

DOSSIER DE COMPÉTENCES

Benoit BAZARD
Software Engineer

Table des matières

Expériences professionnelles.....	2
Amadeus 2018 (en cours).....	2
Stage Dassault Aviation 2017 (2 mois).....	2
Thèse sur le deep learning 2016 – 2017 (10 mois).....	3
Stage Outscale 2015 (3 mois).....	5
Projets d'étude à Telecom SudParis.....	6
Intelligence artificielle d'échecs 2015 (3 mois).....	6
Microscopie à contraste de phase 2016 (3 mois).....	7
Space Invaders 2016 (2 semaines).....	8
Projets personnels.....	9
Cryptopals Challenges 2018.....	9
Nextcloud et synchronisation de calendriers 2018.....	10
Site Web Personnel 2017.....	10
Random Notes Generator 2015.....	10
Formations.....	11
Langues.....	11

Contact

email : benoit.bazard@gmail.com

site web : bbazard.github.io

EXPÉRIENCES PROFESSIONNELLES

AMADEUS

2018 (EN COURS)

Trop récent pour donner des détails.

STAGE DASSAULT AVIATION

2017 (2 MOIS)

Contexte du projet :

- Présentation du machine learning et discussion de son utilité pour l'essai en vol (20 slides).
- Discussion avec les différentes parties prenantes pour mieux comprendre les modes de pannes des capteurs.
- Découverte des projets existants de détection automatique de pannes. Et décision d'où (quels capteurs, quels avions, quels vols) des techniques de machine learning serait le plus efficaces.
- Mise en place d'une architecture de détection automatique de pannes de capteurs.
 - Création d'un data warehouse – les données n'étaient pas centralisées et dans un format créé en interne.
 - Manipulation des données pour un data scheme approprié aux algorithmes.
 - Choix des algorithmes de machine learning pouvant effectuer cette détection de pannes.

L'implémentation n'a pas pu être réalisée à cause d'un problème de santé.

Environnement technique : C, Python

Contexte du projet :

Segmentation des caractères de plaques d'immatriculation avec un réseau de neurones en python pour une entreprise thaïlandaise de vidéosurveillance.

Cette thèse s'inscrit dans le cadre d'un double diplôme passé à l'Asian Institute of Technology.

État de l'art:

Les techniques suivantes de segmentation ont été étudiées :

- Region-Convolutional Neural Network (R-CNN)
- Pixel Projection
- Hidden Markov Chain (HMM)
- Component Analysis
- Long Short Term Memory (LSTM)

Méthodologie :

Pour une thèse, l'approche doit être innovante, il n'est donc pas possible de juste implémenter une des techniques existantes, il a fallu que je conçoive une nouvelle approche. Pour cela je me suis principalement inspiré des techniques de R-CNN et de Pixel Projection.

On extrait d'abord la ligne d'immatriculation de la plaque, puis dans un deuxième temps chaque caractère séparément.

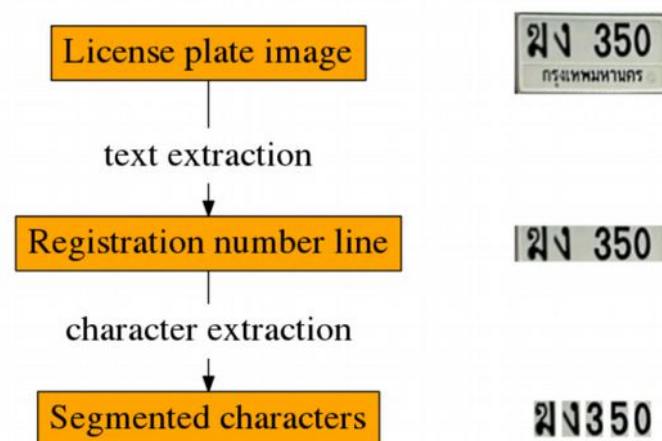
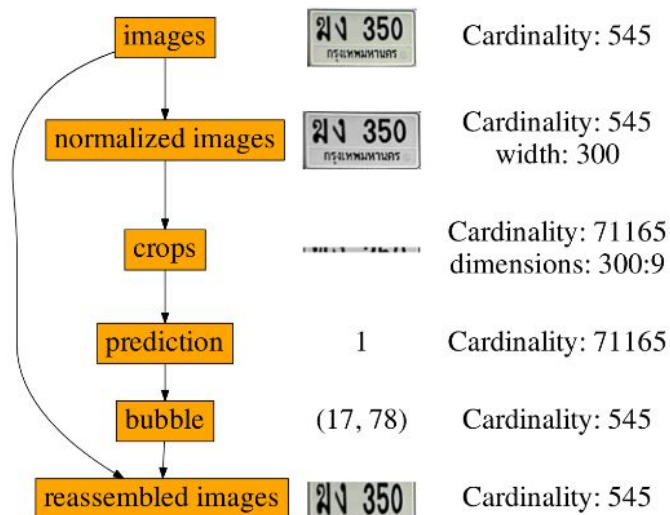


Schéma de la segmentation en deux phases.

Pour cela la plaque est découpée en lamelles se chevauchant. Chaque lamelle passe dans un réseau de neurones qui les classifie en bon (1) ou mauvais (0), la ligne d'immatriculation est reconstituée d'après les bonnes lamelles.



*Ici la ligne d'immatriculation extraite se situe entre la ligne de pixels 17 et 78 de l'image originale.
La prédiction de la lamelle (crop) est bonne (1).*

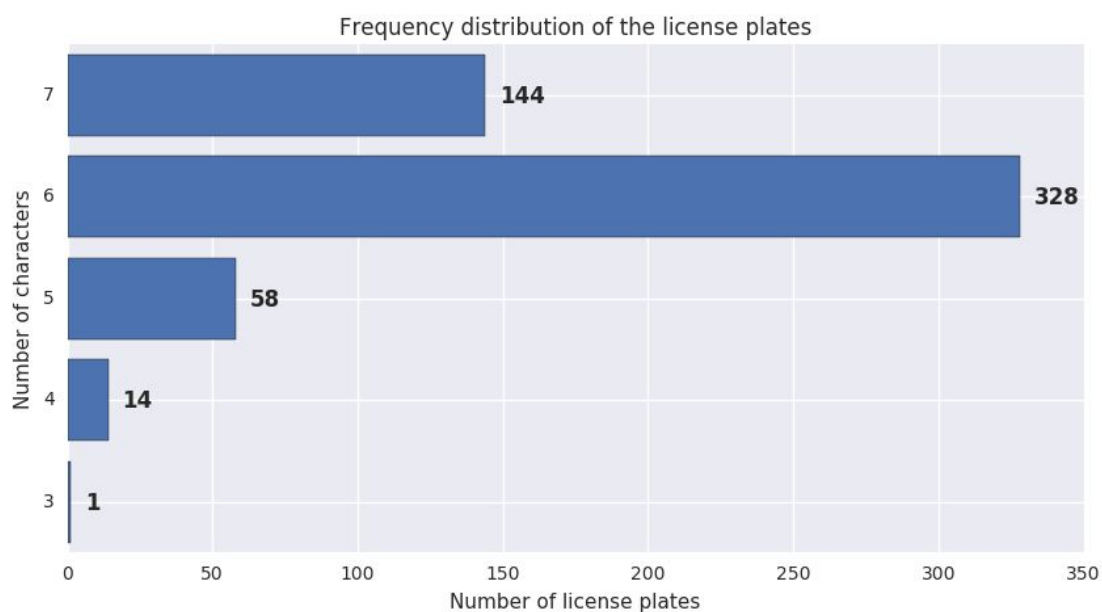
En tout, les 545 plaques d'immatriculation forment 71 165 lamelles pour la première phase.

Un procédé similaire est ensuite appliqué à la ligne d'immatriculation pour récupérer les caractères de la plaque.

Pour plus d'informations et de schémas, la thèse (en anglais) :

<https://github.com/BBazard/thesis/blob/master/bbazardThesis.pdf>

Afin de mieux comprendre le dataset de plaques d'immatriculation qui m'a été fourni pour entraîner les réseaux de neurones, j'ai créé plusieurs visualisations. Le dataset comportait 545 plaques d'immatriculation.



Les plaques d'immatriculation du dataset ont tendance à avoir six caractères.



L'agencement des caractères des plaques d'immatriculation thaïlandaise n'est pas normalisé.

Sur cette dernière image, chaque ligne correspond à une plaque d'immatriculation, chaque segment correspond à la place qu'occupe un caractère horizontalement. Les couleurs servent juste à mieux compter le nombre de caractères.

On remarque que parmi les plaques d'immatriculation à six caractères, ces caractères peuvent être agencés très différemment. La deuxième ligne (frame31) est une plaque d'immatriculation avec trois caractères à gauche et trois caractères à droite alors que la dernière ligne est une plaque d'immatriculation avec deux caractères à gauche et quatre caractères à droite (frame452).

Le code matplotlib des visualisations est disponible à :

<https://bbazard.github.io/visualization>

Environnement technique : Python (numpy, pandas, matplotlib, caffe, scikit-learn), Anaconda, NVIDIA DIGITS, Jupyter, LaTeX

STAGE OUTSCALE

2015 (3 MOIS)

Contexte du projet :

Stage de découverte de l'univers du Cloud, du Big Data et de procédés DevOps.

- Écriture d'un livre blanc sur le Big Data et le Cloud
- Installation Hadoop (MapR) sur le cloud réparti sur trois machines virtuelles par SSH
- Déploiement de machines virtuelles avec Docker, Vagrant, VirtualBox
- Déploiement de Nginx dans le cloud
- Utilisation de boto (interface python pour Amazon EC2, compatible Outscale)
- Website Scraping (Beautiful Soup, Selenium)
- Découverte d'outils : Jira, Ansible, Jenkins

Environnement technique : Python, Beautiful Soup, Selenium, Docker

PROJETS D'ÉTUDE À TELECOM SUDPARIS

INTELLIGENCE ARTIFICIELLE D'ÉCHECS

2015 (3 MOIS)

Contexte du projet :

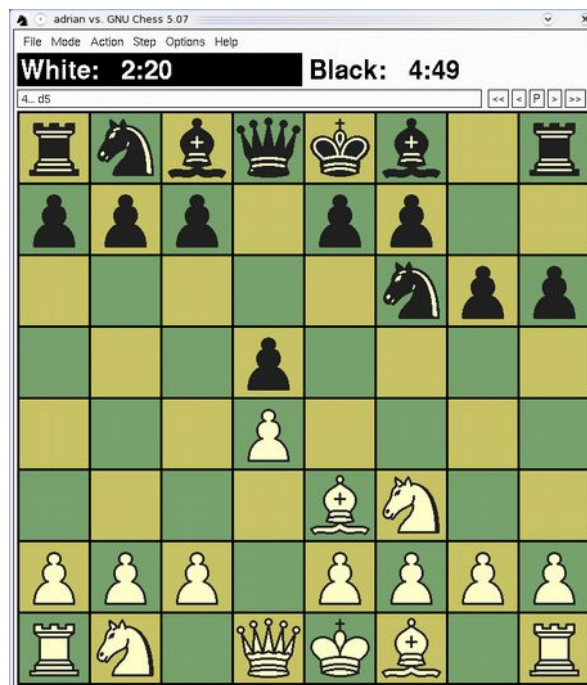
En groupe de quatre personnes, implémentation d'un algorithme min-max avec une heuristique pour déterminer si une position d'échiquier donnée est plus ou moins favorable.

Utilisation du protocole UCI pour pouvoir jouer contre l'intelligence artificielle avec une interface graphique grâce aux deux utilitaires :

- XBoard - X Window System Chess Board
- Polyglot - Chess engine protocol adaptor

Xboard permet d'avoir une interface graphique pour jouer contre l'intelligence artificielle.

Polyglot permet de recevoir par chaîne de caractères l'état de l'échiquier et d'envoyer par chaîne de caractères le prochain coup que l'intelligence artificielle a décidé de jouer.



Capture d'écran du GUI XBoard.

En plus de programmer, j'étais aussi le chef de projet. J'ai réparti la charge de travail en fonction des compétences de chacun, enseigné les bonnes pratiques git, valgrind, lint et revu le code des moins expérimentés pour donner des conseils.

<https://github.com/BBazard/chessaint>

Environnement technique : C, Subversion, Doxygen, Valgrind, gdb, cpplint.py

Contexte du projet :

Un microscope à contraste de phase est capable de prendre plusieurs images d'un même objet en variant uniquement la phase. Il existe un moyen de combiner toutes ces images qui ne diffèrent que par leurs phases pour créer une unique image de très haute résolution. Pour cela on utilise un algorithme qui, entre autres, passe les images dans le domaine de Fourier puis dans le domaine temporel plusieurs fois.

En groupe de deux, implémentation de l'algorithme de Gerchberg-Saxton.

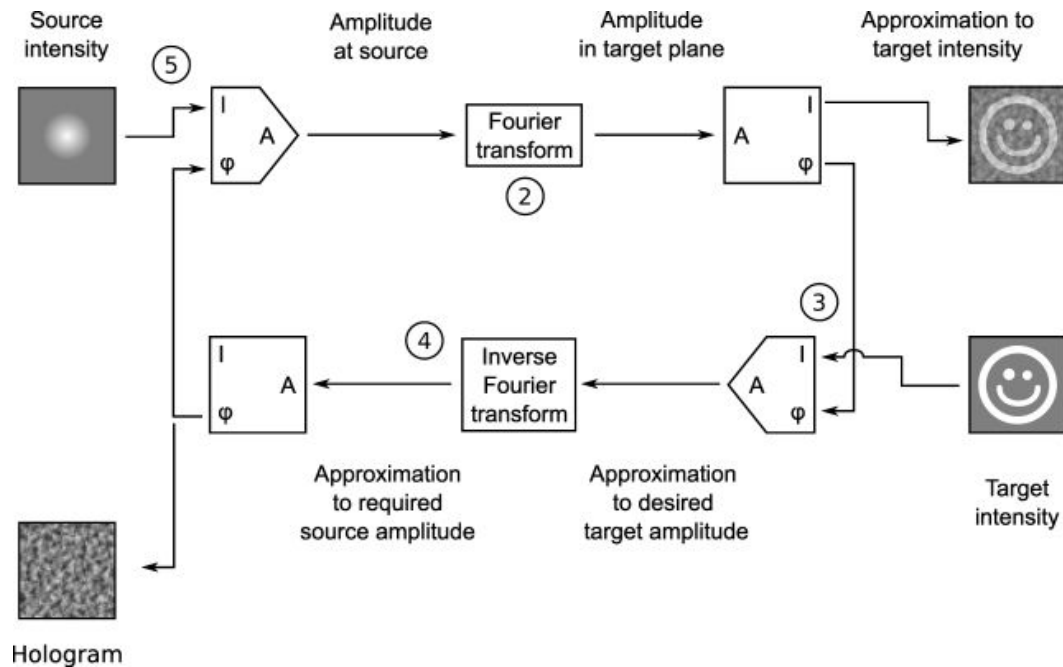


Diagramme de l'algorithme de Gerchberg-Saxton.

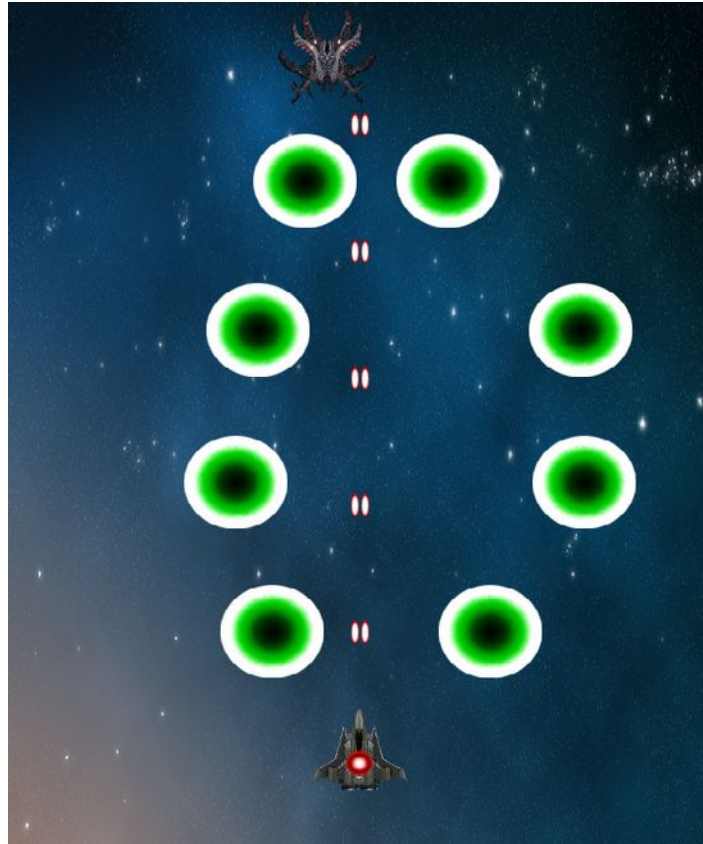
<https://github.com/BBazard/fourierscope>

Environnement technique : C, openmp, libtiff (tiff format), libfftw3 (Fourier Transform), Google Test, Doxygen

Contexte du projet :

Pour un cours sur les C++11 smart pointers.

En groupe de deux, implémentation d'un space invaders clone avec le design pattern Entity Component System. Certains projectiles forment des courbes de lagrange pour les rendre moins prédictibles.



Capture d'écran au cours d'une partie.

Les images des sprites ont été trouvées sur internet.

<https://github.com/BBazard/INTouhou>

Environnement technique : C++

PROJETS PERSONNELS

CRYPTOPALS CHALLENGES

2018

Contexte du projet :

Série 1 à 4 de cryptopals.com une série de challenges cryptographiques.

Implémentation de différentes primitives cryptographiques et attaques dont :

- Detection oracle
- Padding oracle
- Bitflipping attacks
- Byte-at-a-time ECB decryption
- Crack an MT19937 seed
- Clone an MT19937 RNG from its output
- Break a SHA-1 keyed MAC using length extension

Crypto Challenge Set 3

This is the next set of **block cipher cryptography** challenges (even the randomness stuff here plays into block cipher crypto).

This set is **moderately difficult**. It includes a famous attack against CBC mode, and a "cloning" attack on a popular RNG that can be annoying to get right.

We've also reached a point in the crypto challenges where all the challenges, with one possible exception, are valuable in breaking real-world crypto.

17. [The CBC padding oracle](#)
18. [Implement CTR, the stream cipher mode](#)
19. [Break fixed-nonce CTR mode using substitutions](#)
20. [Break fixed-nonce CTR statistically](#)
21. [Implement the MT19937 Mersenne Twister RNG](#)
22. [Crack an MT19937 seed](#)
23. [Clone an MT19937 RNG from its output](#)
24. [Create the MT19937 stream cipher and break it](#)

Capture d'écran de la série 3.

https://github.com/BBazard/cryptopals_challenges

Environnement technique : Python

Contexte du projet :

Mise en place d'une instance nextcloud sur un réseau privée (ip en 192.168.*.*) avec création d'un trusted certificate authority pour avoir une synchronisation du calendrier par SSL entre PC Linux et téléphone Android.

Environnement technique : Docker, Apache, OpenSSL

SITE WEB PERSONNEL

2017

Contexte du projet :

Création d'un site web portfolio : bbazard.github.io

Environnement technique : HTML, CSS, JavaScript

RANDOM NOTES GENERATOR

2015

Contexte du projet :

LilyPond snippet qui génère une partition de musique aléatoire en pdf pour travailler la lecture de notes.



Exemple de mesures générées par le script.

<https://github.com/BBazard/random-notes-generator>

Environnement technique : Scheme, Lilypond

FORMATIONS

2014 – 2018	Diplôme d'ingénieur de Télécom SudParis Évry
2016 – 2017	Master of Science de l'Asian Institute of Technology Bangkok – Thaïlande
2011 – 2014	Préparation aux grandes écoles PT Lycée Chevrollier – Angers
2011	Baccalauréat scientifique, option mathématiques Lycée David d'Angers – Angers

LANGUES

FRANÇAIS	Langue maternelle
ANGLAIS	Courant (TOEIC 955/990 TOEFL 97/120)
PORTUGAIS BRÉSILIEN	Usage intermédiaire (Famille brésilienne)
ALLEMAND	Usage intermédiaire (LV1 au lycée)
JAPONAIS	Usage débutant (LV3 à Télécom SudParis)
THAÏLANDAIS	Usage débutant (Je sais lire le thaïlandais)