

DÉPARTEMENT DE GÉOMATIQUE APPLIQUÉE
FACULTÉ DES LETTRES ET SCIENCES HUMAINES
UNIVERSITÉ DE SHERBROOKE

ESSAI

SEGMENTATION SÉMANTIQUE EN TEMPS RÉEL À
PARTIR D'UN NANO-ORDINATEUR : ÉTUDE DES
PERFORMANCES ET DES LIMITES.

SOUS-TITRE

*Dans cadre de la maîtrise en sciences géographiques cheminement de type cours en géo
développement durable*

VINCENT LE FALHER
LEFV2603

LONGUEUIL
SEPTEMBRE 2020

Table des matières

Liste des figures	1
Liste des tableaux	1
1 Mise en contexte	1
2 Problématique	2
3 Objectifs	3
4 Identification des besoins en termes de données	4
4.1 Stratégie de recherche de données	5
4.2 Conditions d'acquisition des nouvelles données	6
4.3 Synthèse des données	7
5 Processus de gestion de l'essai	11
6 Échéancier	11

Liste des figures

1 Échéancier	13
------------------------	----

Liste des tableaux

1 Tableau des données	8
---------------------------------	---

1 Mise en contexte

Voici le plan qui est utilisé pour rédiger la mise en contexte de l'essai.

- Présentation du projet avec PJCCI ; contexte, objectif et brève synthèse de la solution proposée.
- Mise en contexte de l'essai par rapport au projet avec PJCCI ;
- Introduction à la détection d'objets en temps réel ;
- Introduction aux objets connectés, aux nano-ordinateurs ;
- Introduction aux réseaux de neurones ;
- Introduction à la segmentation sémantique avec les réseaux de neurones convolutifs entiers (FCN) ;
- Conclure avec les avantages et les défis du projet PJCCI.

La compagnie LES PONTS JACQUES CARTIER ET CHAMPLAIN INCORPORÉE (PJCCI) désire évaluer la mise en service de la piste multifonctionnelle (vélos, piétons, etc.) du pont Jacques-Cartier, à Montréal, durant l'hiver. Pour ce faire, la piste doit rester sécuritaire et dégagée, malgré les événements météorologiques.

L'université de Sherbrooke, qui participe à cette initiative, propose de mettre en place sur le pont une plateforme de détection innovatrice qui consiste à installer plusieurs paires d'objets connectés ultralégers et performants (des nano-ordinateurs) à différents endroits du pont. Chacun de ces nano-ordinateurs possède trois différents types de capteurs : vision, son, et météorologiques (température, humidité, etc.). Chaque nano-ordinateur d'une paire perçoit le même environnement, mais d'une perspective différente que son homologue : la caméra pointe vers la même surface, mais d'un autre point de vue ; les sons et les données météorologiques sont captés dans le même voisinage. Les données collectées par les capteurs sont traitées en temps réel par des algorithmes de détections performants qui sont adaptés à ce type de problématiques : les réseaux de neurones, du domaine de l'intelligence artificielle. La déduction de l'état de la surface de la piste (sèche, mouillée, glacée, etc.) se fait en fusionnant les différentes perceptions (multi-cibles) de chaque capteur (multi-capteurs).

L'objet principal de cet essai consiste à étudier la capacité du nano-ordinateur du fabricant NVIDIA, le Jetson nano [17], à exécuter, en temps réel, un modèle de réseau de neurones entraîné à faire de la segmentation sémantique (classification) d'images de haute résolution qui sont perçues avec la caméra. Les résultats de cette étude permettront de déterminer le modèle de réseau de neurones le plus adapté pour répondre aux besoins du volet vision du projet pour PJCCI.

La détection d'objets et de surface en temps réel est de plus en plus précise et efficace depuis que les performances des systèmes informatisés permettent l'exécution d'algorithmes exigeants, en majeure partie depuis l'utilisation des processeurs graphiques "GPU" [6] [8] [2] [11] [22] [13]).

Les systèmes informatiques performants sont de plus en plus miniatures, on parle de nano-ordinateurs et des objets connectés ("Internet of Things" ou "IoT") [4] [21]. Ils permettent la détection en temps réel à des endroits, dans des situations et dans des conditions qui n'étaient pas envisageables il y a encore 10 ans ([22] [3] [1] [4]).

Les réseaux de neurones ont aussi très rapidement progressé depuis 2012 [2], permettant d'offrir des alternatives aux solutions de détection et de classifications, entre autres [20]. Les réseaux de neurones convolutifs entiers ("FCN" en anglais, pour "Fully Convolutional Network") sont les derniers à avoir émergé ("state-of-art") [22] et à profiter au domaine de la vision et de la détection

d'objets ([16] [22]).

La segmentation sémantique est une forme de classification d'image, pixel par pixel, qui tire profit des dernières évolutions de la classification supervisée grâce aux réseaux de neurones convolutifs entiers, et se permet d'être déduite en temps réel avec des nano-ordinateurs ([14] [4]). Les images doivent être de très haute résolution, ce qui nécessite d'avoir à disposition un système informatique capable de fournir une puissance de calcul appropriée, particulièrement pour la manipulation de la mémoire et des nombres flottants pendant l'inférence [15]. Leur application par des nano-ordinateurs est un défi en raison de la faible consommation d'énergie (Watts) et de la puissance de calcul limitée de ces derniers [7].

Pour PJCCI, les avantages d'une telle plateforme seraient multiples, et on peut en énumérer plusieurs, sans se limiter à : contrôler l'épandage de sel ; surveiller les conditions de la piste multifonctionnelle ; suivre les effets du gel et du dégel ; optimiser les coûts des opérations d'entretien (déplacements, quantité) ; offrir aux usagers des conditions d'accès sécurisées et optimales même en hiver ; effets environnementaux atténués ; détecter ce qu'un spécialiste humain ne pourrait pas ou aurait des difficultés à détecter ; prise de décision et gestion proactive ; planification.

D'un autre côté, les défis ne sont pas à sous-évaluer : la détection doit être précise, fiable et consistante, tout cela afin d'assurer aux usagers un service de qualité dans un contexte sécuritaire.

2 Problématique

Voici les problèmes que l'essai va contribuer à résoudre.

- Défi pour le domaine du transport actif et durable d'être soutenu par des solutions technologiques fiables (opérationnelles) pour pouvoir offrir des services de qualité et sécuritaire sur l'ensemble des quatre saisons ;
- Capacités d'un nano-ordinateur, tel que le NVIDIA Jetson nano, de faire de la segmentation sémantique en temps réel avec des vidéos de hautes résolutions et d'images par seconde élevée ;
- Problématiques des solutions existantes pour la segmentation sémantique en temps réel sur nano-ordinateur ;

Dans le cadre du projet pour PJCCI, une plateforme technologique sera mise à la disposition des gestionnaires du pont afin de les aider à prendre les décisions les plus responsables et raisonnables possibles. Mais la mise en opération d'une solution innovante et fiable, qui concilie des algorithmes d'apprentissage profond, du temps réel, des nano-ordinateurs, et des conditions climatiques variables, est complexe. Dans une certaine mesure l'essai va contribuer à la recherche de solutions afin de répondre au défi pour le domaine du transport actif et durable d'être soutenu par des solutions technologiques fiables (opérationnelles), l'objectif étant de pouvoir offrir des services de qualité et sécuritaire sur l'ensemble des quatre saisons.

La seconde problématique que l'essai va contribuer à résoudre concerne les limites d'un nano-ordinateur. Un nano-ordinateur est un ordinateur miniaturisé en taille, mais aussi limité en capacité. Il existe différents fabricants et modèles, de spécifications variées, pour répondre à différents besoins. Le dernier né est le modèle "Jetson nano" du fabricant "NVIDIA ", disponible depuis juin 2019 au prix très abordable de 99\$US, et qui sera le matériel utilisé dans le cadre de cet essai. La compagnie NVIDIA a conçu ce matériel spécialement pour l'inférence de modèles d'apprentissage

profond sur une plateforme mobile (drone) ou proche des données ("edge" en anglais). L'inférence nécessite une architecture et une puissance machine différente de celle nécessaire pour l'entraînement. Les modèles de réseaux de neurones sont adaptés et optimisés pour l'inférence. L'essai va permettre de préciser les capacités du Jetson nano pour l'inférence de diverses architectures de réseaux de neurones convolutifs entiers (FCN en anglais) et la segmentation sémantique en temps réel avec des vidéos de différentes propriétés (résolutions et nombre d'images par seconde). Il existe des tests encourageants ([18] [16] [6]), qui seront utilisés comme modèle, même si ceux-ci sont limités à des types d'application qui ne sont pas les mêmes que pour l'essai.

Il est difficile de trouver des jeux de données pour entraîner les réseaux de neurones convolutifs entiers adaptés à la problématique. La technique de "Data augmentation" permet de démarrer d'un modèle qui a déjà appris avec un jeu d'images important (milliers d'images), et de lui faire apprendre davantage, en lui fournissant un plus petit jeu d'images (centaines d'images) de la nouvelle zone d'étude. Par exemple un modèle peut avoir appris à classifier des images de la Californie, États-Unis. Pour lui permettre de classifier des images de la Ville de Sherbrooke, il est souhaitable de lui fournir un nouveau jeu de données spécifique à cette ville afin qu'il s'adapte (ses paramètres) à cette région. Dans le contexte de cet essai, les données acquises sur le terrain seront fournies aux différents modèles qui seront évalués, et qui seront ré-entraînés avec ce nouveau jeu d'images adapté à la zone d'étude.

La paramétrisation (des "hyper-paramètres") des réseaux de neurones est très "subtile" et "intuitive" et requière de l'expérience. C'est un processus d'essais-erreurs qui est très coûteux en temps, et risqué puisqu'il n'y a aucune garantie de succès. La technique de "Transfer Learning" permet d'hériter d'un modèle qui est déjà entraîné et configuré, et de l'adapter pour répondre à ses besoins. Cette technique permet un gain en temps et en énergie (et en argent) important puisque le temps de conception (architecture et configuration) et le temps d'entraînement, de validation et de tests sont diminués de façon non négligeable. La problématique pour l'essai est de trouver le modèle qui est le plus adapté pour répondre au besoin, et il en existe des milliers [12]. La recherche dans la littérature permet heureusement de limiter les choix et donner des pistes ([22] [16] [18]). La problématique de la conception existe toujours, car le modèle a besoin d'être étudié, adapté et ré-entraîné, jusqu'à l'obtention de résultats probants. Mais la paramétrisation des hyper-paramètres n'est plus nécessaire (supposément), ce qui est très avantageux.

3 Objectifs

L'objectif principal de l'essai est d'évaluer la capacité du nano-ordinateur "NVIDIA Jetson nano" à exécuter, en temps réel, un modèle de réseau de neurones à convolution entier (FCN) permettant la segmentation sémantique d'une piste multifonctionnelle. La segmentation sera limitée à deux classes : une classe pour la piste et une autre pour les autres éléments présents dans l'image. Pour y arriver, différents sous-objectifs ont été établis.

Voici en bref les sous-objectifs, avec un peu plus de détails par la suite pour chacun d'eux :

- Évaluer les limites de la plateforme, matérielle et applicative ;
- Trouver des moyens d'optimiser la plateforme, au besoin, d'un point de vue matériel, mais aussi applicatif ;
- Permettre un accès à distance sécurisé au nano-ordinateur ;
- Documenter l'approche, les tests, et les résultats ;

— Tester le mode multi-cible (optionnel).

Le premier sous-objectif est de déterminer quelles sont les limites de la plateforme, d'un point de vue matériel (GPU, CPU, mémoire, transfert mémoire, consommation, etc.), mais aussi applicatif (entraînement, inférence). Cette phase du projet va permettre d'exécuter différents modèles déjà existants, sans modification, en tenant compte des éléments documentés dans la littérature [16] [22] [18]. Selon le déroulement de cette étape, un ou plusieurs modèles seront sélectionnés.

Un autre sous-objectif est d'optimiser ou d'adapter la plateforme, d'un point de vue matériel, mais aussi applicatif, afin d'avoir les meilleures performances et résultats possibles pendant l'entraînement et l'inférence.

Comme les résultats devront être disponibles en tout temps, une connexion à distance sécurisée devra être mise en place. Cette connexion permettra aussi de pouvoir prendre le contrôle du nano-ordinateur à distance et de l'administrer.

L'approche, les tests, et les résultats seront documentés. Il y aura beaucoup d'activités relatives à la conception et aux tests, le cheminement complet ne sera pas fourni. Une synthèse sera préférée et les informations les plus pertinentes seront incluses. Les détails de l'installation de l'environnement de développement et des applications, librairies et autres dépendances nécessaires seront inclus, ainsi que ceux de la configuration. Dans le cas où l'objectif principal n'est pas atteint, ou partiellement, la/les raison/s de l'échec seront spécifiées et des pistes de solutions potentielles proposées.

Optionnellement, si le temps le permet, le mode multi-cible pourrait être testé. L'idée est d'avoir deux mêmes nano-ordinateurs qui pointent vers la même cible, et de combiner les deux résultats afin d'améliorer les chances de déduction. Il serait envisageable d'utiliser une deuxième cible "virtuelle" sur le même nano-ordinateur dans un premier temps, afin de ne pas dépendre d'un deuxième nano-ordinateur.

4 Identification des besoins en termes de données

Voici les données qui ont été identifiées comme nécessaires.

- jeux de données (images) d'entraînement des modèles existants ;
- vidéos de la zone d'étude ;
- nouvelles images, extraites des vidéos ;

Il existe déjà des modèles de réseaux de neurones dont l'application est la segmentation sémantique. Ces modèles ont été entraînés avec des jeux de données qui sont disponibles gratuitement. Il est planifié utiliser ces modèles déjà pré-entraînés, et les ré-entraîner en bonifiant les jeux de données d'entraînement avec des données locales. Ces jeux de données seront aussi utiles pour valider et tester le résultat du ré-entraînement.

Selon nos recherches, à ce jour, il n'existe pas de modèles de réseaux de neurones qui ont été entraînés pour répondre directement aux objectifs de l'essai. Il sera donc nécessaire de construire des jeux de données représentant la zone d'étude, le "nouveau domaine" de l'apprentissage. Ces nouvelles images permettront d'adapter à ce nouveau domaine les modèles de réseaux de neurones qui auront été sélectionnés (technique "Adaptation domain" en anglais). Les images seront extraites des vidéos acquises dans les conditions décrites ci-après (section 4.2).

Les vidéos sont la source de données principales de l'essai : l'inférence est faite à partir de vidéos en temps réel. La résolution et le nombre d'images par seconde seront à déterminer. Les vidéos seront acquises dans les conditions décrites ci-après (section 4.2).

4.1 Stratégie de recherche de données

Voici les stratégies qui ont été identifiées pour la recherche de données.

- ressources NVIDIA ;
- références en lien avec le sujet ;
- Internet ;
- Association des piétons et cyclistes du pont Jacques-Cartier ;
- travaux d'étudiants de l'Université de Sherbrooke ;
- acquisition des vidéos de la piste multifonctionnelle du pont Jacques-Cartier ;
- discussions avec mon directeur de projet ;

Les ressources mises à disposition par le constructeur du Jetson nano, NVIDIA, seront étudiées pour apprendre et tester le nano-ordinateur. Parmi les plus intéressantes, on peut citer le "Jetson Nano Developer Kit", le "NVIDIA Deep Learning Institute", la communauté Jetson, les tutoriels, les "benchmarks". Des jeux de données sont fournis gratuitement.

En complément des ressources de NVIDIA, deux références scientifiques seront principalement utilisées comme points de départ et comme modèles pour l'essai, car leurs études ont été faites avec le Jetson nano ([16] et [6]). Beaucoup de références ont été publiées ces deux dernières années sur le sujet de la segmentation sémantique, ils existent donc de multiples alternatives inspirantes.

Internet est une mine d'information et de données. Il existe des forums et des blogues dans lesquels des utilisateurs publient leurs expérimentations de la segmentation sémantique en temps réel avec le Jetson nano ([9]), ou plus génériquement la segmentation sémantique. Des sites comme "modelzoo.co" sont des entrepôts de modèles déjà pré-entraînés. Une autre option est d'effectuer une recherche d'images ou de vidéos de la piste multifonctionnelle du pont Jacques-Cartier via les sites de recherche tels que Google.

L'Association des piétons et cyclistes du pont Jacques-Cartier existe depuis de nombreuses années pour promouvoir le transport actif et conserver la piste multifonctionnelle du pont Jacques Cartier ouverte durant l'hiver. Ils fournissent, via des sites Internet, des collections de vidéos et d'images qui pourraient être utilisées. Il serait aussi possible d'entrer en contact avec l'association et leur demander de prendre de nouvelles vidéos. Voir <http://pontjacquescartier365.com>, et <https://www.flickr.com/photos/pontjacquescartier>.

Une autre possibilité serait d'hériter des acquisitions faites par un autre étudiant de l'université de Sherbrooke, soit déjà archivée, soit collectée prochainement. Mon directeur de projet Mickaël G. m'a informé qu'un étudiant de Sherbrooke va avoir besoin de collecter le trafic automobile sur le campus de l'Université de Sherbrooke, à Sherbrooke.

Enfin il y a l'acquisition des vidéos spécifiquement pour le projet PJCCI, tel que documenté à la section 4.2. Comme il n'y a aucune date de planifiée pour la capture des vidéos, l'essai devra s'arranger pour dépendre le moins possible d'elles durant la préparation et le développement, et s'attendre à les recevoir pour le ré-apprentissage et les tests, en fin d'essai.

Tout au long de l'essai, mon directeur Mickaël sera une ressource importante afin de vérifier que les sources de données, les prétraitements et les traitements sont adéquats aux attentes du projet pour PJCCI.

4.2 Conditions d'acquisition des nouvelles données

Les images et vidéos qui seront acquises devront répondre à différents requis, afin de pouvoir adapter le modèle de réseau de neurones dans des conditions les plus proches de la réalité.

- L'objet d'intérêt pour la collecte des données est la surface de la piste multifonctionnelle du pont Jacques-Cartier, à Montréal (Québec, Canada).
- L'acquisition se fera à partir de différents points d'intérêt sur le pont. Ces points d'intérêt seront déterminés par les gestionnaires du projet pour PJCCI.
- L'appareil d'acquisition sera installé sur un trépied pendant la capture afin d'avoir un point de vue stable et constant.
- Pour chaque point d'intérêt, la hauteur du trépied, en plus de son géoréférencement, l'angle de vision et la direction seront documentés. Les points d'intérêt seront géoréférencés afin d'être représentés sur une carte. La précision du géoréférencement devra être assez précise pour que d'autres acquisitions puissent avoir lieu. De plus, au moins trois points de référence sur les bords de l'image devront être déterminés pour garder le même angle de vision et la direction à chaque acquisition.
- La capture sera faite pendant une période assez prolongée, telle que des périodes de 15 ou 30 minutes. La période d'acquisition est à la discrétion des architectes applicatifs du projet pour PJCCI.
- La même surface sera capturée de différents angles en simultané avec deux appareils d'acquisition. Cela permettra d'avoir deux images de la même zone au même moment, mais d'une perspective différente.
- Comme indiqué précédemment, plusieurs points d'intérêts de la piste multifonctionnelle seront capturés, toujours de différents angles en simultané.
- L'acquisition sera faite de jour.
- La luminosité devra être bonne, normale.
- L'acquisition dans des conditions spéciales est exclue, telle que la nuit, une luminosité ou une visibilité extrême, comme aucune, ou médiocre.
- Des images des différentes conditions de la surface seront acquises : sèche, partiellement mouillée, trempée, partiellement recouverte de neige, recouverte de neige, glace noire, "slush", etc. Cela devrait correspondre aux classes du projet avec PJCCI et non celles de l'essai. Il existe de nombreuses références à ce sujet, par exemple [5], [10], [19].
- Les vidéos devront capturer du trafic : piéton, vélo, coureur, poussette, groupe, chiens en laisse, etc. Cela va permettre d'entraîner le modèle lorsqu'il y a des obstacles sur la piste et tester la fiabilité de la segmentation sémantique.
- Les vidéos seront acquises en différentes résolutions : haute (1080p/i), puis de plus en plus faible 760p/i, 576p/i, 480p/i, 360p/i. Cela va permettre de tester les performances avec une perte progressive de la qualité (nombre de pixels).
- Les vidéos seront acquises avec un nombre d'images par seconde différente (FPS, "Frames Per Second" en anglais) : élevés (30FPS), puis de moins en moins rapide, 20FPS, 10FPS, 1FPS. Cela va permettre de tester les performances et déterminer le meilleur compromis

entre résolutions et FPS.

- Il n'est pas important que l'appareil d'acquisition soit le même que celui qui sera utilisé pendant l'essai, car les images sont adaptées par l'environnement de développement ("framework" en anglais) d'apprentissage profond au moment du prétraitement de l'entraînement.
- Comme il y aura une quantité non négligeable de nouvelles données acquises, qui de plus sont des données multimédias de haute résolution, un espace de rangement ("storage" en anglais) suffisant et performant sera nécessaire. Un espace de sauvegarde est de plus indispensable pour conserver les données brutes, mais aussi les différentes versions dues aux traitements. Un espace dans le nuage ("cloud") est une option non négligeable, mais la décision est à la discrétion des architectes applicatifs du projet pour PJCCI.
- Une nomenclature pour le nom des répertoires et des fichiers vidéos sera déterminée, à la discrétion des architectes applicatifs du projet pour PJCCI. Les données utilisées pour l'essai seront gérées à ma discrétion (espace de rangement, copies de sauvegarde, nomenclature).

4.3 Synthèse des données

Voici le tableau de synthèse des données, incluant la référence avec leur réseaux de neurones.

TABLE 1: Tableau des données

	Spécification	Description
1	réseau : U-Net jeu de données : Membrane (origine isbi challenge) nombre d'images : 30 résolution/s : 512x512	C'est le jeu de données pour le réseau U-Net. Il est utilisé dans le benchmark de NVIDIA pour l'inférence avec le Jetson nano. Les images sont de type médicale. À noter que le framework "Keras" s'occupe de l'augmentation de données. https://github.com/zhixuhao/unet/tree/master/data/membrane
2	réseau : SegNet jeu de données : CamVid vidéos : 10 minutes résolution/s : HD	SegNet est un réseau qui a été créé pour la segmentation sémantique de vidéos. Il a été entraîné avec le jeu de données de CamVid, qui procurent des vidéos de la route avec la même perspective que le conducteur du véhicule. Un modèle entraîné est disponible pour le Jetson nano. https://github.com/PengKiKi/camvid
3	réseau : MFANet jeu de données : Cityscapes nombre d'images : 5000 résolution/s : 1280x1024	MFANet est un réseau qui a été créé en 2019 pour la segmentation sémantique sur des appareils tel que le Jetson nano. Il a été entraîné avec le jeu de données de Cityscapes, qui procurent des images de scènes urbaines. Différentes stratégies d'augmentation de données sont utilisées. Des tests ont été fait avec le Jetson nano. leejy@ustb.edu.cn
4	réseau : MAVNet jeu de données : Penstock nombre d'images : 135 résolution/s : 1280x1024	C'est l'un des deux jeux de données pour le réseau MAVNet. Les images sont celles de "conduites forcées", des voies d'eau de régulation, et sont préparées pour la segmentation sémantique. Des tests ont été fait avec le Jetson nano. https://github.com/tynguyen/MAVNet/tree/master/data/TN_penstock
5	réseau : MAVNet jeu de données : Penstock nombre d'images : 135 résolution/s : 1280x1024	C'est l'un des deux jeux de données pour le réseau MAVNet. Les images sont celles de drones volant à l'intérieur d'un bâtiment, et préparées pour la segmentation sémantique. Des tests ont été fait avec le Jetson nano. https://github.com/tynguyen/MAVNet/tree/master/data/perch_drone
6	réseau : RESNet18 jeu de données : Cityscapes nombre d'images : 25 000 résolution/s : 360x720, 512x256, 1024x512, 2048x1024	Cityscapes est un jeu de données qui fournit des images de rues spécifiquement destinées pour la segmentation sémantique. Il peut être utilisé par de nombreux réseaux. RESNet18 a été entraîné avec ce jeu et est disponible en diverses résolutions pour le Jetson Nano. https://github.com/tynguyen/MAVNet/tree/master/data/perch_drone

	Spécification	Description
7	réseau : RESNet18 jeu de données : DeepScenes nombre d'images : 15 000 résolution/s : 576x320, 864x480	DeepScene propose un modèle et un jeu de données. Le modèle est entraîné avec différents jeux de données, comme Cityscapes, SUN-RGBD, Synthia. Le jeu de données fournit des images de forêt, qui est destinée pour la segmentation sémantique. RESNet18 a été entraîné avec ce jeu et est disponible en deux résolutions pour le Jetson Nano. http://deepscene.cs.uni-freiburg.de
8	réseau : RESNet18 jeu de données : Multi-Human nombre d'images : 25 043 résolution/s : 512x320, 640x360	Le jeu de données Multi-Human fournit des images contenant des humains, et qui est destinée pour la segmentation sémantique. RESNet18 a été entraîné avec ce jeu et est disponible en deux résolutions pour le Jetson Nano. https://lv-mhp.github.io/dataset
9	réseau : RESNet18 jeu de données : Pascal VOC nombre d'images : 11 530 résolution/s : 320x320, 512x320	Le jeu de données Pascal VOC fournit des images de classes variées tel que des personnes, des animaux, des véhicules, et des objets classiques, et qui peut être utilisé pour la segmentation sémantique. RESNet18 a été entraîné avec ce jeu et est disponible en deux résolutions pour le Jetson Nano. http://host.robots.ox.ac.uk/pascal/VOC/voc2012/index.html
10	réseau : RESNet18 jeu de données : SUN RGB-D nombre d'images : 10 335 résolution/s : 512x400, 640x512	Le jeu de données SUN RGB-D fournit des images de scènes d'intérieur de maison, et qui est destiné pour la segmentation sémantique. RESNet18 a été entraîné avec ce jeu. https://synthia-dataset.net
11	réseau : DeepScene jeu de données : Synthia nombre d'images : 220 000 résolution/s : 1280x760	Le jeu de données Synthia fournit des images (et vidéos) de scènes de rue comme celui de Cityscapes, et qui est destiné pour la segmentation sémantique. DeepScene a été entraîné avec ce jeu. Il n'a pas été testé avec le Jetson Nano. http://3dvision.princeton.edu/datasets.html
12	jeu de données : Association des piétons et cyclistes pont Jacques-Cartier nombre d'images : 313 résolution/s : variées	L'Association des piétons et cyclistes du pont Jacques-Cartier a une collection d'images et de vidéos de la piste multifonctionnelle du pont Jacques-Cartier. Ce n'est pas un jeu de données qui est prêt à être utilisé pour l'apprentissage tel-qu'il, il doit être préparé. Mais c'est une source de données qui est très importante pour l'essai. Il est envisagé de contacter l'association au besoin afin de leur demander leur collaboration pour la collecte d'autres d'images ou vidéos. https://www.flickr.com/photos/pontjacquescartier http://pontjacquescartier365.com/videos-pont-jacques-cartier

	Spécification	Description
13	<p>jeu de données : images et vidéo sur Internet</p> <p>nombre d'images : entre 30-50</p> <p>résolution/s : variées</p>	<p>Internet est une source de données non négligeable en terme de données. Quelques images et vidéos de la piste multifonctionnelles du pont Jacques-Cartier, autres que celles fournies par L'Association des piétons et cyclistes du pont Jacques-Cartier, sont disponibles. Ce n'est pas un jeu de données qui est prêt à être utilisé pour l'apprentissage tel-quel, il doit être préparé. Mais c'est une source de données qui est très importante pour l'essai.</p> <p>https://google.ca</p>
14	<p>jeux de données : Kaggle</p>	<p>Le site Kaggle offre une vingtaine de jeux de données offert par la communauté pour faire de la segmentation sémantique, et qui sont prêt à être utilisé. Les jeux de données n'ont pas été évalués.</p> <p>https://www.kaggle.com/search?q=</p>

5 Processus de gestion de l'essai

L'environnement nuagique Office 365 (Office, OneDrive) sera utilisé afin de partager le rapport, les analyses, les références, les résultats et les discussions.

La méthodologie Agile sera utilisée afin d'être efficace et proactive dans le développement et le suivi des tâches. L'application Web "Trello" sera utilisée pour la gestion des tâches.

Le code source sera conservé dans le contrôleur de source gitHub en tant que projet public. Le code pourra ainsi être facilement partagé avec le directeur et accessible publiquement et gratuitement. Le type de la licence n'a pas encore été décidé.

Au niveau de la communication, Outlook365, le téléphone, les messages textes et Skype seront utilisés au besoin. Skype permettra de partager l'écran.

Ayant de l'expérience professionnelle dans une grande compagnie où la collaboration avec des employés localisés à travers le Canada est très bien implantée, et les déplacements limités, les rencontres physiques avec le directeur seront peu nombreux. Il s'agira sûrement de se rencontrer au moment de concevoir l'architecture, puis afin de clarifier certains points et incertitudes, ou l'utilisation d'un tableau blanc sera alors grandement utile. Autrement les outils de collaboration existant en ligne remplissent amplement leur rôle.

Les horaires disponibles pour communiquer par téléphones seront entre 14h et 18h, entre le lundi et le vendredi. Certaines soirées de semaine seront probablement mises à profit, au besoin.

L'échange de documents par courriel sera fortement déconseillé. Les commentaires pourront être échangés par courriel, par téléphone, message texte, par chat, mais il est attendu qu'ils soient intégrés directement dans les documents en ligne.

Je serais responsable de la gestion des outils de collaboration et de la planification des rencontres.

Le directeur sera responsable de savoir collaborer avec ces outils sans toutefois avoir à les maîtriser.

Le projet sera étalé sur deux sessions, c'est-à-dire 38 semaines. Les efforts sont estimés entre 1 à 3 jours par semaine, l'équivalent de 40 à 90 jours d'effort (250 à 500 heures). Étant donné que je suis travailleur à temps plein (37.5/semaine), père de famille de deux jeunes enfants (8 & 10.5) et propriétaire d'une maison, les soirs de semaines seront dédiés à l'essai, entre 21h00 et 23h00 (10 heures étalées sur 5 jours). Les fins de semaine seront principalement utilisées pour se reposer et "décrocher". Une seule semaine de pause est prévue (vacances scolaires de mars). Je planifie prendre entre cinq et douze jours (vendredi et/ou lundi) de congés durant la semaine, au besoin.

6 Échéancier

Hypothèses

- L'échéancier a été conçu de façon à concilier ma vie de famille, mon travail, mes activités personnelles et la réalisation de mon essai.
- La revue de la littérature a déjà été commencée bien avant le début de la session.
- Le plan de travail a été complété pendant le cours GAE723 avant le début de la session.
- Les sources de données ont déjà été choisies et vérifiées.

- L'environnement de travail et de développement a déjà été choisi et évalué.
- Des prises de notes et brouillons sont prises au fur et à mesure de l'exécution, et mises à jour. Les parties qui peuvent être rédigées sans risque d'être retouchées peuvent être complétées. Sinon la rédaction finale et propre de l'essai sera faite dans la dernière partie de l'échéancier.
- L'inscription se fait à la session d'hiver.
- Le début de l'essai commence la deuxième semaine de janvier.
- L'inscription à l'essai est faite le 21 janvier.
- Le dépôt initial est planifié dans la semaine du 24 août.
- Le dépôt final est planifié dans la semaine du 28 septembre.
- L'essai se déroule sur deux sessions, hivers et été, échelonné sur une durée de 38 semaines.
- Il y a une semaine de pause d'incluse, pendant la semaine de relâche scolaires 29 mars.
- Mes vacances d'été sont planifiées fin Août, du 14 août au 03 septembre.
- Le directeur a deux semaines pour réviser les parties qui lui sont remises. Un temps est alloué avec lui afin de réviser ses commentaires.
- Il est planifié revoir l'échéancier régulièrement, surtout pour s'adapter aux événements non sollicités et inattendus.

une présentation de style Gant, liste des activités et des semaines; doit être lisible et tenir sur une page (landscape) lisible

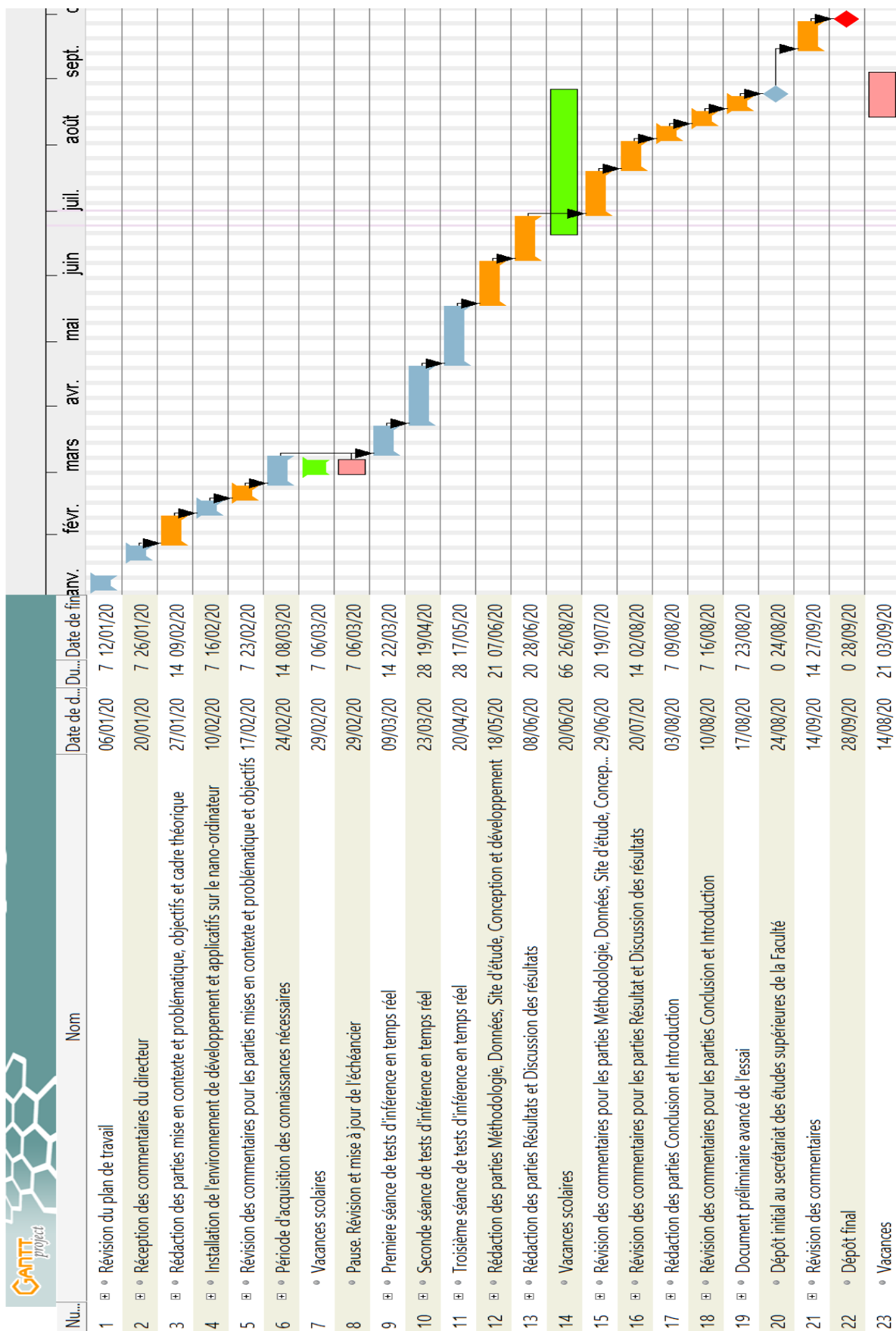


FIGURE 1 – Échéancier

Bibliographie

- [1] S. ABOUZAHER, M. SADIK et E. SABIR. "IoT-Empowered Smart Agriculture : A Real-Time Light-Weight Embedded Segmentation System". In : *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (2017), p. 319-332. DOI : 10.1007/978-3-319-68179-5_28.
- [2] A. BEAM. *Deep Learning 101 - Part 1 : History and Background*. 2017. URL : https://beamandrew.github.io/deeplearning/2017/02/23/deep_learning_101_part1.html.
- [3] M. BERNAS, B. PLACZEK et A. SAPEK. "Edge Real-Time Medical Data Segmentation for IoT Devices with Computational and Memory Constrains". In : *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (2017), p. 119-128. DOI : 10.1007/978-3-319-67077-5_12.
- [4] B. BLANCO-FILGUEIRA et al. "Deep Learning-Based Multiple Object Visual Tracking on Embedded System for IoT and Mobile Edge Computing Applications". In : *IEEE Internet of Things Journal* (juin 2019), p. 5423-5431. DOI : 10.1109/JIOT.2019.2902141.
- [5] L. CHENG, X. ZHANG et J. SHEN. "Road Surface Condition Classification Using Deep Learning". In : *Journal of Visual Communication and Image Representation* (2019). DOI : 10.1016/j.jvcir.2019.102638.
- [6] C. P. CHONG, C. A. T. SALAMA et K. C. SMITH. "Real-Time Edge Detection and Image Segmentation". In : *Analog Integrated Circuits and Signal Processing* (1992), p. 117-130. DOI : 10.1007/BF00142412.
- [7] M. COPEL. *What's the Difference Between Deep Learning Training and Inference ?* Août 2016. URL : <https://blogs.nvidia.com/blog/2016/08/22/difference-deep-learning-training-inference-ai/>.
- [8] T. DETTMERS. *Deep Learning in a Nutshell : History and Training*. Déc. 2015. URL : <https://devblogs.nvidia.com/deep-learning-nutshell-history-training/>.
- [9] F. DUSTIN. *Realtime Semantic Segmentation on Jetson Nano in Python and C++*. Oct. 2019. URL : <https://www.linkedin.com/pulse/realtime-semantic-segmentation-jetson-nano-python-c-dustin-franklin>.
- [10] Liping FU et al. "A Risk-Based Approach to Winter Road Surface Condition Classification". In : *Canadian Journal of Civil Engineering* (mar. 2017), p. 182-191. URL : <https://www.nrcresearchpress.com/doi/full/10.1139/cjce-2016-0215>.
- [11] JIACONDA. *A Concise History of Neural Networks*. Avr. 2019. URL : <https://towardsdatascience.com/a-concise-history-of-neural-networks-2070655d3fec>.
- [12] J. Y. KOH. *Model Zoo - Deep Learning Code and Pretrained Models for Transfer Learning, Educational Purposes, and More*. 2018. URL : <https://modelzoo.co/>.
- [13] A. KURENKOV. *A 'Brief' History of Neural Nets and Deep Learning*. 2015. URL : <https://www.andreykurenkov.com/writing/ai/a-brief-history-of-neural-nets-and-deep-learning/>.

- [14] J. LONG, E. SHELHAMER et T. DARRELL. “Fully Convolutional Networks for Semantic Segmentation”. In : *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Juin 2015, p. 3431-3440. DOI : 10.1109/CVPR.2015.7298965.
- [15] M. MODY et al. “Low Cost and Power CNN/Deep Learning Solution for Automated Driving”. In : *Proceedings - International Symposium on Quality Electronic Design, ISQED*. 2018, p. 432-436. DOI : 10.1109/ISQED.2018.8357325.
- [16] T. NGUYEN et al. “MAVNet : An Effective Semantic Segmentation Micro-Network for MAV-Based Tasks”. In : *arXiv :1904.01795 [cs]* (juin 2019). URL : <http://arxiv.org/abs/1904.01795>.
- [17] NVIDIA. *Jetson Nano*. Mar. 2019. URL : <https://developer.nvidia.com/embedded/jetson-nano>.
- [18] NVIDIA. *Jetson Nano : Deep Learning Inference Benchmarks*. Avr. 2019. URL : <https://developer.nvidia.com/embedded/jetson-nano-dl-inference-benchmarks>.
- [19] Guangyuan PAN et al. “Winter Road Surface Condition Recognition Using a Pre-Trained Deep Convolutional Neural Network”. In : (), p. 13.
- [20] D. PATHAK et M. EL-SHARKAWY. “Architecturally Compressed CNN : An Embedded Realtime Classifier (NXP Bluebox2.0 with RTMaps)”. In : *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*. Jan. 2019, p. 0331-0336. DOI : 10.1109/CCWC.2019.8666495.
- [21] N. SHARMA, M. SHAMKUWAR et I. SINGH. *The History, Present and Future with Iot*. Intelligent Systems Reference Library. Springer Science and Business Media Deutschland GmbH, 2019. ISBN : 18684394 (ISSN). DOI : 10.1007/978-3-030-04203-5_3. URL : https://www.scopus.com/inward/record.uri?eid=2-s2.0-85060126104%5C&doi=10.1007%5C%2F978-3-030-04203-5_3%5C&partnerID=40%5C&md5=1878f15afc39fbca5142a5f680e0f3c7.
- [22] J. ZHENG et al. “Real-Time Semantic Segmentation Network for Edge Deployment”. In : *Proceedings of 2019 Chinese Intelligent Systems Conference*. Sous la dir. d’Y. JIA, J. DU et W. ZHANG. Springer Singapore, 2020, p. 243-249. ISBN : 978-981-329-697-8 978-981-329-698-5. DOI : 10.1007/978-981-32-9698-5_28. URL : http://link.springer.com/10.1007/978-981-32-9698-5_28.