data$Emp_cam[i] == "Priesse" ) {
    Massif[i] <- "St Hubert"
  } else
    next
}
data <- cbind(Massif, data)
#NB: Massif = NA correspondent aux photos du Brame du Cerf!, pas prises en compte dans ce travail
# Enregistrer pour ne pas devoir refaire tourner le code
#write.csv2(data, "C:\\Users\\pc\\Documents\\GxABT\\TFE\\Experimentations\\Database_lot1.csv")
#-----

# EVITER REDONDANCE

#-----
# Créer base non redondante pour chaque massif
# ICI: St Hubert
```

Annexes

```
# ouvrir Database_lot1 et ne garder que St-Hubert

data3 <- read.csv2(file='Database_lot1.csv',header=TRUE)
data3 <- subset(data3, Massif=="St Hubert")

#####
#####

#On sait que:
# Chaque sequence = 2 photos, mais il peut y avoir plus que deux lignes si plus d'un objet dans la photo
# Toujours un nombre paire de lignes avec date et moment similaires (car les deux photos)
# pour ne rien compter en double, il faudrait faire la moyenne (/2) de chaque séquence.
# Si on tombe sur des chiffres décimaux, on ajuste à l'entier supérieur
# Exemple: il trouve un sac à dos sur l'une des deux photos,
#il y a plus de chances que le sac à dos soit vraiment là

#####
#####

# Je groupe les données selon séquence et classe et je les additionne
data2 <- data3 %>% group_by(Emp_cam,Date,Heure,Minute,Seconde,class) %>% add_tally()
# je fais la moyenne des deux photos, en arrondissant à l'entier supérieur
#-----
# ATTENTION : ne pas runner cette ligne plusieurs fois sinon on fait des moyennes de moyennes !
#-----
data2$n <- ceiling(data2$n/2)

#Je supprime les lignes redondantes
data2 <- data2 %>% distinct(Emp_cam,Date,Heure,Minute,Seconde, class, .keep_all = TRUE)

# NE TRAVAILLER QUE SUR LES PERSONNES
# *****
# il faut aussi éviter les erreurs d'année (certaines photos sont datées en 2013)
sthub <- subset(data2, data2$class == "person" & (data2$Annee == 2018 | data2$Annee == 2019) )
sthub <- sthub[,-c(1,12,13,14)]
# *****
#Maintenant, il faut supprimer les multiples de quand les gens stagent devant la caméra
# *****
#Tout d'abord, on va classer les occurences dans le temps
sthub <- sthub[order(sthub$Emp_cam, sthub$Annee, sthub$Mois,sthub$Jour,sthub$Heure,sthub$Minute,sthub$Seconde),]

#Les définir comme une série chronologique (Package lubridate)
time <- paste(sthub$Heure, sthub$Minute, sthub$Seconde, sep="-")
time <- paste(sthub$Date,time,sep=" ")

time <- ymd_hms(time)

#ajout à la database
```

Annexes

```
sthub <- cbind(time,sthub)

# REPETER CE BLOC JUSQU'A CE QU IL N Y AI PLUS DE "YES"
# _____
# _____
#calcul intervalle de temps entre les lignes
itv<- c()

for (i in 1:nrow(sthub2)){

  itv_tmp <- interval(sthub2$time[i],sthub2$time[i+1])
  itv <- rbind(itv,itv_tmp)
  rm(itv_tmp)
}

itv <- ifelse(itv >= 0,itv,9999)
itv <- ifelse(is.na(itv), 9999,itv )

#suppression colonne pour la mettre à jour
sthub2$itv <- NULL
sthub2 <- cbind(sthub2, itv)

#mise à jour colonne redondance (PAS DE TEMPS MINIMUM CHOISI à 20 SECONDES)
sthub2$redondance <- as.factor(ifelse(sthub2$itv <= 20, "yes","no"))

#enlever les lignes

for (i in 1:nrow(sthub2)){

  if (sthub2$redondance[i] == "yes"){

    if(sthub2$n[i] >= sthub2$n[i+1]){
      sthub2 <- sthub2[-(i+1),]

    }else{
      sthub2 <- sthub2[-i,]
    }

  } else
    next
}
# _____
# _____
# FIN DU BLOC

# CREATION BASE DE DONNEES CLEAN
```

Annexes

```
#-----  
#write.csv2(sthub2, "C:\\Users\\pc\\Documents\\GxABT\\TFE\\Experimentations\\sthubert.csv" )  
  
# END
```

Annexe 3 : Mask R-CNN

Le modèle Mask R-CNN (Figure 23) est très similaire à son prédécesseur, Faster R-CNN (précédemment R-CNN puis Fast R-CNN). Comme son nom l'indique, « Mask » signifie qu'en plus de la détection d'objet par boîtes, ce dernier est segmenté de façon précise, pixel par pixel formant donc ce masque (Figure 22). Les R-CNN sont un groupe de réseaux neuronaux convolutionnels où le filtre passe sur des régions de l'image, pas sur l'image entière (Region based) (Girshick et al., 2014). Cela permet de pouvoir détecter et classifier plusieurs objets sur

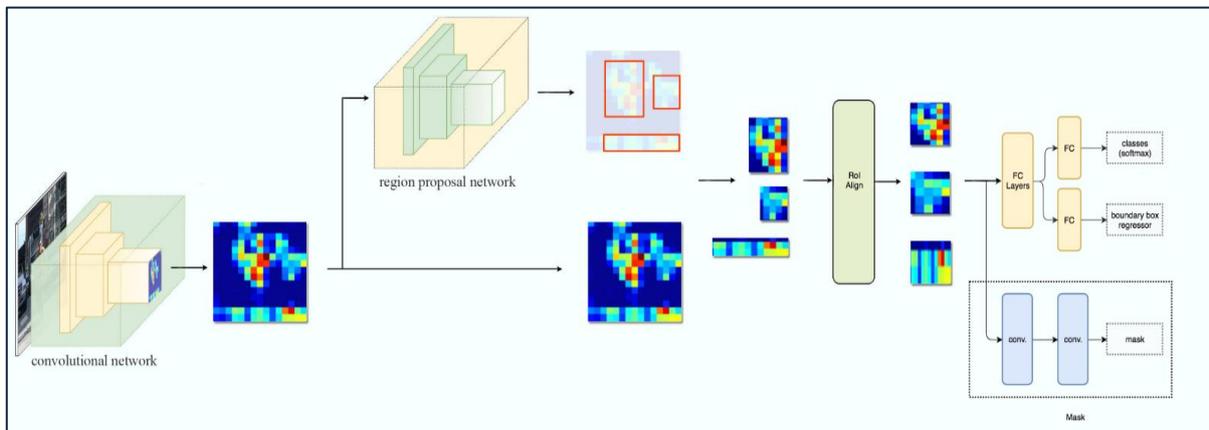
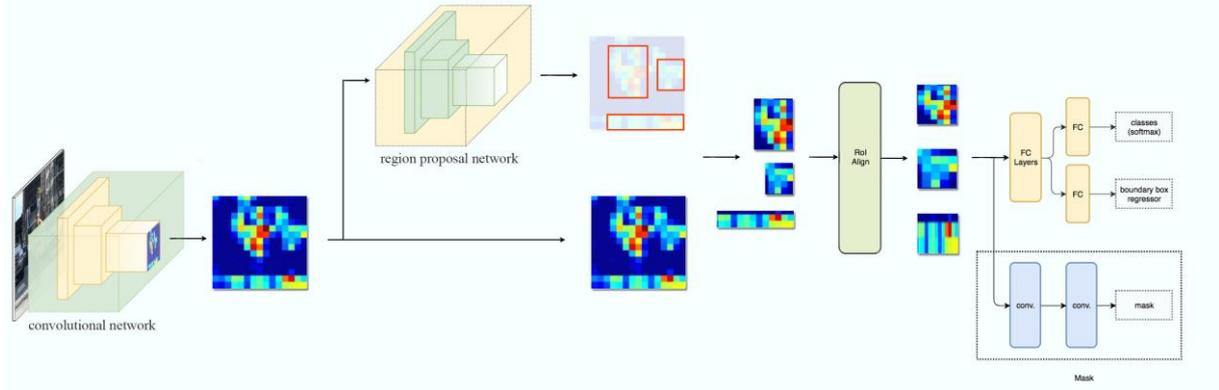


Figure 23: Architecture de Mask R-CNN. Très similaire au modèle Faster R-CNN (Ren et al., 2015), la différence majeure se situe dans les deux couches convolutionnelles supplémentaires permettant de créer le masque, cette partie est visible dans le rectangle en pointillés en bas à droite du schéma.



Figure 22: Exemple de photo traitée par Mask R-CNN, Photo prise à la Baraque Michel, Hautes Fagnes, le 29 janvier 2019.

une seule image et de minimiser le nombre de calculs.



1) Détection des objets

Pour cette partie du modèle, Mask R-CNN est similaire à son prédécesseur, Faster R-CNN (Ren et al., 2016). Il est donc nécessaire d'expliquer ce dernier (Figure 25). Cette architecture est née du succès des R-CNN grâce à la recherche sélective des régions d'intérêt (Region of Interest, ROI) (Uijlings et al., 2013), qui présentait un défaut majeur : la vitesse d'analyse des images puisqu'elle générait 2000 ROI par image. L'auteur a alors amélioré le processus dans l'architecture qu'il a appelé Fast R-CNN (Girshick, 2015). Dans celle-ci, il a simplement inversé deux étapes : l'image entière passe dans le CNN, le réseau donne ensuite

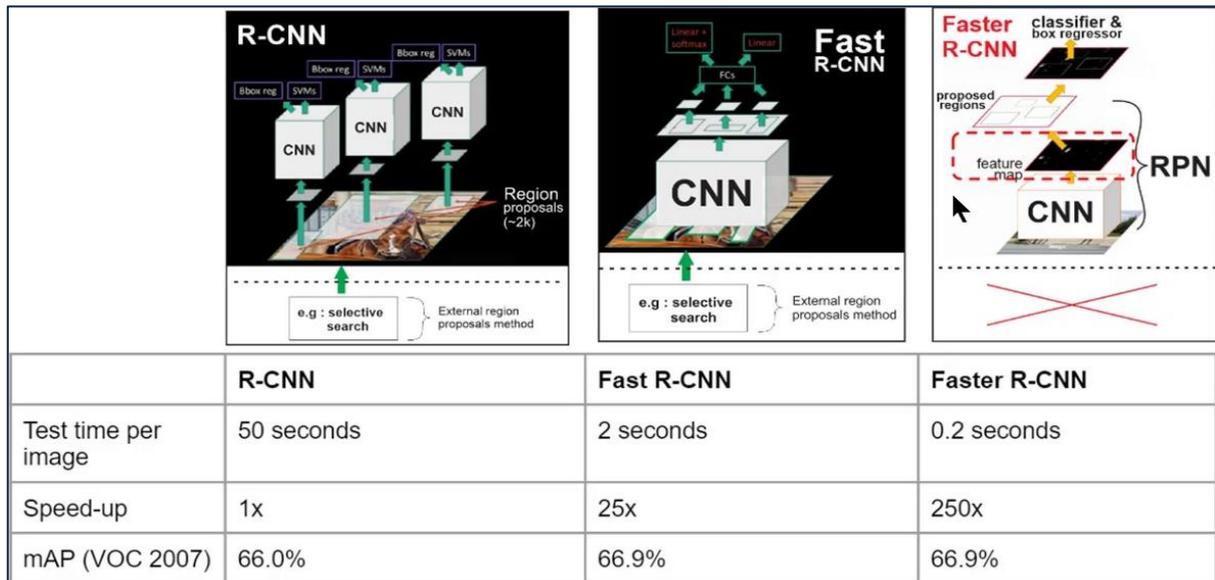


Figure 25: Comparaison des modèles de R-CNN (Li et al., 2015)

une carte des caractéristiques qui ont le plus d'impact dans l'image. Le modèle détermine les ROI grâce à cela. On évite donc de faire passer chaque ROI dans un CNN séparé.

Ce second CNN est appelé Region Proposal Network (RPN). A la place de faire passer des filtres sur l'image de base comme dans un CNN classique, on démarre ici de la carte de correspondance issue du 1^{er} CNN. De là, on fait passer une fenêtre carrée, à partir de laquelle est généré un set de neuf ancres qui permettent d'ajuster les bounding boxes par

backpropagation (Ren et al., 2016) . Ces ancres varient par leur échelle et leur ratio largeur/hauteur (Figure 26).

Finalement, ces régions passent dans les dernières couches entièrement connectées du modèle. Ces dernières servent à préciser les dimensions des bounding boxes par régression et à classifier l'objet se trouvant à l'intérieur. Ici grâce à une fonction softmax puisqu'il peut y a plus de deux classes dans l'image. En effet, la fonction softmax étend la décision binaire issue d'une fonction logistique, à une décision binaire pour chaque classe, avec l'unité comme somme des probabilités de l'ensemble des classes (Mikolov et al., 2011). C'est donc la meilleure façon de classifier un objet parmi de nombreuses classes.

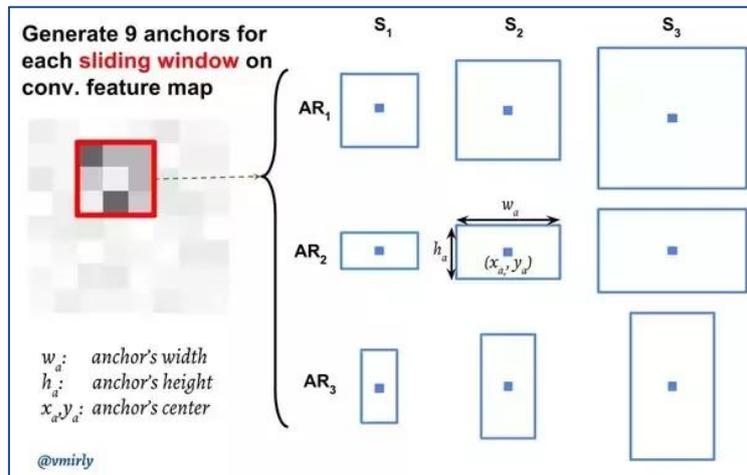


Figure 26: Dans le Region Proposal Network, l'amélioration de la vitesse de process est due à la génération de 9 ancres pour chaque fenêtre qui passe sur la carte de correspondance. Cela permet d'ajuster la précision des bounding boxes permettant la détection des objets dans le modèle général (Mirialili. 2017)

2) Segmentation des objets

Cette composante de l'architecture de Mask R-CNN utilise un réseau entièrement convolutionnel (Fully Convolutionnal Network, FCN, Figure 27) (Long et al., 2015). A la

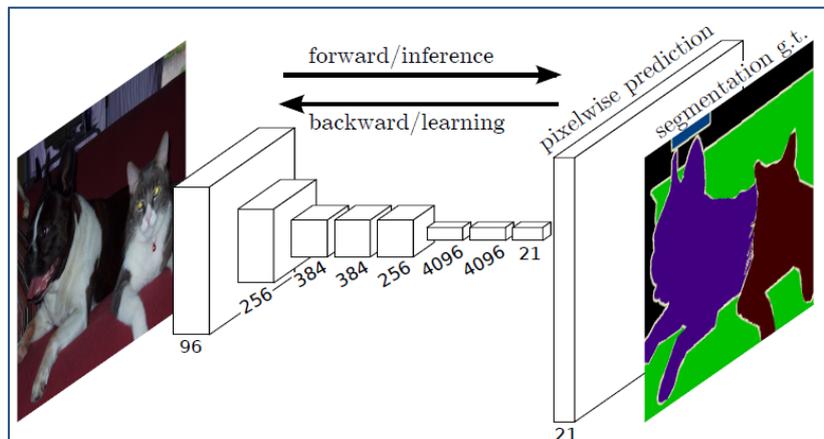


Figure 27: Fully convolutionnal network (FCN) permettant de segmenter les objets détectés par Mask R-CNN (Long et al., 2015)

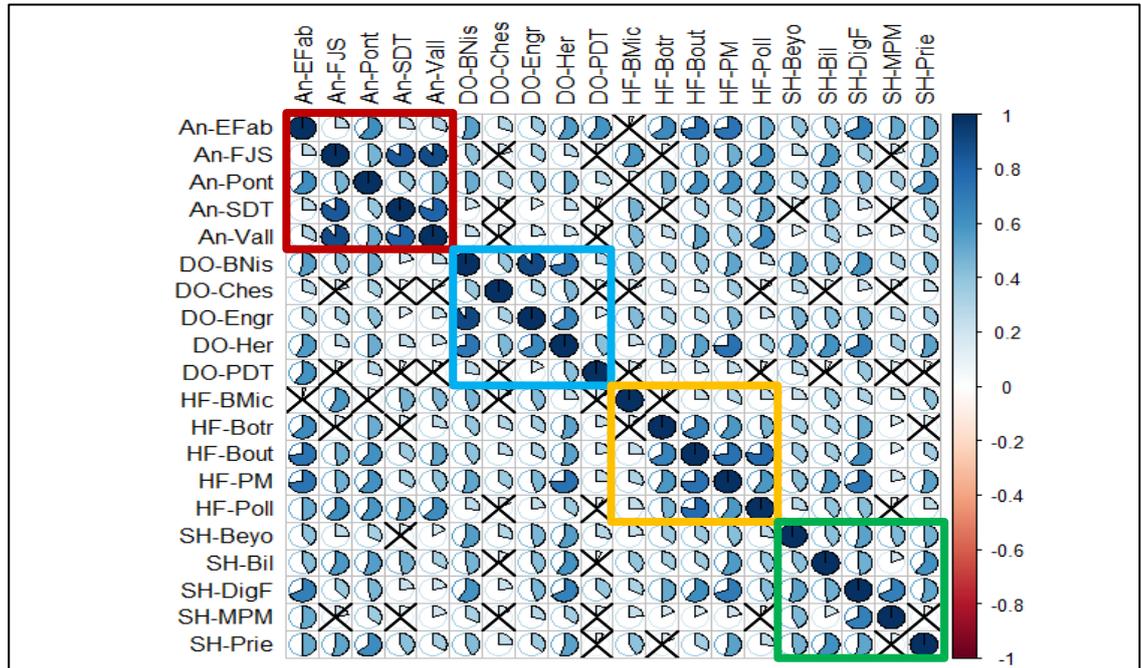
Annexes

différence du CNN classique, le FCN ne présente pas de couches entièrement connectées (fully connected layers) à la fin du réseau, mais bien des couches convolutionnelles. La classification des objets se fait donc pixel par pixel et cette responsabilité appartient aux couches convolutionnelles. Cela permet de prendre en compte la position des pixels sur l'image, puisque ceux-ci ne sont plus vectorisés comme dans une couche entièrement connectée. On utilise donc des informations locales plutôt que globale (Tsang, 2018).

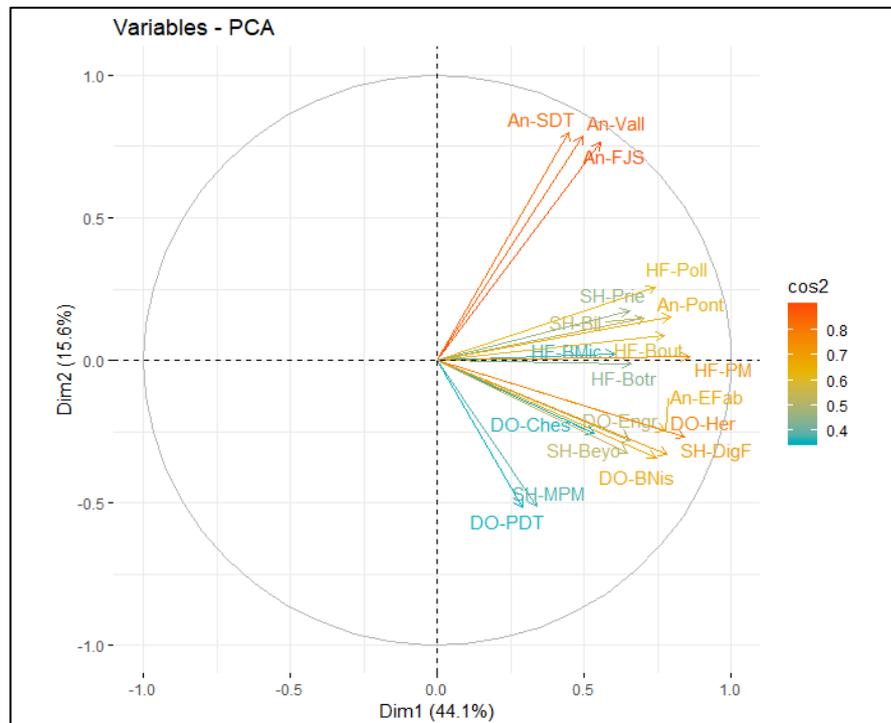
Annexe 4 : Liste des classes connues par Mask R-CNN

0	empty	24	zebra	47	cup	69	desk
1	person	25	giraffe	48	fork	70	toilet
2	bicycle	26	hat	49	knife	71	door
3	car	27	backpack	50	spoon	72	tv
4	motorcycle	28	umbrella	51	bowl	73	laptop
5	airplane	29	shoe	52	banana	74	mouse
6	bus	30	eye glasses	53	apple	75	remote
7	train	31	handbag	54	sandwich	76	keyboard
8	truck	32	tie	55	orange	77	cell phone
9	boat	33	suitcase	56	broccoli	78	microwave
10	traffic light	34	frisbee	57	carrot	79	oven
11	fire hydrant	35	skis	58	hot dog	80	toaster
12	street sign	36	snowboard	59	pizza	81	sink
13	stop sign	37	sports ball	60	donut	82	refrigerator
14	parking meter	38	kite	61	cake	83	blender
15	bench	39	baseball bat	62	chair	84	book
16	bird	40	baseball glove	63	couch	85	clock
17	cat	41	skateboard	64	potted plant	86	vase
18	dog	42	surfboard	65	bed	87	scissors
19	horse	43	tennis racket	66	mirror	88	teddy bear
20	sheep	44	bottle	67	dining table	89	hair drier
21	cow	45	plate	68	window	90	toothbrush
22	elephant	46	wine glass				
23	bear						

Annexe 5 : Résultats intermédiaires de l'étude des relations entre les caméras par rapport à leur nombre de photos



Graphique 18: Matrice de corrélations montrant l'évolution du nombre de photos par caméra en fonction du temps. Le test de significativité est réalisé par le coefficient de Pearson ($\alpha= 0,05$). Les relations non significatives sont barrées. Les carrés de couleurs représentent les différents massifs



Graphique 19: Analyse en composante principale de la relation entre les différentes caméras au fil du temps – 1er plan factoriel

Annexe 6 : Résultats du test de la qualité du modèle

Tableau 8: Résultats du test de la qualité du modèle sur 3.200 photos. Pour chaque classe (Personnes, Chiens, Sacs à dos, Véhicules, Vélos), des matrices de confusion ont été créées. *1* : mauvaise classification des chiens ; *2* : cause de non-détection des personnes ; *3* : vélo confondu avec une moto

Massif	Nom Cam	Classe	V+	F-	F+	%detection	%classification
Anlier	Etang	Personnes	172	18	3	90,5%	98,3%
Anlier	Etang	Chiens	0	1	0	0,0%	100,0%
Anlier	Etang	Sacs à dos	1	4	5	20,0%	16,7%
Anlier	Etang	Véhicules	0	0	1	100,0%	0,0%
Anlier	Etang	Vélo	2	6	0	25,0%	100,0%
Anlier	FJS	Personnes	120	10	0	92,3%	100,0%
Anlier	FJS	Chiens	0	0	0	100,0%	100,0%
Anlier	FJS	Sacs à dos	3	8	1	27,3%	75,0%
Anlier	FJS	Véhicules	14	0	0	100,0%	100,0%
Anlier	FJS	Vélo	0	7	0	0,0%	100,0%
Anlier	Pont	Personnes	171	20	5	89,5%	97,2%
Anlier	Pont	Chiens	0	0	0	100,0%	100,0%
Anlier	Pont	Sacs à dos	10	15	1	40,0%	90,9%
Anlier	Pont	Véhicules	0	0	18	100,0%	0,0%
Anlier	Pont	Vélo	14	10	0	58,3%	100,0%
Anlier	S_d_Tir	Personnes	62	41	6	60,2%	91,2%
Anlier	S_d_Tir	Chiens	0	0	2	100,0%	0,0%
Anlier	S_d_Tir	Sacs à dos	0	14	0	0,0%	100,0%
Anlier	S_d_Tir	Véhicules	2	0	2	100,0%	50,0%
Anlier	S_d_Tir	Vélo	4	0	0	100,0%	100,0%
Anlier	Vallee	Personnes	67	5	1	93,1%	98,5%
Anlier	Vallee	Chiens	3	0	0	100,0%	100,0%
Anlier	Vallee	Sacs à dos	11	3	1	78,6%	91,7%
Anlier	Vallee	Véhicules	0	0	2	100,0%	0,0%
Anlier	Vallee	Vélo	10	16	0	38,5%	100,0%
HF	B_Michel	Personnes	638	64	13	90,9%	98,0%
HF	B_Michel	Chiens	8	9	0	47,1%	100,0%
HF	B_Michel	Sacs à dos	94	36	3	72,3%	96,9%
HF	B_Michel	Véhicules	0	0	0	100,0%	100,0%
HF	B_Michel	Vélo	0	0	0	100,0%	100,0%
HF	Botrange	Personnes	240	7	1	97,2%	99,6%
HF	Botrange	Chiens	14	0	2	100,0%	87,5%
HF	Botrange	Sacs à dos	26	36	9	41,9%	74,3%
HF	Botrange	Véhicules	6	2	2	75,0%	75,0%
HF	Botrange	Vélo	33	9	0	78,6%	100,0%
HF	Bout	Personnes	101	2	2	98,1%	98,1%
HF	Bout	Chiens	4	0	0	100,0%	100,0%
HF	Bout	Sacs à dos	26	17	0	60,5%	100,0%
HF	Bout	Véhicules	9	3	0	75,0%	100,0%
HF	Bout	Vélo	1	1	0	50,0%	100,0%

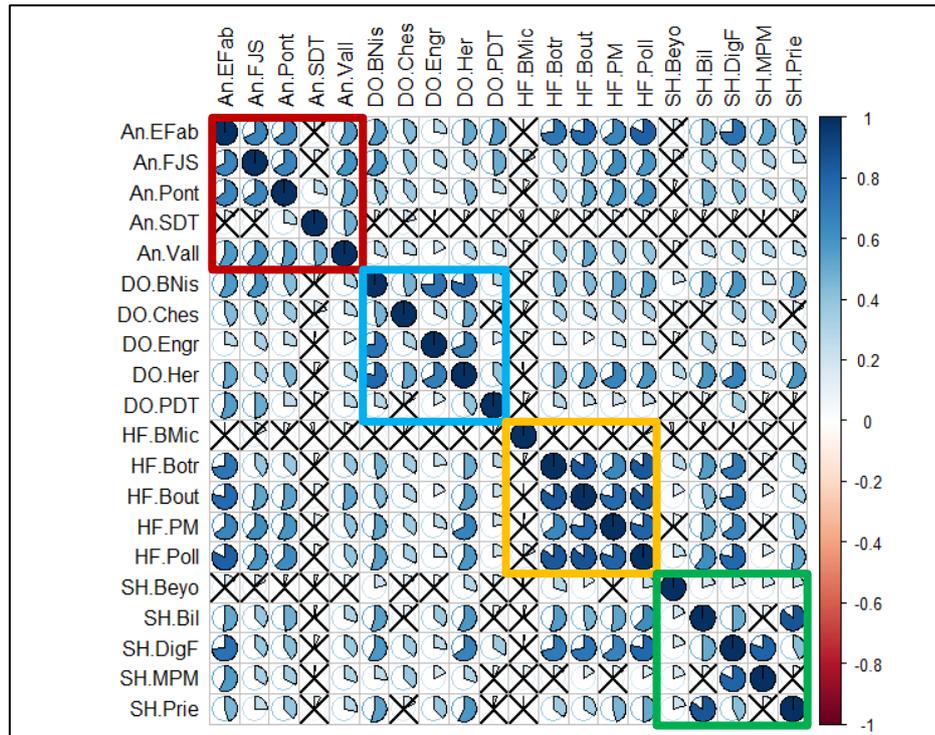
Annexes

HF	Polleur	Personnes	216	33	4	86,7%	98,2%
HF	Polleur	Chiens	3	0	1	100,0%	75,0%
HF	Polleur	Sacs à dos	35	15	1	70,0%	97,2%
HF	Polleur	Véhicules	3	0	3	100,0%	50,0%
HF	Polleur	Vélo	4	1	0	80,0%	100,0%
HF	Pont_Marie	Personnes	355	35	21	91,0%	94,4%
HF	Pont_Marie	Chiens	0	8	1	0,0%	0,0%
HF	Pont_Marie	Sacs à dos	71	15	38	82,6%	65,1%
HF	Pont_Marie	Véhicules	0	0	3	100,0%	0,0%
HF	Pont_Marie	Vélo	0	2	0	0,0%	100,0%
PNDO	Cheslé	Personnes	212	42	8	83,5%	96,4%
PNDO	Cheslé	Chiens	1	3	0	25,0%	100,0%
PNDO	Cheslé	Sacs à dos	24	20	5	54,5%	82,8%
PNDO	Cheslé	Véhicules	0	0	8	100,0%	0,0%
PNDO	Cheslé	Vélo	0	0	0	100,0%	100,0%
PNDO	Engreux	Personnes	264	15	7	94,6%	97,4%
PNDO	Engreux	Chiens	9	8	0	52,9%	100,0%
PNDO	Engreux	Sacs à dos	31	37	7	45,6%	81,6%
PNDO	Engreux	Véhicules	0	0	48	100,0%	0,0%
PNDO	Engreux	Vélo	2	2	0	50,0%	100,0%
PNDO	Herou	Personnes	288	21	4	93,2%	98,6%
PNDO	Herou	Chiens	0	1	0	0,0%	100,0%
PNDO	Herou	Sacs à dos	19	26	4	42,2%	82,6%
PNDO	Herou	Véhicules	0	0	3	100,0%	0,0%
PNDO	Herou	Vélo	0	0	1	100,0%	0,0%
PNDO	Nisramont	Personnes	253	18	13	93,4%	95,1%
PNDO	Nisramont	Chiens	4	0	0	100,0%	100,0%
PNDO	Nisramont	Sacs à dos	36	15	21	70,6%	63,2%
PNDO	Nisramont	Véhicules	0	0	3	100,0%	0,0%
PNDO	Nisramont	Vélo	7	4	0	63,6%	100,0%
PNDO	P_D_Tailles	Personnes	199	18	6	91,7%	97,1%
PNDO	P_D_Tailles	Chiens	8	20	0	28,6%	100,0%
PNDO	P_D_Tailles	Sacs à dos	14	4	2	77,8%	87,5%
PNDO	P_D_Tailles	Véhicules	8	5	0	61,5%	100,0%
PNDO	P_D_Tailles	Vélo	3	1	1	75,0%	75,0%
Saint Hubert	Beyoli	Personnes	143	49	12	74,5%	92,3%
Saint Hubert	Beyoli	Chiens	0	21	0	0,0%	100,0%
Saint Hubert	Beyoli	Sacs à dos	12	19	6	38,7%	66,7%
Saint Hubert	Beyoli	Véhicules	15	14	3	51,7%	83,3%
Saint Hubert	Beyoli	Vélo	1	0	1	100,0%	50,0%
Saint Hubert	Bilaude	Personnes	102	25	3	80,3%	97,1%
Saint Hubert	Bilaude	Chiens	0	0	0	100,0%	100,0%
Saint Hubert	Bilaude	Sacs à dos	7	9	0	43,8%	100,0%
Saint Hubert	Bilaude	Véhicules	37	9	55	80,4%	40,2%
Saint Hubert	Bilaude	Vélo	1	3	2	25,0%	33,3%
Saint Hubert	Diglette	Personnes	261	7	4	97,4%	98,5%
Saint Hubert	Diglette	Chiens	8	17	0	32,0%	100,0%
Saint Hubert	Diglette	Sacs à dos	29	22	9	56,9%	76,3%

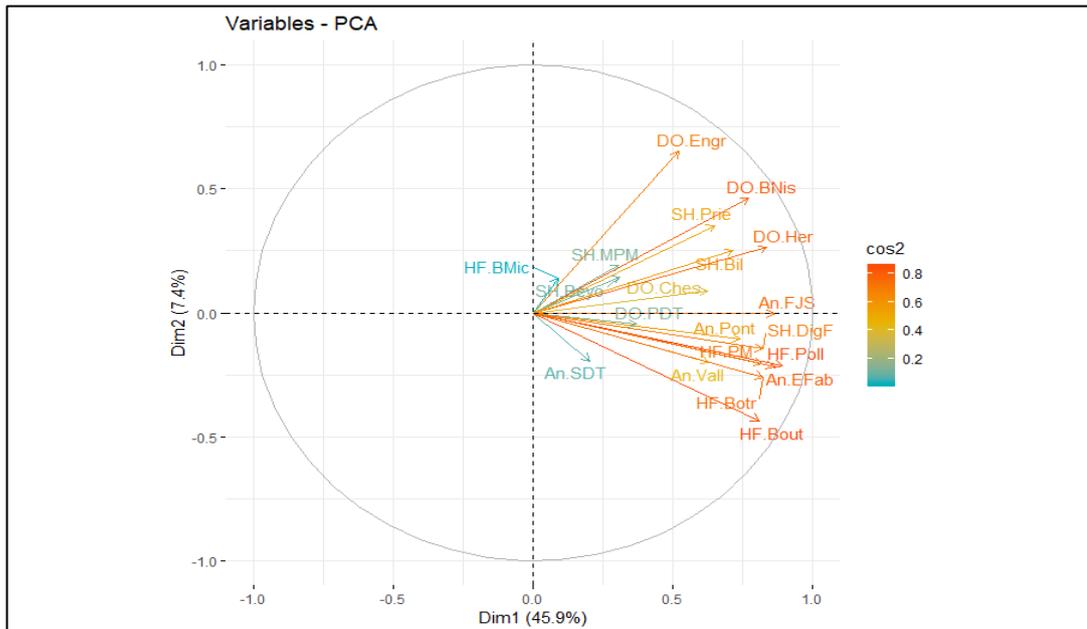
Annexes

Saint Hubert	Diglette	Véhicules	16	4	11	80,0%	59,3%
Saint Hubert	Diglette	Vélo	4	0	0	100,0%	100,0%
Saint Hubert	Masblette	Personnes	341	7	5	98,0%	98,6%
Saint Hubert	Masblette	Chiens	8	12	2	40,0%	80,0%
Saint Hubert	Masblette	Sacs à dos	40	13	0	75,5%	100,0%
Saint Hubert	Masblette	Véhicules	48	3	1	94,1%	98,0%
Saint Hubert	Masblette	Vélo	2	0	1	100,0%	66,7%
Saint Hubert	Priesse	Personnes	171	34	3	83,4%	98,3%
Saint Hubert	Priesse	Chiens	1	2	0	33,3%	100,0%
Saint Hubert	Priesse	Sacs à dos	5	15	1	25,0%	83,3%
Saint Hubert	Priesse	Véhicules	9	0	6	100,0%	60,0%
Saint Hubert	Priesse	Vélo	0	0	0	100,0%	100,0%

Annexe 7 : Résultats intermédiaires de l'étude des relations entre caméras pour le comptage du nombre de personnes



Graphique 20: Matrice de corrélation représentant l'évolution du nombre de personnes détectées et classifiées pour chaque caméra en fonction du temps. Les tests ont été réalisés grâce au coefficient de Pearson ($\alpha = 0,05$)



Graphique 21: Analyse en composante principale. Evolution du nombre de personnes capturées par chaque caméra en fonction du temps

Annexe 8: Résultats du test de la méthode de subsampling et extrapolation de Lupp et al. (2016) sur les données d'AGRETA

	Directory	Date	Nb photos	Nb personnes	Remarque
1	Barrage_Nisramont	15-08-18	224	331	
2	Barrage_Nisramont	16-08-18	322	547	
3	Barrage_Nisramont	26-07-18	120	186	
4	Barrage_Nisramont	05-10-18	66	96	
5	Baraque_Michel	27-10-18	0	0	carte sd pleine avant
6	Baraque_Michel	30-10-18	101	135	
7	Baraque_Michel	09-09-18	752	1590	
8	Baraque_Michel	18-08-18	1456	1584	faux déclenchements
9	Beyoli	11-12-18	2	0	
10	Beyoli	14-11-18	52	56	
11	Beyoli	15-11-18	10	14	
12	Beyoli	04-09-18	18	10	
13	Bilaude	06-10-18	358	704	
14	Bilaude	19-10-18	62	33	faux déclenchements
15	Bilaude	07-10-18	100	273	
16	Bilaude	06-12-18	66	4	voitures
17	Botrange	01-09-18	204	84	faux déclenchements
18	Botrange	16-08-18	236	338	
19	Botrange	28-08-18	1688	128	faux déclenchements
20	Botrange	31-12-18	184	508	
21	La Béole	19-09-18	428	126	faux déclenchements
22	La Béole	19-01-19	278	764	
23	La Béole	18-11-18	272	758	
24	La Béole	08-09-18	192	360	
37	Chesle	19-09-18	8	4	
38	Chesle	19-08-18	50	186	
39	Chesle	30-07-18	32	47	
40	Chesle	16-11-18	8	0	
41	Diglette_Fourneau	21-10-18	174	346	
42	Diglette_Fourneau	28-10-18	102	244	
43	Diglette_Fourneau	28-12-18	88	168	

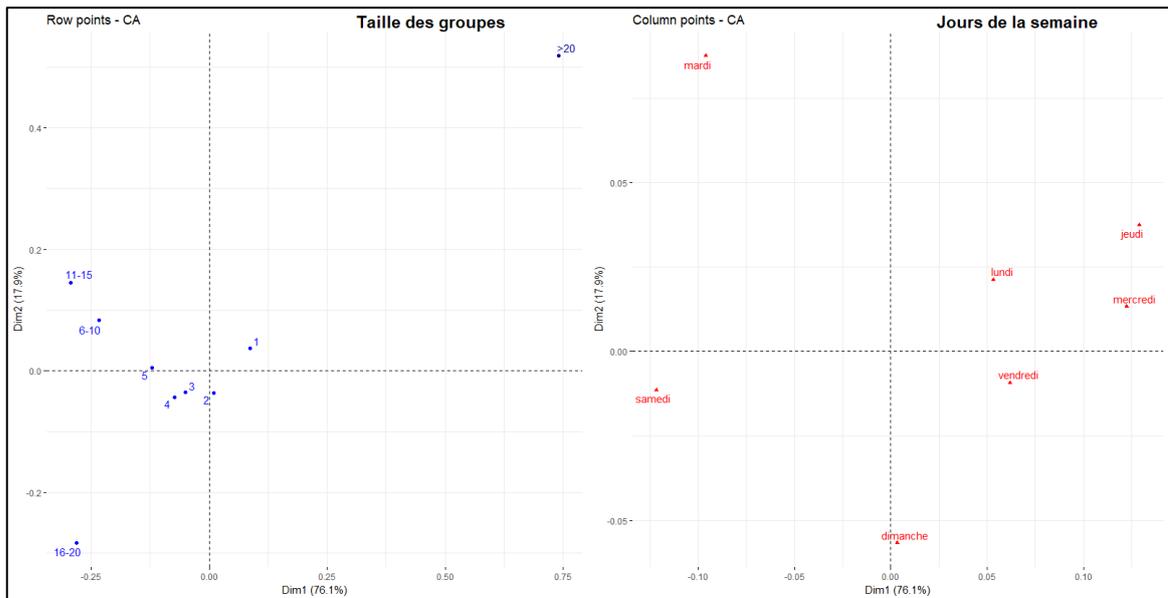
Annexes

44	Diglette_Fourneau	04-09-18	40	28	
45	Engreux	20-09-18	58	189	
46	Engreux	29-09-18	170	727	
47	Engreux	30-09-18	110	286	
48	Engreux	25-08-18	146	404	
49	Etang_Fabrique	24-08-18	112	186	
50	Etang_Fabrique	30-08-18	108	125	
51	Etang_Fabrique	17-11-18	144	199	
52	Etang_Fabrique	02-01-19	126	161	
53	Fagne_Jean_Simon	22-08-18	1238	2	faux declenchement
54	Fagne_Jean_Simon	26-08-18	838	64	faux declenchement
55	Fagne_Jean_Simon	23-08-18	1664	21	faux declenchement
56	Fagne_Jean_Simon	28-08-18	176	4	faux declenchement
57	Herou	18-07-18	216	625	
58	Herou	24-07-18	200	402	
59	Herou	01-11-18	228	693	
60	Herou	16-10-18	82	189	
61	Masblette_Pont_Mauricy	04-08-18	28	59	
62	Masblette_Pont_Mauricy	20-01-19	40	92	
63	Masblette_Pont_Mauricy	04-09-18	20	8	
64	Masblette_Pont_Mauricy	16-09-18	34	67	
65	Plateau_Des_Tailles	20-01-19	108	223	
66	Plateau_Des_Tailles	14-09-18	10	3	
67	Plateau_Des_Tailles	31-10-18	62	84	
68	Plateau_Des_Tailles	12-08-18	958	1522	
69	Polleur	09-11-18	84	148	
70	Polleur	21-10-18	592	1041	
71	Polleur	13-10-18	534	755	
72	Polleur	30-09-18	514	902	
73	Pont	03-11-18	22	27	
74	Pont	05-01-19	20	22	
75	Pont	24-08-18	42	79	
76	Pont	08-09-18	46	76	
77	Pont_Marie	14-10-18	255	2055	
78	Pont_Marie	25-08-18	464	1429	
79	Pont_Marie	03-10-18	0	0	bizarre

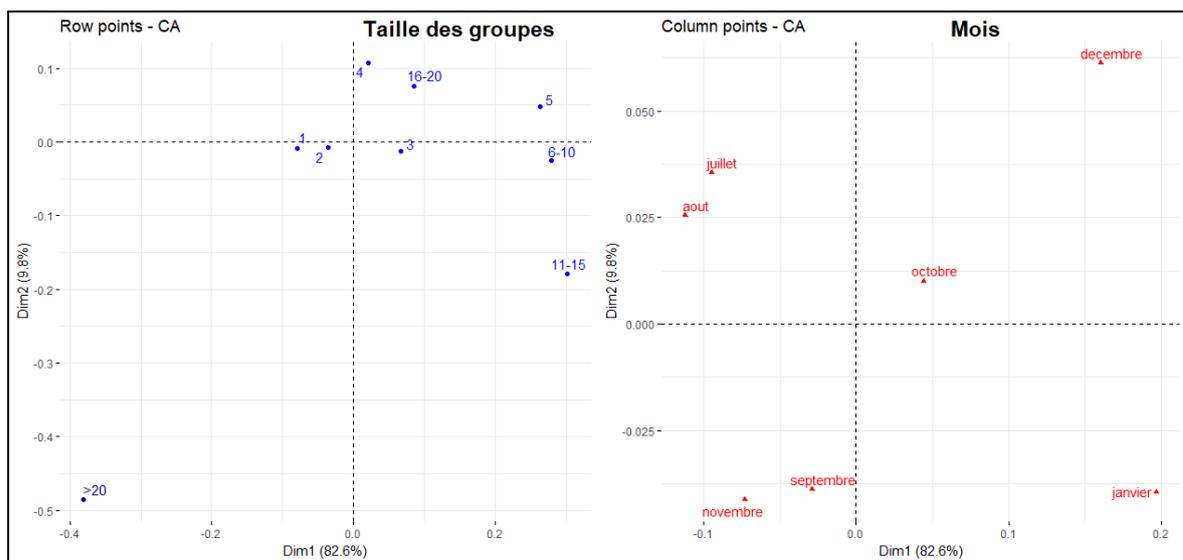
Annexes

80	Pont_Marie	02-08-18	103	73	
81	Priesse	03-10-18	96	124	
82	Priesse	21-09-18	260	427	
83	Priesse	16-01-19	64	66	
84	Priesse	29-09-18	538	1285	
85	Stand_De_Tir	20-09-18	362	8	faux déclenchements
86	Stand_De_Tir	26-08-18	126	96	faux déclenchements
87	Stand_De_Tir	21-09-18	8	0	faux déclenchements
88	Stand_De_Tir	22-09-18	10	11	
89	Vallee	12-01-19	18	14	
90	Vallee	25-09-18	62	0	faux déclenchements
91	Vallee	12-09-18	242	4	faux déclenchements
92	Vallee	19-09-18	278	20	faux déclenchements

Annexe 9 : Résultats complémentaires concernant l'étude de la taille des groupes



Graphique 29: Biplots représentant les correspondances entre les tailles des groupes (à gauche) et les jours de la semaine (à droite).



Graphique 30: Biplots représentant les correspondances entre les tailles des groupes (à gauche) et les mois (à droite).

Annexe 10 : Critique de la position de chaque caméra

Tableau 9: Critique de la position de chaque caméra. "1": condition remplie; "0": à améliorer. L'angle et le lieu reste le critère le plus important. Le lieu correspond à la simplicité du chemin (pas de croisements). La hauteur influence la qualité de la classification par le modèle

code	Angle et lieu	Risque d'obstruction	Risque de stagnation	Hauteur	Score
An-EFab	0	1	0	1	2
An-FJS	1	1	1	1	4
An-Pont	0	0	0	0	0
An-SDT	0	1	1	1	4
An-Vall	0	1	1	1	4
DO-BNis	0	1	1	1	3
DO-Ches	0	0	1	1	2
DO-Engr	0	1	1	0	2
DO-Her	0	0	0	1	1
DO-PDT	1	1	1	1	4
HF-BMic	0	1	1	1	3
HF-BoTr	1	0	0	1	2
HF-Bout	0	1	0	1	2
HF-PM	0	1	1	1	3
HF-Poll	1	1	0	1	3
SH-Beyo	1	0	1	0	2
SH-Bil	1	1	1	1	4
SH-DigF	1	1	1	0	3
SH-MPM	0	1	1	1	3
SH-Prie	0	0	1	1	2
SOMME	7	14	14	16	

Annexe 11 : Comparaison entre les données caméras et Proximus

Tableau 10: Coefficient de Pearson ($\alpha=0,05$) comparant l'évolution au cours du temps entre le nombre de personnes comptées par l'indice et le nombre de personnes comptées par Proximus. En jaune, les coefficients de Pearson supérieur à 0,5.

Origin	Behavior	UL_Anlier	UL_DeuxOurthes	UL_DeuxOurthesCoeur	UL_HautesFagnesBords	UL_HautesFagnesCoeur	UL_SH_Feuillus	UL_SH_Ouvert	UL_Taillecoeur
Resident	All	-0,133	0,488	-0,257	0,428	0,015	-0,033	0,061	0,031
Resident	Longstops	-0,135	0,48	-0,277	0,45	-0,002	-0,033	0,061	0,067
Resident	Shortstops	0,069	0,037	0,098	-0,008	0,039			-0,068
Local	All	-0,247	0,421	-0,282	0,403	-0,072	0,029	0,015	0,253
Local	Longstops	-0,276	0,436	-0,33	0,304	-0,201	-0,071	-0,081	0,192
Local	Shortstops	-0,169	0,364	-0,217	0,391	0,081	0,102	0,114	0,27
Regional	All	-0,177	0,537	0,022			0,022	-0,011	0,235
Regional	Longstops	0,029	0,539	0,009			0,031	0,03	0,182
Regional	Shortstops	-0,288	0,468	0,027			0,012	-0,028	0,223
National	All	0,643	0,742	0,296	0,807	0,414	0,373	0,619	0,206
National	Longstops	0,632	0,728	0,25	0,735	0,388	0,371	0,435	0,165
National	Shortstops	0,545	0,753	0,348	0,842	0,387	0,348	0,64	0,223
International	All	0,309	0,621	0,288	0,461	0,139	-0,025		-0,003
International	Longstops	0,225	0,558	0,298	0,15	-0,004			-0,008
International	Shortstops	0,312	0,64	0,233	0,677	0,253	-0,025		0,002
All	All	0,148	0,708	0,018	0,783	0,231	0,249	0,415	0,246
All	Longstops	0,094	-0,064	0,454	-0,185	0,654	0,579	0,456	0,687
All	Shortstops	0,09	0,719	0,166	0,801	0,325	0,256	0,455	0,263

Cette Annexe n'est pas décrite dans le travail et est uniquement destinée à l'équipe d'AGRETA. Pour de plus amples d'informations, me contacter à gguidosse@gmail.com.