

État de l'art - Méthodes de suivi innovantes de la fréquentation

Mai 2023

Les besoins du PNE sur la connaissance de la fréquentation	2
Les technologies	3
Technologies IA	3
DeepFaune	3
Camtrap	4
Timelapse	4
Pièges IA	5
WildCount	5
SenseCAP	5
Hack the planet - IA camera trap	6
Solutions de traitement des données de piège photo	7
GeoCam	7
Wildlife Insights	7
Conservation AI	8
Trapper	8
Agouti	8
WildTrax	8
Autres	9
Affluences	9
Flux Vision	9
Outdoorvision	9
OpenDataCam	10
Camelot	10
Liste non exhaustive	11
Les acteurs	12
Techniques & Scientifiques	12
Point juridique	14
Conclusion	15

Aurélien Coste

Les besoins du PNE sur la connaissance de la fréquentation

Le suivi de la fréquentation dans le Parc national des Écrins est une thématique importante pour la gestion de celui-ci. Depuis les 50 ans d'existence du Parc, des enquêtes ont été menées et des outils ont été mis en place pour produire des données de suivi. Cependant, les dispositifs actuels ne suffisent pas pour avoir un suivi efficace et à long terme de la fréquentation. De plus, avec l'arrivée de nouvelles technologies, il convient de réaliser un nouvel état de l'art concernant ces nouveaux procédés innovants de suivi de la fréquentation.

L'analyse des besoins du Parc national des Écrins a été réalisée par Maëlle Bouchez l'an dernier lors de son stage. En voici une synthèse. Les enjeux par rapport à la fréquentation sont différents selon les secteurs et selon les services au siège du Parc. Il est important de qualifier la fréquentation, car le constat partagé est que le problème de ces dernières années n'est pas le nombre de visiteurs, mais leurs pratiques et leurs comportements. La fréquentation hivernale a plus d'impact, mais est bien moins canalisée, cela reste important d'avoir un suivi de la fréquentation même en hiver. Le dispositif actuel ne fournit que des informations, sans analyse, il y a un manque de connaissances sur comment et pourquoi les visiteurs viennent. La chaîne d'organisation pour accueillir les visiteurs doit être plus cohérente. Les données sont actuellement récoltées à la main ce qui augmente considérablement le temps de traitement. Il faudrait pouvoir communiquer à distance les informations des capteurs et pièges photos. Il faut donc consolider le système actuel de suivi par l'utilisation efficace des outils déjà en place et par l'intégration de nouveaux outils.

On a alors besoin d'un ou des outil(s) qui permettrait à son utilisateur d'annoter les médias, de les archiver, d'analyser les données d'annotation, d'exporter ces données, de gérer les dispositifs et les sites où sont placés les dispositifs.

Pour assurer le suivi de la fréquentation à long terme, il convient de s'appuyer sur une structure solide qui comprend toute la chaîne de valeur des données. Voici un schéma réalisé par Maëlle qui image parfaitement la chaîne des données.

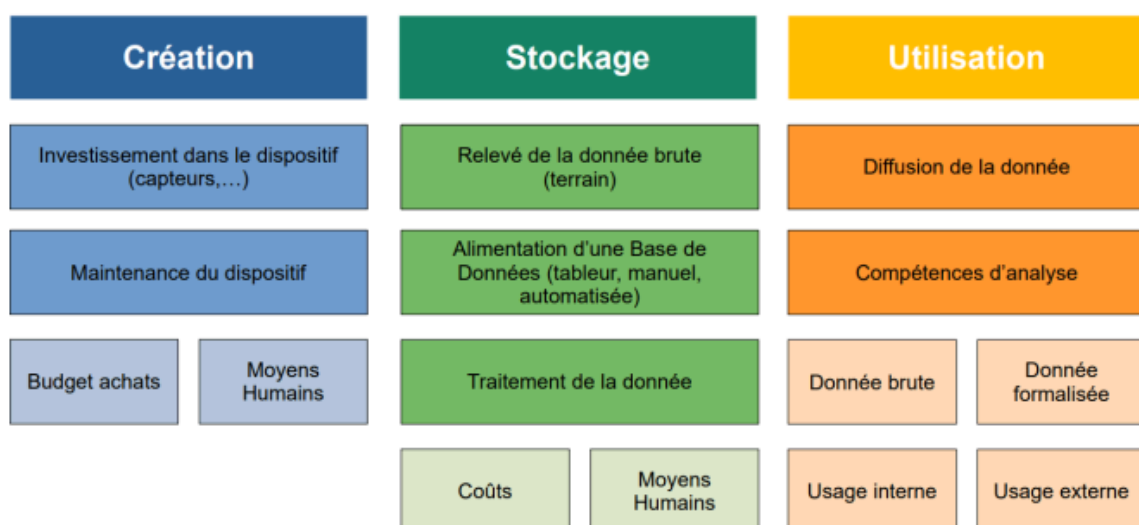


Figure 1 - Schéma de la chaîne des données

Les technologies

Au travers de cet état de l'art, plusieurs catégories d'outils vont être analysées. Il en va de définir ces catégories pour mieux visualiser l'ensemble de l'existant.

La première catégorie est **Technologie IA**, il s'agit des outils qui proposent le traitement des médias (images et vidéos) via une IA (qui est à 99.9 % du temps un ou des réseaux de neurones). Seuls, les outils de cette catégorie ne permettent pas de répondre à tous nos besoins. Néanmoins, ils sont importants à comprendre et à étudier car certains de ces outils sont utilisés dans des solutions d'autres catégories.

La catégorie suivante est **Piège IA**, il s'agit de l'ensemble des outils qui proposent une solution de piège photo avec IA embarquée. Ces outils vont prendre des photos sur place, analyser ces photos via l'IA et transmettre les données par un réseau LoRaWAN. Cette catégorie répond à plus de besoin que la catégorie précédente car elle contient la prise de photo et donc l'acquisition des données à traiter.

La troisième catégorie est celle des **solutions de traitement des données de piège photo**, cela comprend l'ensemble des outils qui proposent des applications où l'on peut importer des médias, les stocker, exporter des données et parfois utiliser l'IA pour reconnaître les taxons dans nos médias. Il en existe énormément, chacune à quelques spécificités mais sont assez ressemblantes dans le principe et dans les fonctionnalités.

Technologies IA

DeepFaune

DeepFaune est un projet collaboratif dirigé par le CNRS avec la collaboration du LBBE (Laboratoire de Biométrie et Biologie Évolutive) et du CEFE (Centre d'Écologie Fonctionnelle et Évolutive) qui utilise de l'IA pour avoir un outil de reconnaissance automatique de la faune en Europe et particulièrement en France. L'outil mis à disposition est un logiciel qui permet à n'importe quel utilisateur de classer des images ou des vidéos de pièges photo en local sur son ordinateur. DeepFaune utilise deux étapes pour la détection. La première étape consiste à identifier un élément dans l'image. Pour cela, MegaDetectorV5 est utilisé, il permet de trier les images en quatre catégories qui sont animaux, humains, véhicules ou rien. Cet outil coupe les images d'animaux ce qui les rend plus simples à classer pour l'étape suivante. Les scientifiques ont décidé de ne garder que les images avec un score de confiance supérieur à 0,80. Comme l'outil est open source c'est une valeur que chacun peut changer selon son besoin. À savoir que la version logiciel utilise YoloV5 car il est plus rapide que son homologue MegaDetectorV5.

La seconde étape de la détection consiste à utiliser un modèle CNN qui est EfficientNetV2 pour classer les images d'animaux parmi 22 espèces. Le modèle calcul un score de prédiction pour chaque classe et retourne l'espèce avec le score le plus haut.

La version actuelle de DeepFaune a pu être entraînée sur 418 726 images d'animaux. 90 % de ces images ont été utilisées pour le training set et 10 % pour le test set.

DeepFaune est donc un outil très puissant mais qui ne répond pas à tous les besoins. On peut citer par exemple que DeepFaune ne peut pas compter le nombre d'individus sur une image. Il semblerait après contact avec l'équipe de DeepFaune qu'il existe une version en interne capable de compter le nombre d'individus d'une seule espèce dans une image. Il ne reconnaît pas les vélos qu'il classe en tant qu'humain. Néanmoins, l'outil est capable de

traiter des images qui auraient été floutées au préalable par un autre outil, ce qui permet une anonymisation des individus.

Camtrap

Camtrap est un outil qui permet le traitement des médias en local ou sur un serveur NAS. Il s'agit d'une application interactive avec des modules de traitement automatique des médias. Les modules ont pour rôle d'analyser automatiquement les médias et d'en extraire des caractéristiques exploitées ensuite par l'application. L'application est une web app qui permet de consulter les médias annotés et/ou filtrés à partir des modules de traitement, de saisir des observations et de visualiser les données enregistrées par l'outil. Il existe deux modules de traitement, ExifTool qui extrait les métadonnées (date précise, durée) et MegaDetector comme pour DeepFaune qui trie les images en quatre catégories. Il existe un mode observation qui permet de visionner et de traiter des images uniquement des images qui n'ont pas encore été contrôlées par des humains ou alors des images qui ont déjà été traitées par des humains. L'outil est plus orienté gestion du workflow dans son ensemble qu'uniquement classification IA des médias comme le font certains outils.

L'outil est développé par Vincent Prunet du Parc national du Mercantour. Il prévoit d'intégrer un module de classification automatique dans une future v2. Cela signifie que l'outil n'est pas encore capable de reconnaître l'espèce de l'animal présent sur l'image. Après contact avec Vincent Prunet, j'ai eu la confirmation que le développement de l'outil était en pause, mais qu'il compte reprendre le travail dessus cet été.

Timelapse

Timelapse est un logiciel gratuit développé par l'université de Calgary au Canada. L'outil affiche les médias, extrait les métadonnées voulues par l'utilisateur et catégorise les images inutilisables. Il permet d'importer, d'afficher et d'interroger les données de reconnaissance d'images produites par MegaDetector. C'est à l'utilisateur ensuite, de parcourir les images et d'indiquer de quelle espèce est l'animal sur l'image ainsi que le nombre s'il y a plusieurs animaux. Dans le cas où il y aurait plusieurs espèces différentes sur la même image, l'outil permet de dupliquer l'image pour la compter une fois avec l'espèce A et une seconde fois avec l'espèce B. Il peut extraire les données dans une table en CSV. Il y a de la documentation et une vidéo de démonstration pour l'utilisation de l'outil. Le défaut de l'outil c'est qu'il n'utilise pas l'IA excepté le fait qu'il est capable d'analyser les images provenant de MegaDetector. On peut alors trier dans un premier temps nos images avec MegaDetector puis les importer dans Timelapse. Un autre défaut est qu'il ne semble plus mis à jour et que l'outil ne fonctionne que sur Windows.

Pièges IA

WildCount

WildCount est un projet de recherche mené par Georges Quenot et Didier Donsez. Le projet vise à produire une caméra de chasse bon marché pour reconnaître et compter la présence d'humains (anonymes) et d'animaux dans les zones sauvages et protégées.

Ce produit contient donc un ensemble de composants (Sony SPresense, Grove PIR sensor, Microchip RN2483, 2x 18650 batteries rechargeables, 2x 18650 supports de batterie, Kerlink iStation, 5 dBi antenne) qui permet de prendre une photo, reconnaître l'espèce sur la photo, transmettre via LoRaWAN l'information à l'utilisateur via Node-RED, influxdb puis Grafana qui fait l'interface avec l'utilisateur.

Pour reconnaître l'espèce, WildCount utilise un modèle de réseau de neurones appelé MobileNet. Ce modèle est particulièrement adapté pour être utilisé dans des systèmes embarqués car un microcontrôleur suffit pour l'exécuter. C'est également un avantage d'un point de vue consommation d'énergie puisqu'il consomme moins que DeepFaune par exemple qui est pour le moment plus efficace mais qui n'est pas fait pour tourner sur des systèmes embarqués.

Le projet avance lorsqu'il y a des étudiants ou stagiaires mobilisés. La dernière équipe mobilisée remonte à quelques mois. Il n'y a qu'un seul prototype qui a été testé une fois dans le Valbonnais de novembre à avril. Le prototype a été installé sur un sentier pas trop fréquenté, car le prototype est assez gros. Il faut également un réseau LoRa pour tester le prototype, c'est pourquoi il a été installé proche du bureau pour que la communication puisse se faire avec une seule antenne.

SenseCAP

La gamme de produits SenseCAP est produite par Seeed Studio, entreprise spécialisée dans la recherche de technologie IoT (Internet of Things). Il s'agit de produits industrialisés, principalement des capteurs de température, d'air et autres ainsi qu'un modèle combinant une caméra et l'IA pour différentes utilisations. L'outil possède ses propres modèles d'IA grâce à TinyML que l'utilisateur peut sélectionner via une interface selon ses besoins. Il y a par exemple la détection d'objets, le comptage d'objets (notamment de personnes, ce qui est intéressant pour le suivi de la fréquentation) ou la classification d'image. On peut également mettre notre propre modèle IA dans l'outil. On peut choisir le score de confiance à partir duquel les informations seront envoyées, selon la FAQ, le modèle de détection d'humain est à 70% d'identifications correctes. Dans la documentation, il y a des informations concernant les températures que l'appareil peut supporter (-40°C à 85°C) ainsi que sur la durée de fonctionnement (jusqu'à 10 ans) et la portée maximale de transmission (10 km). L'appareil est alimenté par une unique pile LiSOC12 ER34615 (plus de détails dans la doc), selon leurs tests, l'autonomie serait de 9 mois. La sécurité des données est gérée par le "Edge Computing". Pour configurer le piège, il y a une application à installer, il faut ensuite connecter le téléphone à l'appareil en Bluetooth. Selon le réseau LoRaWAN utilisé, il faut un compte sur la plateforme et adapter la fréquence de transmission, il en existe 7 dans l'appareil, il y a un tableau récapitulatif dans la documentation. L'outil utilise OTAA pour rejoindre le réseau LoRa, mais il est compatible avec ABP. Plusieurs modes de transmissions sont disponibles avec plus ou moins de paquets d'acquittements. Pour la visualisation des données, c'est par la plateforme SenseCAP Portal qui utilise Azure. Sur

cette plateforme, on peut visualiser nos appareils ainsi que les statistiques/données récupérées par nos différents pièges. Il y a également une API avec trois moyens différents pour récupérer les données (HTTP, MQTT et Websocket).

L'inconvénient de ce produit est qu'il ne dispose pas de beaucoup d'avis de la part d'utilisateurs qui l'auraient testé. Le défaut majeur de l'outil est son mode de fonctionnement. L'appareil se réveille, collecte les valeurs et les transmet via LoRaWAN, par défaut l'intervalle de réveil est de 60 minutes, mais peut être modifié de 1 à 1 440 minutes. Je pense que l'entreprise a d'abord produit ses capteurs de température, d'air etc puis s'est penché sur la détection d'image avec l'IA, mais n'a pas réfléchi à la problématique de la prise de photo par un détecteur de mouvements. La prise de photo à intervalles réguliers fonctionne parfaitement pour la prise de mesures, mais n'est pas efficace pour l'analyse d'image. Prendre une photo toutes les minutes (ce qui correspond à l'intervalle minimal) n'est pas suffisant.

J'ai envoyé un mail au support pour leur poser la question de la détection de mouvement pour la prise d'image, en effet, ils m'ont confirmé ce que j'ai pu lire dans la doc.

Enfin, malgré une documentation en anglais correcte, l'utilisation de l'outil nécessite des compétences en informatique (notamment en LoRa).

Il existe un produit similaire à SenseCAP, il s'agit du Milesight VS121, créé par une entreprise spécialisée dans la commercialisation de produits LoRa. En revanche, cet outil est adapté pour une utilisation indoor, cela ne correspond donc pas à nos besoins. De plus, il n'y a encore aucun retour sur ce produit, il semble être très récent.

Hack the planet - IA camera trap

Hack the planet est une organisation à but non lucratif composée de deux ingénieurs néerlandais. Ils ont développé une caméra qui utilise du machine learning pour analyser les photos en temps réel et alerter les rangers par WhatsApp si quelque chose est détecté. Il s'agit d'une caméra standard à laquelle est relié un mini-ordinateur. Ce mini-ordinateur est capable de récupérer les images de la caméra puis d'utiliser l'IA pour classer ce qu'il y a sur l'image. Le système est équipé d'une liaison montante par satellite et peut fonctionner partout dans le monde sans dépendre d'un réseau mobile ou LoRa. L'outil est principalement utilisé pour lutter contre le braconnage et les conflits avec les éléphants qui s'approchent trop près des villages au Gabon.

Il semblerait qu'il existe très peu de solutions de ce style, néanmoins certains sites web proposent des DIY (Do It Yourself) ou des tutoriels pour créer son propre piège IA. Cela semble donc réalisable d'avoir son propre piège photo avec IA à la maison mais pour un cadre plus sérieux de suivi de la fréquentation comme souhaite faire le Parc national des Écrins ce n'est pas envisageable d'utiliser ce genre de piège photo homemade.

Cependant, on peut remarquer que les entreprises spécialisées dans les produits LoRa s'intéressent de plus en plus aux pièges photo avec IA. Leurs produits ne répondent pas encore à nos besoins, mais c'est quelque chose qu'il faudra vérifier dans le futur lorsque leurs produits auront évolué.

Solutions de traitement des données de piège photo

Un facteur important dans le choix de la plateforme est de savoir si on veut que nos données puissent être rendues publiques. Certaines plateformes ont été développées pour maintenir des workflows privés tandis que d'autres sont orientées vers des dépôts publics. Les plateformes diffèrent également par leurs facilités d'utilisation, la performance du modèle ainsi que des fonctionnalités intégrées.

GeoCam

GeoCam anciennement FASTcam est une application web open source créée par Natural Solutions qui permet de gérer le stockage et d'annoter les médias importés. L'outil n'est pas une solution cloud, c'est-à-dire que chaque organisation aura son application GeoCam sur son serveur qu'il peut modéliser à souhait et les données restent dans l'organisation. Pour l'annotation des médias, on pourra compter le nombre d'espèces et/ou d'animaux présents sur une même image ainsi qu'annoter manuellement des informations supplémentaires tel que le sexe ou le comportement. Il permet également de créer des projets où l'on peut renseigner des pièges et des sites. Cet outil est très récent, il a vu le jour pendant le deuxième trimestre de 2022. Sa conception a été très réfléchie, il existe un backlog de GeoCam avec toutes les fonctionnalités que l'outil souhaite avoir. C'est très prometteur car ce backlog est vraiment complet et prendra notamment en compte le format des données. Il utiliserait les normes standard telles que TAXREF ainsi que la norme CTMS pour les pièges photo. GeoCam pourrait extraire des données d'occurrences transmissibles à GeoNature car un travail de compatibilité serait fait de manière à intégrer dans la synthèse de GeoNature les données produites dans GeoCam. Cependant l'outil n'est pas encore fonctionnel et son développement semble avoir été ralenti par manque d'investissement. Le dernier commit remonte à janvier 2022 mais le développement ne semble pas avoir été arrêté car des issues sont ouvertes par les développeurs sur le GitHub.

Wildlife Insights

Wildlife Insights est la référence en matière de plateforme web dans le domaine du traitement automatisé des images. Il est possible d'importer et de télécharger ses images sur la plateforme afin que les espèces soient automatiquement identifiées grâce à l'IA. La plateforme propose également un très bon outil d'examen et d'annotation des images pour corriger et/ou vérifier si l'IA a correctement classifié l'espèce. Il s'agit de leur propre IA qui a été entraînée en utilisant la technologie TensorFlow de Google. L'IA est très performante et permet de reconnaître plus de 1000 espèces différentes. Il est important de prendre en compte que l'outil est utilisé au niveau mondial et donc qu'il se peut qu'il ait plus de mal à reconnaître la faune française et même européenne. L'outil n'est pas disponible en français et n'est pas open source. À savoir que le partage des données est obligatoire après une période de 48 mois.

Conservation AI

Conservation AI est une plateforme web développée à l'Université John Moores de Liverpool. L'outil dispose de différents modèles IA pour identifier les humains, les voitures et les espèces (principalement celles du Royaume-Uni, d'Afrique du Sud et d'Amérique du Nord). L'utilisateur upload ses images sur la plateforme, il choisit le modèle de reconnaissance, ensuite les images vont être classées par le modèle puis les résultats seront disponibles sur la plateforme. L'outil ne gère pas le stockage des médias, les images publiées sont disponibles pour tous et il n'y a pas de moyen de vérifier et corriger ce que l'IA a prédit.

Trapper

Trapper est une application web développée par l'OSCF (Open Science Conservation Fund) basé en Pologne, il permet le traitement de médias (images et vidéos). L'outil propose le traitement automatisé des médias par l'IA qui a été entraîné par eux-mêmes. Il existe énormément de documentation et est open source mais l'outil n'est pas disponible en français.

Agouti

Agouti est une application web à l'initiative de l'Université de Wageningen (Pays-Bas) et de l'Institut de recherche sur la nature et la forêt (INBO). L'outil propose d'importer des médias, de les annoter manuellement, de les analyser en ligne via leur outil et enfin d'exporter les fichiers de données dans des formats standardisés. L'outil est disponible en français et dans de nombreuses autres langues européennes. Il n'est pas open source et une contribution est demandée pour les organisations à but non lucratif qui souhaitent utiliser Agouti.

WildTrax

WildTrax est une plateforme en ligne développée par l'ABMI (Alberta Biodiversity Monitoring Institute) qui permet de stocker, gérer et analyser les données que l'on peut importer. L'outil permet de gérer les images, les vidéos et les audios. MegaDetector est également pris en charge pour filtrer les images et l'outil met à disposition un package R, wildRtrax pour analyser et normaliser les données, qui est gratuit à condition de disposer d'une licence MIT © 2020. WildTrax n'utilise pas d'IA pour accélérer la phase d'analyse des images et n'est pas un outil open source.

Autres

Cette section présente les outils qui sont assez originaux et qui ne rentrent pas dans l'une des trois précédentes catégories.

Affluences

Affluences est une entreprise spécialisée dans la gestion de l'affluence depuis 2014. Ils fournissent plusieurs technologies pour gérer l'affluence selon le cas d'usage, transports, bibliothèques universitaires, sites naturels et la liste n'est pas exhaustive. La solution pour les sites naturels est expliquée sans rentrer dans les détails techniques. Une organisation peut se doter du système de comptage d'affluences pour avoir une vision précise de tous les flux de visiteurs. L'outil est capable de différencier adultes et enfants, voitures, vélos et animaux. Ils offrent également un outil de système de file d'attente virtuelle et de réservation par créneaux. La solution est équipée d'un algorithme d'analyse prédictive d'affluence demi-heure par demi-heure, à la journée ou à la semaine. Il y a une Web API pour intégrer les prévisions sur nos propres outils ainsi qu'un back-office pour exploiter les indicateurs d'affluence de nos sites.

Ils mettent à disposition deux types de compteurs. Le premier est dit compteur IA mais il nécessite une alimentation électrique continue. Le second est dit compteur thermique petit et discret. Ils sont étanches et résistent aux fortes températures, ils n'ont pas besoin d'une alimentation électrique continue.

L'utilisation de Affluences au mont Ventoux permet de connaître en direct le taux d'occupation des parkings via des compteurs infrarouges et vidéos 3D. En revanche, il n'est pas question de réserver sa place comme aux calanques à Marseille.

Flux Vision

Flux Vision est un outil développé par Orange business qui permet d'analyser des données de fréquentation, de provenance et de profils sociodémographiques. Flux Vision permet de mesurer et d'analyser la fréquentation et les flux de déplacements sur une zone géographique donnée. Les données sont le volume et la durée de fréquentation du lieu, les profils socio-démographiques, le lieu de provenance ainsi que l'emplacement de l'hébergement et le nombre de nuitées. L'outil s'appuie sur des données GPS en plus des données issues du réseau mobile Orange. La combinaison des deux sources offre une précision de l'ordre de 50 m. Les données et statistiques sont anonymes et respectent les recommandations CNIL et RGPD. L'outil semble intéressant mais fonctionne-t-il tout aussi bien en cœur de parc là où le réseau mobile est moins développé.

Outdoorvision

Créé en 2019, Outdoorvision est un projet piloté par le Pôle Ressource National Sports de Nature (PRNSN). Cette plateforme est destinée aux gestionnaires des territoires. Elle cartographie les flux des pratiquants de marche à pied, trail, VTT et ski en récupérant leurs traces GPS. La récupération s'effectue via des applications : certaines sont récoltées directement (Visorando, VisuGPX, trail CONNECT, IGNrando, OpenRUNNER), d'autres par une connexion du compte de l'utilisateur avec Outdoorvision via l'application (Garmin, Suunto, Polar) ou alors par un dépôt direct de l'utilisateur sur la plateforme Outdoorvision. L'objectif avec cet outil est d'avoir une vue macro des pratiques sportives, de façon continue

dans le temps. C'est un outil d'aide à la décision locale : on peut constater l'accord ou le désaccord de l'activité des visiteurs avec l'offre du territoire et adapter cette dernière.

OpenDataCam

OpenDataCam est un outil open source qui se connecte à une caméra. Il fait tourner en temps réel un algorithme qui compte et suit les objets en mouvement. L'algorithme utilisé est Yolo dont j'ai déjà parlé dans la présentation de DeepFaune. OpenDataCam intègre une interface qui permet à l'utilisateur d'avoir le contrôle sur les données, il n'y a pas de connexion cloud. Pour l'installation, il y a une bonne documentation et l'outil fonctionne via une interface web locale qui permet de visualiser ce que voit la caméra en temps réel, de mettre en place des outils pour compter et de visualiser les trajectoires de tous les objets.

OpenDataCam peut lire des fichiers locaux en format mp4 ou alors via une caméra USB et pour les caméras à distance, il faut que le flux vidéo soit en format MJPEG et diffusé via le protocole RSTP.

L'outil reconnaît tout ce que peut reconnaître Yolo, c'est-à-dire une petite centaine d'objets. Il y a également un mode pour voir et exporter les données comptées ainsi que les trajectoires. Il existe une API pour récupérer les données pour les utiliser dans notre application web par exemple. L'outil ne sauvegarde aucune image, tout est en temps réel et les données sauvegardées sont uniquement des chiffres ainsi que des trajectoires en format JSON ou CSV. L'outil peut être adapté selon les besoins de chaque organisation en modifiant son code. L'avantage majeur de cet outil est qu'il peut compter le nombre d'objets vus ainsi que connaître le sens de passage. En revanche, l'outil ne supporte pas plusieurs inputs, il faut déployer plusieurs instances de OpenDataCam pour gérer plusieurs caméras, de plus l'outil nécessite un humain avec des compétences techniques en informatique pour gérer l'outil.

Camelot

Camelot est un outil développé par deux développeurs et est encore mis à jour. Il s'agit d'un logiciel installable facilement en local. On peut créer une organisation dans laquelle on peut gérer des enquêtes, des sites, des caméras et exporter des rapports. Une enquête est composée de caméras et permet d'importer des médias. Une fois les médias importés on peut visualiser et annoter les images et vidéos mais à la main uniquement, il n'y a pas d'IA. Cependant on peut utiliser leur service en ligne où après s'être enregistré et avoir upload des images, Camelot lancera automatiquement MegaDetector sur ces images. À noter que l'outil est disponible uniquement en anglais mais il y a une bonne documentation en ligne (en anglais également).

Liste non exhaustive

Cette section présente très brièvement les autres outils que j'ai survolés au fil des recherches et qui sont moins intéressants du fait qu'il ne soit plus mis à jour, pas gratuit ou qu'il existe mieux.

Aardwolf : n'est plus à jour, pas très professionnel car un seul dev et aucune structure derrière. C'est un projet homemade.

CameraBase : n'est plus à jour depuis 2015.

CamtrapR : est un package R qui permet la gestion et l'extraction des pièges photo. Il permet de créer un fichier R pour manipuler les données liées aux pièges photo, c'est se compliquer la tâche d'écrire du code R alors qu'il existe des solutions plus simples.

CPW Warehouse : est une application personnalisée de Microsoft Access conçue pour faciliter l'archivage, l'identification, la synthèse et l'analyse des données photographiques. Il faut une version d'Access pour utiliser ce produit qui n'est plus mis à jour depuis 2019.

DigiKam : est un logiciel avancé open source de modification d'images avec des outils d'import, de gestion, d'édition et de partage d'images. Ce n'est donc pas une solution complète.

eMammal : est une application web de gestion des données et d'archivage pour les projets de piège photo par le SCBI (Smithsonian's national zoo & conservation biology institute). L'outil est plutôt orienté stockage et partage qu'analyse des images.

iNaturalist : est une plateforme web communautaire pour le partage de données de la faune ainsi qu'un réseau social où les utilisateurs peuvent communiquer entre eux. Tout le monde peut participer et identifier les images déposées par les autres. On est donc dépendant de la participation des utilisateurs pour la classification.

Ocapi : est un logiciel de développement d'IA pour l'analyse de données issues de piège photo développé par TerrOïko. Ce n'est donc pas une solution complète mais on peut peut-être utiliser leur outil d'entraînement d'IA en cas de besoin.

Snoopy : est une plateforme d'étude qui date de 2015. Elle n'est plus mise à jour et l'interface est un peu vieille.

WildID : est un outil payant qui permet de traiter les données d'images provenant de pièges photo en trois étapes simples. Upload, identification automatique via IA et vérification. De plus l'IA est entraînée avec des images d'Afrique du sud donc des espèces différentes de celles qu'on peut trouver en France ce qui signifie que la classification par l'IA sera moins efficace.

Wild.ID : est une application logicielle développée au San Diego Supercomputer Center. L'outil permet de gérer et traiter des médias de pièges photo. L'outil n'est plus mis à jour depuis 2019 et ne contient pas d'IA.

ViXen : est un outil open source qui n'est plus mis à jour depuis 2019 et qui permet d'annoter manuellement les médias et autres fonctionnalités classiques.

Zooniverse : est une plateforme participative qui permet de faire participer les utilisateurs à l'identification des espèces. On est donc très dépendant des utilisateurs pour la classification, car il n'y a que la participation.

Les acteurs

Techniques & Scientifiques

Pour la reconnaissance d'image automatisée, une technique s'est rapidement démarquée des autres depuis les années 2010, il s'agit des réseaux de neurones artificiels. Cependant, cette idée n'est pas récente, on entend parler du premier réseau de neurones artificiel depuis 1943. C'est ce qu'on appelle les réseaux de neurones profonds qui émergent depuis les années 2010 avec l'arrivée du deep learning.

Il existe plusieurs types de réseau de neurones profonds, celui qui permet la reconnaissance d'image est appelé "réseau de neurones convolutif" en anglais CNN.

Il est très difficile d'entraîner des réseaux de neurones avec des entrées (inputs) à haute dimension comme les images. En informatique, une image est la plupart du temps composée de trois dimensions. Une dimension pour la hauteur, une pour la largeur et une pour le code RGB. Le modèle CNN consiste à appliquer des couches de convolutions puis des couches de pooling pour réduire la dimension de l'image tout en gardant les informations de cette image. On obtient ainsi en sortie (output) un vecteur qui est la représentation vectorielle des caractéristiques de l'image. On peut alors appliquer un réseau de neurones classique dit "vanilla" en anglais avec ce vecteur, voici un schéma d'un réseau de neurones vanilla.

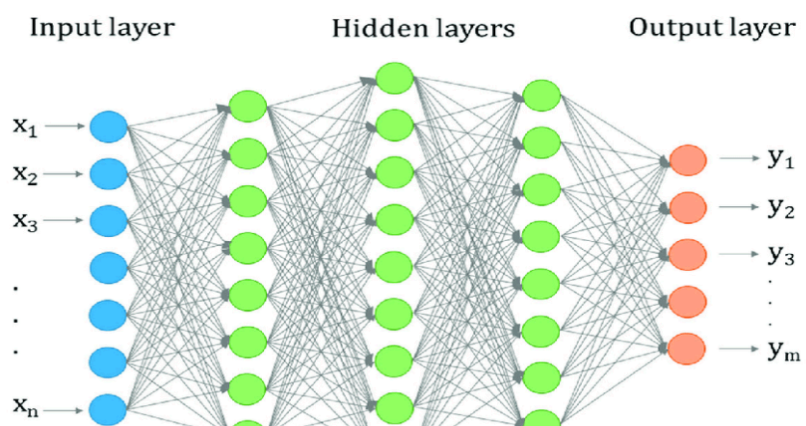


Figure 2 - Schéma d'un réseau de neurones classique dit vanilla

Un réseau de neurones classique est composé de plusieurs couches. La première couche est la couche d'entrée (input layer). Ensuite, il y a une ou plusieurs couches cachées (hidden layers), c'est ce qu'on peut appeler l'algorithme. Le nombre de couches et de neurones par couches est arbitraire. Enfin, il y a la couche de sortie (output layer) qui représente le résultat.

Chaque neurone est lié à tous les autres neurones de la couche précédente et chaque neurone à son biais sauf pour l'input layer.

La mécanique des réseaux de neurones vanilla est la suivante, on assigne un poids à chaque lien entre les neurones. Pour calculer la valeur d'un neurone, on fait la somme cumulée des poids et des valeurs de la couche précédente. On additionne ensuite le biais et enfin, on utilise une fonction, en général la fonction ReLU (Rectified linear unit) pour uniformiser les valeurs. Au départ de l'apprentissage d'un réseau, les poids et les biais sont

aléatoires et c'est ensuite le réseau lui-même qui va modifier ses poids et ses biais pour augmenter la précision de la classification de l'image, c'est le phénomène de propagation arrière (backpropagation).

Le deep learning fonctionne mieux avec de nombreux (par exemple, des millions) points de données étiquetés. L'apprentissage profond (deep learning) peut néanmoins profiter à ces projets grâce à l'apprentissage par transfert, dans lequel un réseau de neurones peut d'abord être formé sur des images disponibles dans d'autres grands ensemble de données, puis être formé sur un autre ensemble de données plus petit (notre cas au PNE). Les connaissances acquises sur le premier ensemble de données sont ainsi réaffectées à la classification du second ensemble de données, plus petit.

Pour réaliser l'apprentissage par transfert, l'ensemble de données ImageNet contient 1,3 million d'images étiquetées pour 1 000 catégories (des objets synthétiques tels que les bicyclettes et les voitures aux catégories d'animaux sauvages tels que les loups et les lions). Cet ensemble de données est couramment utilisé dans la recherche sur l'apprentissage par transfert. Cela dit, pour l'apprentissage par transfert il a été démontré que plus les classes d'images sont similaires entre les ensembles de données de départ et d'arrivée, plus l'apprentissage par transfert est efficace.

Les réseaux de neurones profonds ont donc une place importante dans les solutions innovantes d'identification et de classification automatisée d'images. La technologie continue d'évoluer et des pistes de travail sont actuellement étudiées par différents acteurs scientifiques. Notamment, l'étude de l'apprentissage actif, c'est un cycle dans lequel les humains étiquettent uniquement les images pour lesquelles le réseau n'est pas confiant, puis ces images sont ajoutées à l'ensemble de données, le réseau est réentraîné et le processus se répète. Une autre piste est la gestion automatique des images multi-espèces, certains groupes de recherche ont déjà des résultats assez satisfaisant sur la détection et le comptage de différentes espèces au sein d'une même image.

Dans un papier à propos des solutions de plateformes de traitement des données de piège photo à l'aide de l'IA publié par Juliana Vélez et John Fieberg en collaboration avec Hila Shamon, Paula J. Castiblanco-Camacho, William McShea, Michael A. Tabak, Carl Chalmers et Paul Fergus, il y a des conclusions très intéressantes sur les modèles de reconnaissance d'images. Il existe un défi commun associé à la reconnaissance d'images à l'aide de l'IA, c'est d'avoir une bonne précision lorsqu'on utilise de nouvelles images que le modèle n'a jamais vues auparavant. Les scientifiques constatent une variation de la précision selon les espèces. Néanmoins, l'IA est un outil très performant qui accélère et facilite le travail d'un opérateur humain. Pour la mise en œuvre d'un flux de travail entièrement automatisé, il faut s'assurer que la classification des espèces soit parfaite et cela nécessite d'entraîner un modèle existant avec ses propres données d'entraînement qui doivent être larges et fiables. Les modèles d'IA continuent à être mis à jour à l'aide de nouvelles données, ce qui devrait permettre d'améliorer les performances du modèle au fil du temps. À ce jour, peu de travaux ont été menés pour développer des modèles d'IA capables d'identifier des caractéristiques individuelles (sexe ou âge) ou des comportements (si les animaux se nourrissent, se déplace ou se reposent). Le deep learning jouera un rôle important dans la prédiction de ces caractéristiques et comportements une fois que davantage de données auront été collectées et rendues disponibles pour l'entraînement de nouveaux modèles encore plus performants.

Point juridique

Il n'existe pas de réglementation spécifique au système de piège photo. Il est possible de proposer un raisonnement par analogie avec le cadre juridique de la vidéo-protection.

Premièrement, la différence matérielle entre le piège photo et la vidéo-protection. On parle de piège photo lorsque l'appareil comporte un système de déclenchement automatique de prise de vues instantanées. La pose du piège doit se faire sans dégradation du support (arbre, poteau...). On parle de vidéo-protection lorsqu'une caméra est couplée à un moniteur et que les images prises de façon continue peuvent être visionnées sur un écran de télévision ou d'ordinateur.

Il existe une différence dans la réglementation en fonction du lieu. Si c'est un lieu ouvert au public ou un lieu privé accessible au public sans autorisation, le Code civil (art. 9) s'applique et prévoit que les prises de vues sont autorisées, mais que la diffusion des images est interdite sauf autorisation expresse préalable de la personne concernée et sauf dans le cadre d'une procédure pénale. Dans un lieu privé, les prises de vues et la diffusion sont interdites sauf autorisation expresse préalable du propriétaire. À noter qu'il est interdit de filmer la voie publique, même pour assurer la sécurité de sa clôture ou de ses biens. Seules les autorités publiques peuvent filmer la voie publique sous certaines conditions.

Pour être valablement mis en œuvre, un dispositif de captation d'image à des fins de police de l'environnement doit entrer dans un cadre juridique défini. En l'absence de réglementation spécifique et afin de garantir la recevabilité des images issues de pièges photo, le plus simple et le plus sûr juridiquement reste de faire autoriser expressément par le Parquet la mise en place de pièges photo dans le cadre d'une enquête préliminaire concernant la voie publique ou un lieu accessible au public sans autorisation où ont été régulièrement constatées des infractions.

Dans notre cas, il ne s'agit pas de faire de la police de l'environnement. L'installation de pièges photo est motivée uniquement à des fins de développement de la connaissance de la fréquentation à des fins scientifiques et de gestion des flux.

En résumé, il n'existe pas de réglementation spécifique aux pièges photo, il est nécessaire de faire des analogies avec la réglementation concernant la vidéo-protection. De plus, les textes traitent des usages et limites des pièges photo en matière de lutte pour la protection de l'environnement, ce qui n'est pas ce que le Parc national souhaite faire. C'est pour ces raisons que la proposition d'une doctrine pour le suivi de la fréquentation par piège photo associé à l'intelligence artificielle doit être écrite pour clarifier les choses.

Conclusion

Au travers de cet état de l'art, nous avons pu avoir une vision d'ensemble sur les différentes technologies et solutions existantes pour un suivi innovant de la fréquentation. Après avoir fait un récapitulatif de nos besoins au sein du Parc national des Écrins. L'étude plus ou moins approfondie de 30 technologies et solutions me permet d'avoir une bonne compréhension du sujet et d'établir une première version d'un protocole à mettre en place pour un suivi innovant de la fréquentation au Parc national des Écrins.

Comme nous avons pu le voir, il ne faut pas uniquement s'intéresser à la classification et à l'identification des images, mais à l'entièreté du workflow et donc au moyen du stockage des images prises par les pièges photo. Commençons par le point de départ du workflow, les pièges photos. Dans la section "Pièges IA", nous avons pu voir que les solutions de pièges photographiques avec intelligence artificielle embarquée sont encore au stade de prototype ou ne semblent pas répondre à nos attentes. On pourrait alors tester le prototype de WildCount. Bien que cette solution soit très prometteuse, nous ne pouvons pas nous contenter de tester uniquement cette solution, qui plus est ne possède qu'un seul prototype. Il s'agirait donc plutôt de coupler au test de WildCount, un système semi-automatisé pour le suivi de la fréquentation. La question d'un système entièrement automatisé peut se poser, mais cela nécessite une IA parfaitement entraînée avec des images provenant du PNE si on veut une précision quasi parfaite. Sans cette IA, il n'est pas préférable d'automatiser entièrement le système, car les prédictions de l'IA ne seront pas suffisamment élevées pour se permettre de ne pas avoir un opérateur humain qui vérifie ces prédictions.

Pour le système semi-automatisé, il faut également avoir des pièges photo, pour le modèle, nous pouvons nous baser sur l'expérience du CREA Mont Blanc, où ils ont utilisé des pièges photos Moultrie qui sont d'entrée de gamme avec une durée de vie de 4 à 5 ans puis ils ont été remplacés par des Reconyx qui sont des produits haut de gamme, c'est la référence des pièges photos. Il existe également la marque Bushnell qui se place en milieu de gamme. Une fois les pièges photos choisis et en notre possession, il faut les installer sur place et établir avec les agents de terrain des secteurs un protocole de récupération des cartes SD ainsi que de changement des piles. Ensuite, les images seraient transmises sur une machine possédant l'outil DeepFaune ainsi que l'outil de floutage "deface" pour les images d'humains. Les images seraient alors classées par catégories parmi humain, voiture, vide et chaque espèce reconnues par DeepFaune. Une fois les images identifiées et classifiées, il nous reste à les ranger dans une base de données organisée. À terme, nous pourrions utiliser GeoCam pour gérer notre ensemble de pièges photos ainsi que pour le stockage de nos images en projets distincts. Néanmoins, comme l'outil est encore au début de son développement, il est envisageable de creuser la piste intermédiaire d'utiliser une base de données interne au parc en attendant que GeoCam soit opérationnel.

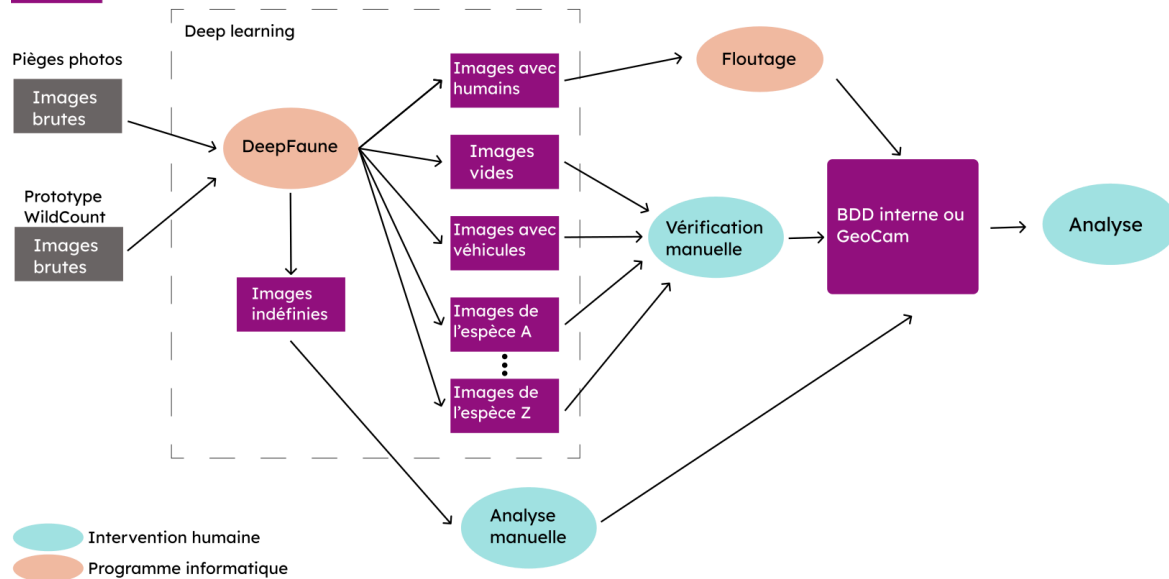


Figure 3 - Schéma du chemin des données des pièges photos dans un protocole semi-automatisé

La mise en place de ce protocole semi-automatisé pourrait permettre un gain de temps énorme pour l'analyse des images provenant des pièges photos. Selon un papier du Proceedings of the National Academy of Sciences (PNAS), pour étiqueter manuellement ~5,5 millions d'images, il faut près de 14,6 ans d'efforts à raison de 40 heures par semaine. Cet effort pourrait être réaffecté à des images plus difficiles ou à des problèmes plus ardues. Les gains de temps proviennent de trois sources : le filtrage automatique des images vides, l'acceptation des informations extraites automatiquement des images pour lesquelles le réseau a une grande confiance, et la fourniture aux étiqueteurs humains d'une liste triée de suggestions du modèle afin qu'ils puissent rapidement sélectionner les espèces, les dénombrements et les descriptions corrects. Toutefois, c'est encore difficile de quantifier le gain de temps qu'apporte un outil d'identification et de classification automatique.