



HAUTE ÉCOLE
D'INGÉNIERIE ET DE GESTION
DU CANTON DE VAUD
www.heig-vd.ch

HEIG-VD

GLEAM - RAPPORT INTERMÉDIAIRE

La Terre de nuit vue de l'espace

Antoine FRIANT
Haute École d'Ingénierie et de Gestion du
Canton de Vaud
Yverdon-les-Bains, VD, CH
antoine.friant@gmail.com

27 mai 2018

1 Cahier des charges

1.1 Résumé du problème

Les données géographiques sont nécessaires pour la prise de décisions importantes. Cependant la fiabilité et la disponibilité de ces données ne sont pas homogènes dans le temps et selon le lieu. Certaines de ces données ont une forte corrélation avec la lumière perçue par les satellites pendant la nuit.

Grâce à l'apprentissage automatique (*machine learning*), il est possible d'entraîner un réseau de neurones sur des données d'une date et d'un lieu connus pour reconstituer une carte de données géographiques à partir d'une image satellite nocturne.

Le travail à effectuer consiste à explorer différents types de données géographiques afin d'en choisir un, et faire de la prédiction sur ce type de données grâce à un réseau de neurones.

1.2 Objectifs

Le TB consiste dans un premier temps à explorer les données suivantes :

- Images satellites nocturnes de la Terre,
- Population humaine,
- Population animale,
- Densité végétale,
- PIB,

Et toutes autres données jugées pertinentes dans le but d'entraîner un réseau de neurones capable de prédire une estimation d'une donnée utile, à partir d'une image satellite de la terre de nuit.

La réalisation d'une application qui entraîne et exploite ce réseau de neurones est l'objectif de la seconde partie du TB.

Le but final est de pouvoir estimer, grâce au machine learning, des informations dont on ne possède pas de données à jour. Et cela à partir d'images satellites de nuit récentes, ou d'une combinaisons de ces images avec une autre donnée à jour.

1.3 Limitations

L'application sera compatible avec Windows 10 et Archlinux, et nécessitera l'installation de bibliothèques tierces (telles que Keras et TensorFlow). Elle ne possèdera pas nécessairement d'interface utilisateur.

L'utilisateur sera responsable de fournir les données à l'application dans un format supporté.

1.4 Description fonctionnelle

L'application prend en argument au moins deux jeux de données géographiques de format imposé : une image satellite nocturne et un autre type de donnée à déterminer au cours du projet. Après un long temps d'entraînement (une semaine au maximum, dépend de la machine utilisée).

Une fois le modèle généré, il est sauvegardé et réutilisable sur une autre image satellite nocturne (d'une date et/ou d'une région différente). Lorsque le modèle est appliqué sur une image satellite, une carte est recrée, affichant le résultats des prédictions.

Par exemple, si au cours du travail de bachelor il s'avère que la population par kilomètre carré est une donnée utile et utilisable, l'application devra prendre en argument une image satellite nocturne ainsi qu'une carte des populations de même taille et de même résolution pour entraîner le réseau de neurones. Une fois le modèle généré, l'application devra être capable de régénérer une approximation de la carte de population par kilomètre carré à partir d'une image satellite.

1.5 Délais

15 juin 2018 : Rapport intermédiaire

27 juillet 2018 : Rapport final et application fonctionnelle

Entre le 3 et le 14 septembre 2018 : Soutenance du travail de bachelor

Table des matières

1	Cahier des charges	i
1.1	Résumé du problème	i
1.2	Objectifs	i
1.3	Limitations	i
1.4	Description fonctionnelle	ii
1.5	Délais	ii
2	Résumé	1
3	Introduction	2
4	Exploration des données	3
5	Développement	4
6	Conclusion	5
7	Bibliographie	6
8	Authentification	7
9	Symboles et abréviations	8
10	Figures	9
11	Annexes	10

2 Résumé

3 Introduction

Les produits d'imagerie satellite sont devenus abondants et largement accessible au cours des vingt dernières années. De nombreux satellites prennent des photographies de la Terre à chaque heure du jour *et de la nuit*. Ces observations nocturnes révèlent des caractéristiques peu évidentes de jour, parfois même cachées. Les routes apparaissent, les villes montrent leurs lumières, même les bateaux de pêche aveuglent les océans avec des projecteurs pour attirer les poissons.

La disponibilité, la résolution et l'uniformité de la qualité de ces données contraste fortement avec le manque de fiabilité d'autres informations géographiques utiles lors de prises de décisions importantes. Par exemple, la densité de la population est une estimation précise en Suisse mais très approximative au Kenya. D'autres mesures intéressantes incluent : la consommation en électricité, les émissions de CO₂, la couverture végétale et la présence de faune. Les lumières nocturnes observées depuis l'espace donnent des indications sur chacune de ces mesures alors qu'elles peuvent manquer dans une région à une date donnée.

Le but de ce projet est d'extraire autant d'information que possible de l'imagerie satellite nocturne en utilisant l'apprentissage automatique (*machine learning*) sous la forme de réseau de neurones.

4 Exploration des données

Source de données (nightlights, poprasters)

Problème de format pour la superposition des rasters

QGIS pour la visualisation

Recherche de corrélation, graphes

GPD, groupement par pays, filtré par développement économique, graphes

deltas de luminosité par pays, incomplet

Données encore pas explorées : human footprint, ...

5 Développement

réseau de neurones à convolution pour faire de la régression => chercher des cas de problèmes similaires

installation : keras, tensorflow, setup nvidia

métaparamètres du réseau de neurones

6 Conclusion

Travail effectuer, travail restant

7 Bibliographie

8 Authentication

9 Symboles et abréviations

10 Figures

11 Annexes