



UNIVERSITÉ
LAVAL

FACULTÉ DES SCIENCES ET GÉNIE

IFT-3580 Stage en informatique 2
Été 2022 - À distance
Baccalauréat en informatique

Stagiaire de recherche en Vision par ordinateur

Laboratoire de Vision et Systèmes Numériques
LVSN

Rapport de fin de stage

6 septembre 2022

Destinataire

Département des stages en milieu pratique

Travail réalisé par :
Isabelle EYSSERIC
NI : 111189609

Superviseur de stage :
Jean-François Lalonde

1 Résumé

Le présent rapport présente le travail réalisé par Isabelle Eysseric, stagiaire en recherche en vision par ordinateur au sein du laboratoire de vision et systèmes numériques, le LSVN[1] du département de génie électrique et de génie informatique de l'université Laval au cours de la session d'été 2022 dans le cadre des travaux de recherche du projet Sentinelle Nord sur le design biophilique[2], attrait inné de l'humain envers la nature, en Arctique afin de concilier bien-être et performance énergétique.

Le stage consistait à rechercher et tester différents algorithmes de segmentation sur des panoramas 360 degrés du jeu de données du laboratoire afin de pouvoir décrire le contenu des scènes et de comprendre l'impact de chaque élément sur l'aspect photopique/mélanopique de la lumière. Ultimement, ce jeu de donnée calibré permettra d'apprendre à prédire ces mesures à partir d'une seule image.

Remerciements

Isabelle Eysseric souhaite remercier toute l'équipe du LSVN[3], celle de l'Institut de l'Intelligence des Données[4], le IID, et celle de Jean-François Lalonde[5] pour son soutien et son aide, mais plus particulièrement son superviseur Jean-François Lalonde[6], professeur et expert en vision par ordinateur, en apprentissage machine (ML), en réalité augmentée (AR) et en réalité virtuelle (VR) et son étudiant à la maîtrise Christophe Bolduc, expert en vision numérique.

Sans oublier Julie Buquet pour son aide avec la base de données Matterport3D[7] et Frédéric Fortier-Chouinard pour celle avec le Secure Shell (SSH) et le serveur du laboratoire.

Liste des symboles et abréviations

Abréviation	Définition
ULaval	Université Laval
LVSN	Laboratoire de vision et systèmes numériques de ULaval
IID	Institut d’Intelligence des Données de l’Université ULaval
GRAAL	Groupe de recherche en apprentissage automatique de ULaval
REPARTI	Regroupement pour étude environnements partagés intelligents répartis
FRQ-NT	Font de Recherche du Québec, domaine de Nature et Technologie
IA	Intelligence artificielle
ML	Apprentissage machine
DL	Apprentissage profond
RL	Apprentissage par renforcement
NLP	Apprentissage automatique de la langue
AR	Réalité augmentée
VR	Réalité virtuelle
IR	Vision infrarouge
2D, 3D	2 dimensions, 3 dimensions
HDR	Haute plage dynamique
RGB	Rouge Vert Bleu
SSH	Secure Shell
GAN	Réseau Antagoniste Génératif
DCGAN	Réseau Antagoniste Génératif à Convolution Profond
CNN	Réseaux de Neurones Convolutifs
R-CNN	Réseau de Neurone Convolutif basé sur les Régions
Fast R-CNN	Réseau de Neurone Convolutif basé sur les Régions Rapide
Faster R-CNN	Réseau de Neurone Convolutif basé sur les Régions plus Rapide
Mask R-CNN	CNN basé sur les Régions qui prédit, extension de Faster R-CNN
Mesh R-CNN	CNN basé sur les Régions qui prédit, extension de Mesh R-CNN
DCNN	Réseau de neurones à convolution profonde
FCN	Réseau entièrement convolutionnel
CRF	Champs aléatoires conditionnels
CRF-RNN	CRFs Entièrement connecté
RPN	Réseau de proposition de région

Liste des tableaux et des figures

Figure	Titre	Page
Figure 1	Génération d'images en couleur et génération d'images hybrides	9
Figure 2	Tutoriel DCGAN avec Pytorch	10
Figure 3	Les types de segmentation sémantique	11
Figure 4	Architecture d'un réseau pour la segmentation sémantique	11
Figure 5	Jeu de données de Matterport3D et COCO Détection	12
Figure 6	Modèle DeeplabV3 de Pytorch et modèle HoHoNet	13
Figure 7	Image originale 9C4A9529 avec les trois résultats	15

Table des matières

1	Résumé	2
2	Introduction	6
2.1	Présentation personnelle	6
2.2	Présentation de l'organisation	6
2.3	Présentation du stage et de son environnement	6
3	Responsabilités et tâches du stagiaire	7
3.1	Rôle et contribution du stagiaire	7
3.2	Objectifs et problématiques	7
3.3	Théorie	7
3.4	Méthodologie	7
3.5	Échéancier et contraintes	8
3.6	Description des tâches et travaux effectués	8
3.6.1	Phase : Formations	8
3.6.2	Phase : Projet sur la lumière HDR dans les panoramas	10
3.6.3	Groupe de lecture les vendredis	13
3.7	Résultats	13
3.8	Comparaison	14
4	Développement et renforcement des compétences	14
4.1	Techniques	14
4.2	Scientifiques	15
4.3	Communication	15
4.4	Réflexion sur la formation pratique	15
4.5	Réflexion sur la formation théorique	15
4.6	Bilan des acquis	15
5	Conclusion	16
6	Référence	17

2 Introduction

2.1 Présentation personnelle

La stagiaire, Isabelle Eysseric, était à sa dernière année du programme de baccalauréat en informatique et avait déjà accumulé la totalité de ses crédits, c'est-à-dire 90, lors du début de son stage en mai 2022. Elle avait également cumulé assez de crédits dans des cours d'intelligence artificielle afin d'obtenir une concentration en Systèmes intelligents.

2.2 Présentation de l'organisation

Le Laboratoire de vision et systèmes numériques (LVSN), du Département de génie électrique et de génie informatique, de la Faculté des sciences et de génie, de l'Université Laval regroupe les travaux de huit professeurs et de près de soixante étudiants gradués, chercheurs post-doctoraux et professionnels de recherche autour du thème de la vision par ordinateur et de ses applications industrielles et biomédicales.

L'organisation se situe au pavillon Adrien-Pouliot de l'université Laval, au local 00116-00118 du département de génie électrique et de génie informatique de la Faculté des sciences et génie. Plus précisément au 1065, avenue de la Médecine, à Québec, dans la province du Québec, Canada, G1V 0A6. L'organisation compte une trentaine de personnes.

2.3 Présentation du stage et de son environnement

Le LVSN constitue un environnement de recherche dynamique et stimulant, bénéficiant d'une infrastructure très développée. Ses membres poursuivent des collaborations multiples aux plans national et international. Le LVSN fait partie du Regroupement pour l'étude des environnements partagés intelligents répartis (REPARTI[8]) du FRQ-NT[9] et compte deux chaires de recherche. Des séminaires et conférences y sont périodiquement organisés et les travaux sont publiés dans la littérature spécialisée internationale.

Le projet de Sentinelle Nord de 2020 est celui qui vise à générer de nouvelles stratégies architecturales tenant compte des cycles lumineux des cinq saisons inuites liées au bien-être et aux bâtiments à faible consommation d'énergie dans des climats extrêmes. Les LIDAR fourniront les données tridimensionnelles nécessaires pour utiliser la réalité augmentée comme outil de conception, pertinent pour la science architecturale et de bâtiment. Les solutions architecturales incluront ainsi des espaces physiques (captures existantes d'Ikaluktutiak) et numériques (nouvelles propositions de design) qui coexistent et interagissent en temps réel grâce à la réalité virtuelle et augmentée.

L'étudiante assistait un étudiant à la maîtrise dans la réalisation de sa tâche de calibration colorimétrique d'un jeu de données de panoramas à haute plage dynamique (HDR) pour le développement de techniques simplifiant la capture mélanopique et photopique d'une scène. Le jeu de données calibré permettra d'apprendre à prédire ces mesures à partir d'une seule image. Les

données colorimétriques sont essentielles pour comprendre l'effet de l'éclairage d'une scène sur l'humain, autant en concentration qu'en appréciation d'un espace.

Celle-ci avait pour mandat la segmentation sémantique d'images panoramiques afin de pouvoir décrire le contenu des scènes du jeu de donnée et de comprendre l'impact de chaque élément sur l'aspect photopique et mélanopique de la lumière.

3 Responsabilités et tâches du stagiaire

3.1 Rôle et contribution du stagiaire

Le rôle principal associé au poste de stagiaire au sein de l'équipe du projet de Sentinelle Nord et de l'étudiant à la maîtrise était de faire la segmentation sémantique sur les données panoramiques du laboratoire LVSN. Pour ce faire, il fallait rechercher et tester des algorithmes d'intelligence artificielle appliqués à la segmentation sémantique dans le domaine de la vision par ordinateur, collecter et traiter les données nécessaires à l'exécution de ces algorithmes, organiser et présenter les résultats des expérimentations tout en assistant l'étudiant gradué, Christophe Bolduc, dans ses travaux de recherche.

3.2 Objectifs et problématiques

Étant donné que la lumière intérieure n'est pas uniforme, elle change en fonction de la position de l'observateur et de la source lumineuse, l'objectif était donc de combiner des mesures denses afin de pouvoir prédire le type de lumière à un endroit précis de l'image et avoir une meilleure compréhension de l'éclairage dans celle-ci avec la segmentation sémantique.

3.3 Théorie

Étant un stage en recherche en informatique, le travail s'est effectué en trois grandes étapes : la revue littéraire sur les meilleurs algorithmes de segmentation puis les tests afin de faciliter l'étudiant dans son projet et l'atteinte de son objectif.

3.4 Méthodologie

Pour faire au mieux ce stage, quelques méthodes de travail ont été mises en place. Au niveau de l'organisation, l'outil Google Document était utilisé afin de définir les tâches et suivre l'avancement du projet. Au niveau de la communication, les outils comme Slack[10], Zoom[11] et PowerPoint[12] de Microsoft étaient utilisés. Étant principalement à distance, les membres de l'équipe utilisaient quelques outils afin de s'entraider, Slack pour communiquer à l'écrit, Zoom à l'oral, tandis que PowerPoint permettait de présenter les résultats du projet.

3.5 Échéancier et contraintes

L'étudiante n'avait pas d'échéancier précis pour réaliser ses tâches. L'échéance était calculée ou ajusté chaque semaine après chaque présentation en fonction des tâches réalisées.

3.6 Description des tâches et travaux effectués

3.6.1 Phase : Formations

L'objectif pour la première partie du stage était de se familiariser avec la vision par ordinateur ainsi que la manipulation d'images dans le langage python. Pour se faire, l'étudiante a dû participer à plusieurs conférences, suivre quelques formations sur différents sites web et lire un livre sur le domaine concerné.

- **Formation sur l'intelligence artificielle**

Le Bootcamp IID[13] se donnait à distance les 2 et 3 mai par des étudiants qui en font l'usage au quotidien, accompagnait de périodes d'échange, de réseautage et de discussion, permettait de se doter de bases communes en IA, et d'élargir leurs compétences pratiques.

Au programme il y avait huit présentations qui couvriraient différents aspects des métiers de l'intelligence artificielle : l'introduction à l'apprentissage profond (DL) et Pytorch[14], le traitement automatique de la langue naturelle (NLP), la reproductibilité en apprentissage automatique, la gestion des résultats mais aussi des tutoriels pour l'utilisation de Calcul Canada[15], pour la vision par ordinateur et pour l'apprentissage par renforcement (RL) et pour finir la promotion d'une IA responsable et éthique.

- **Formation sur la vision par ordinateur et le traitement de l'image**

La photographie algorithmique[16] est un domaine qui se situe entre la vision artificielle, l'infographie, et la photographie. Elle utilise des techniques de calcul permettant de compenser les limites des appareils photo traditionnels et améliorer la façon dont on capture et manipule les données visuelles ainsi que la façon dont on interagit avec elles. Le cours permet d'explorer la formation des images, les modèles d'appareils, la création de panoramas, la synthèse de texture, les images à haute plage dynamique (HDR), le redimensionnement intelligent, la métamorphose, etc. Il s'oriente sur l'application concrète des concepts par de nombreux travaux pratiques, ainsi qu'un projet personnel.

L'étudiante a dû visionner les capsules pour les modules sur les pixels, le filtrage, la haute plage dynamique et l'apprentissage profond mais aussi faire les deux travaux pratiques du cours, un sur l'alignement automatique et un sur le filtrage. Elle a dû également faire la lecture des chapitres sur la formation d'images, le traitement d'images et l'apprentissage profond du livre *Computer Vision : Algorithms and Applications*[17] de Richard Szeliski.

La stagiaire réalisa deux travaux pratiques dans ce cours. Pour le premier (a), elle généra automatiquement une image couleur à partir des plaques de verre numérisées de la collection Prokudin-Gorskii[18]. Pour ce faire, elle dû extraire les trois canaux de couleurs, les chevaucher l'un par-dessus l'autre puis les aligner pour que leur combinaison forme une image couleur en RGB. Un algorithme fut créé pour une seule échelle et un autre pour des échelles multiples. Pour le second (b), celle-ci réalisa un travail sur la manipulation d'images dans le domaine des fréquences. Une partie était sur l'accentuation des détails d'images en utilisant la technique de l'accentuation (sharpening), l'autre partie était sur la génération d'images hybrides, dont l'interprétation dépend de la distance entre l'image et l'observateur. La technique utilisée fut celle de l'article de SIGGRAPH[19] écrit par Oliva, Torralba et Schyns en 2006. Cet effet a été réalisé à l'aide des fréquences de deux images superposées. Une image accentue les hautes fréquences donc perceptibles de proche et l'autre accentue les basses fréquences perceptibles de loin.

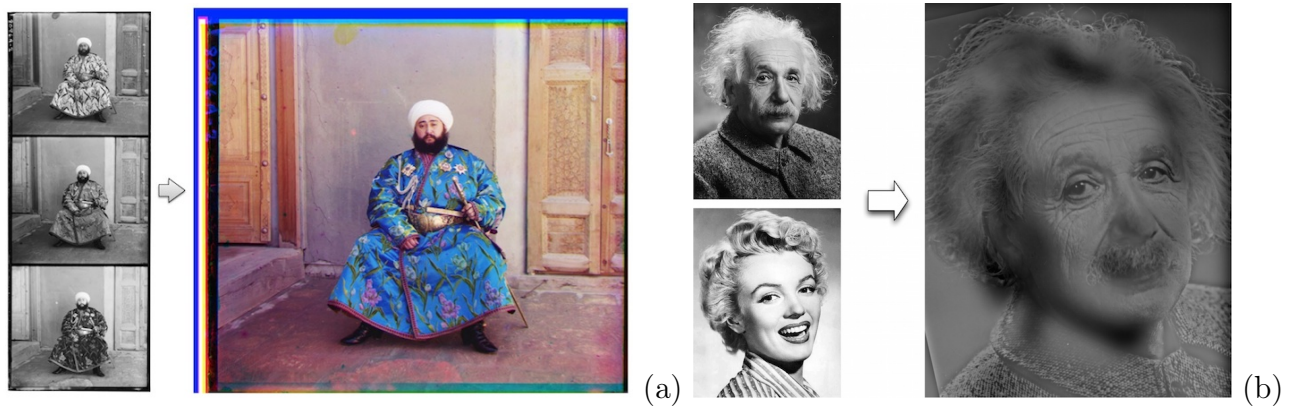


FIGURE 1 – Génération d'images en couleur et génération d'images hybrides

- **Formation sur l'apprentissage machine et apprentissage profond**

Une formation en apprentissage machine sur le site web Kaggle[20] a été suivie par la stagiaire. Celle-ci portait sur les idées de base en apprentissage automatique, la création et la validation de modèles, l'exploration et le traitement des données mais aussi la pipeline. L'étudiante a lu le chapitre concernant l'apprentissage profond du livre de vision par ordinateur de Richard Szeliski.

- **Formation sur les bibliothèques d'apprentissage profond en Python**

Une formation avec les bibliothèques TensorFlow[21] et Keras[22] a été suivie afin d'élargir ses connaissances dans ce type de bibliothèques. Elle lui a appris à créer et former des réseaux de neurones pour les données structurées, à mieux comprendre le sur et sous ajustement des données, la descente de gradient stochastique, l'abandon et la normalisation des lots mais aussi le classement binaire. Quant à la formation sur la bibliothèque Pytorch a été suivie puisqu'elle est utilisée par le laboratoire de recherche. Elle a pu y apprendre les bases, les tenseurs et la construction de modèles à l'aide de notebooks et de vidéos. La manipulation des images et des vidéos avec la bibliothèque TorchVision de PyTorch fut également apprise à

l'aide de tutoriels. Les méthodes d'apprentissage apprises durant cette formation étaient le réglage fin de la détection d'objets (Fine Tuning), l'apprentissage par transfert (Transfert Learning) ainsi que la génération d'exemples contradictoires avec un réseau antagoniste génératif (GAN) et la génération de nouvelles célébrités après lui avoir montré des photos de nombreuses célébrités réelles avec un réseau antagoniste génératif à convolution profonde (DCGAN).



FIGURE 2 – Tutorial DCGAN avec Pytorch

3.6.2 Phase : Projet sur la lumière HDR dans les panoramas

- **Revue de littérature sur la segmentation sémantique**

Dans cette deuxième partie du stage, l'étudiante a dû lire plusieurs articles scientifiques et articles de blogs dans le domaine de la vision par ordinateur et plus précisément sur la tâche de segmentation sémantique afin de mieux comprendre les attentes du projet. Pour se faire, il a été nécessaire de se renseigner sur les types de segmentation sémantique et les techniques utilisées afin d'y parvenir. La segmentation sémantique étant une tâche de prédiction dense, c'est-à-dire que l'étiquette de chaque pixel est prédite dans l'image en sortie, l'usage de l'apprentissage en profondeur fut donc nécessaire. L'architecture des réseaux de neurones en U est couramment utilisée pour ce genre de problème.

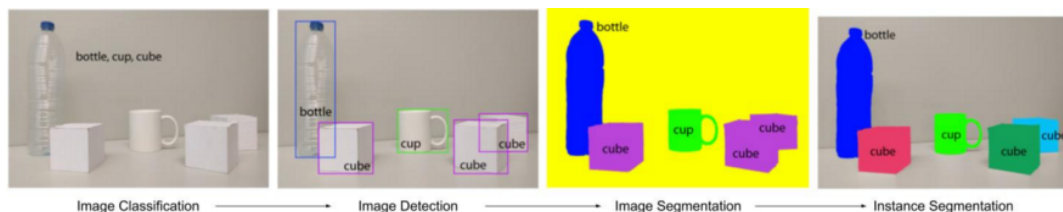


FIGURE 3 – Les types de segmentation sémantique

- **Recherche d’algorithmes pour la segmentation sémantique**

Les différents réseaux de neurones pour le traitement des images ont été étudiés comme les Réseaux de Neurones Convolutifs (CNN) et ses variantes pour la segmentation, les R-CNN, Fast R-CNN, Mask R-CNN et Mesh R-CNN mais aussi les modèles Deeplab et HoHoNet.

À la suite de ça, deux modèles ont été retenus : Deeplab[23] qui est actuellement l’état de l’art pour la segmentation sémantique d’instance et le modèle HoHoNet[24] qui traite également les images de panoramas 360 degrés. La force du modèle Deeplab de Google est pour ses performances au niveau de la segmentation sur les images panoramiques tandis que celle pour le modèle HoHoNet est sur l’estimation de profondeur en plus de faire la segmentation sémantique sur les images panoramiques à haute résolution.

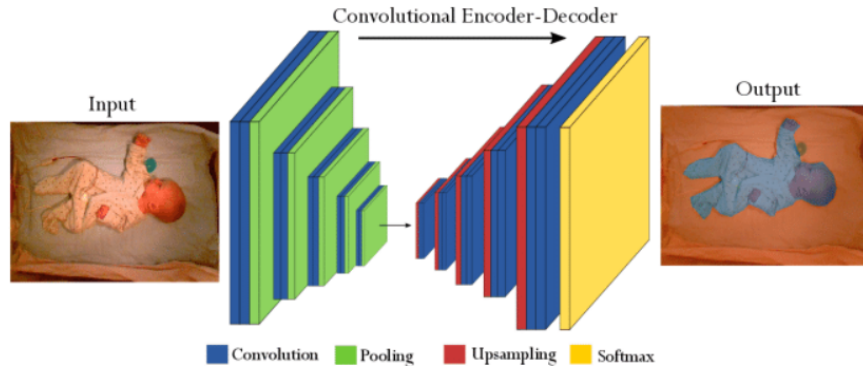


FIGURE 4 – Architecture d’un réseau pour la segmentation sémantique

- **Collecter des données nécessaires à l’exécution de ces algorithmes**

Pour commencer, le jeu de données de COCO Détection[25] fut téléchargé afin de tester le modèle Deeplab puis éventuellement celui de Matterport3D. Pour se faire, il a fallu se connecter au serveur du laboratoire et télécharger localement les données nécessaires à la tâche de segmentation puisque le jeu de données au complet était bien trop gros (1 TB). Le jeu de données de Stanford2D3D[26] est celui utilisé par le modèle HoHoNet. Éventuellement les données panoramiques du laboratoire LVSN ont été téléchargées afin de produire les résultats de segmentation et d’estimation de profondeur et de position.

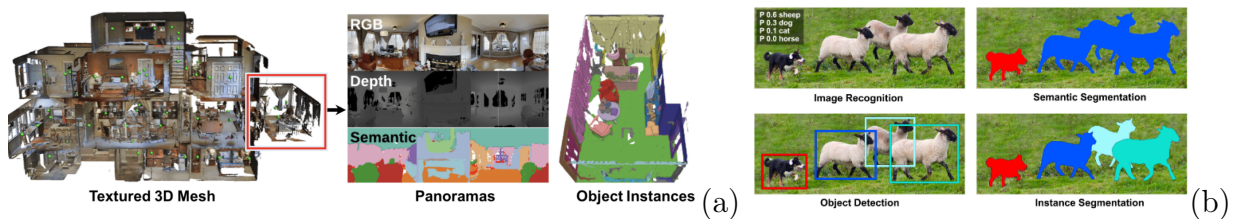


FIGURE 5 – Jeu de données de Matterport3D et COCO Détection

- **Tester des algorithmes pour la segmentation sémantique**

Tout d’abord, l’étudiante testa le modèle Deeplab de Pytorch sur quelques images, puis sur le jeu de données COCO Détection. Pour se faire elle dû installer l’application de COCO afin de pouvoir lire les images avec leur annotations et les manipuler.

Par la suite, elle essaya de tester le jeu de données de Matterport3D mais la génération des masques pour chaque image ne put fonctionner. Elle rencontra quelques problèmes, essaya de les résoudre, de changer de stratégie pour finalement réaliser qu’il n’était pas possible de remplir cette tâche. Le manque de support pour Matterport3D et le langage obsolète utilisé (Lua[27]) par l’application ne permettait pas de réaliser en temps cette tâche.

N’ayant plus beaucoup de temps pour finir le mandat, il fut décidé de changer de jeu de données et de modèle afin de réaliser les tests sur les données du laboratoire. Le modèle HoHoNet avec le jeu de données Stanford2D3D fut sélectionné. Le modèle fut testé sur un échantillon de 50 images panoramiques du laboratoire LVSN puis finalement sur le jeu de données au complet, c’est-à-dire sur 2280 images panoramiques. Trois résultats pour chaque images furent produits, la segmentation sémantique, l’estimation de profondeur et celle de la position.

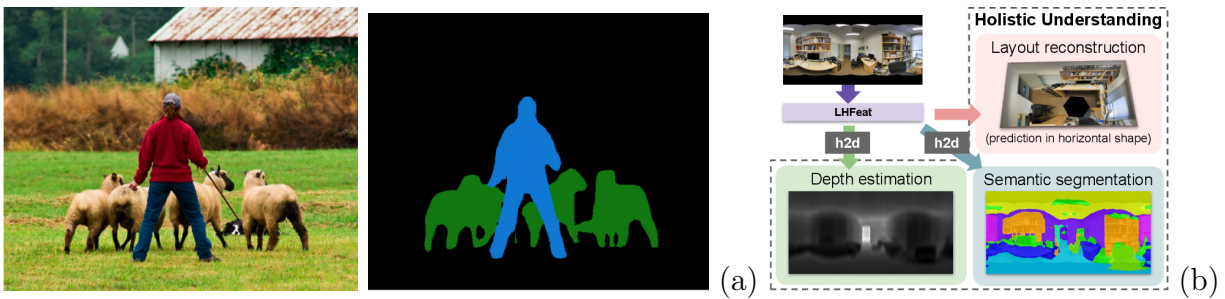


FIGURE 6 – Modèle DeeplabV3 de Pytorch et le modèle HoHoNet

- **Organiser et présenter les résultats des expérimentations**

Chaque semaine, la stagiaire présentait les résultats des expérimentations à son superviseur et son étudiant à la maîtrise ce qui permettait par la suite de réajuster les expérimentations.

- **Assister un étudiant gradué dans ses travaux de recherche**

Les travaux de recherche de l’étudiant gradué consistait à capturer le plus d’information durant la prise de photos panoramiques afin d’entraîner un modèle à détecter l’information manquante.

3.6.3 Groupe de lecture les vendredis

En parallèle aux travaux, pour se tenir à jour dans les travaux scientifiques du domaine de la vision par ordinateur mais plus précisément au niveau des photographies numériques, l'étude d'un article scientifique soigneusement sélectionné était faite chaque vendredi après-midi avec tous les membres sous la direction du professeur Jean-François Lalonde. Le matin, Isabelle prenait le temps d'étudier l'article en question afin de pouvoir participer à la discussion faite sur celui-ci. Pour finir, un vote était fournis afin de critiquer cet article.

La liste des articles étudiés durant le stage :

- 13 mai : *Hierarchical Text-Conditional Image Generation with CLIP Latents*[28]
- 20 mai : *Gauge Equivariant Convolutional Networks and the Icosahedral CNN*[29]
- 27 mai : *Learning to Factorize and Relight a City*[30]
- 03 juin : *Block-NeRF : Scalable Large Scene Neural View Synthesis*[31]
- 10 juin : *Spatial-Temporal Transformer for Dynamic Scene Graph Generation*[32]
- 17 juin : *Vision Transformer with Deformable Attention*[33]
- 24 juin : *Denoising Diffusion Probabilistic Models*[34]
- 01 juillet : *Instant Neural Graphics Primitives with a Multiresolution Hash Encoding*[35]
- 08 juillet : *Scene Representation Transformer : Geometry-Free Novel View Synthesis Through...*[36]
- 15 juillet : *Partial success in closing the gap between human and machine vision*[37]
- 22 juillet : *Diffusion Models Beat GANs on Image Synthesis*[38]
- 29 juillet : *Deep Unsupervised Learning using Nonequilibrium Thermodynamics*[39]
- 05 aout : *Paint By Word*[40]
- 12 aout : *Dynamic Routing Between Capsules*[41]
- 19 aout : *Leaning Invariances in Neural Networks*[42]
- 26 aout : *Learning Transferable Visual Models From Natural Language Supervision*[43]

3.7 Résultats

Résultats et analyse : Finalement l'étudiante réussit à trouver, tester et segmenter l'ensemble de données de panoramas du laboratoire de recherche LVSN. Ses résultats vont permettre de pouvoir décrire le contenu des scènes du jeu de données et comprendre l'impact de chaque élément sur l'aspect photopique et mélanopique de la lumière et éventuellement créer de meilleurs espaces de vie pour les gens qui travaillent et habitent dans le grand nord. Ultimement, le jeu de donnée calibré permettra d'apprendre à prédire ces mesures à partir d'une seule image.

Discussion : Pour la phase de formation, l'étudiante réalisa que celle-ci aurait pu être plus courte ou suivie en parallèle du projet afin de ne pas manquer de temps sur le projet. Pour la phase du projet, celle-ci réalisa qu'à l'avenir elle prendrait le temps de mieux comprendre l'ensemble de données et leurs annotations afin d'éviter de perdre du temps à le faire fonctionner sur un modèle compte tenu du temps limité pour remplir le mandat. Une fois le changement de modèle fait, les tests ont été beaucoup plus rapides et le mandat pu être réalisé dans les temps.

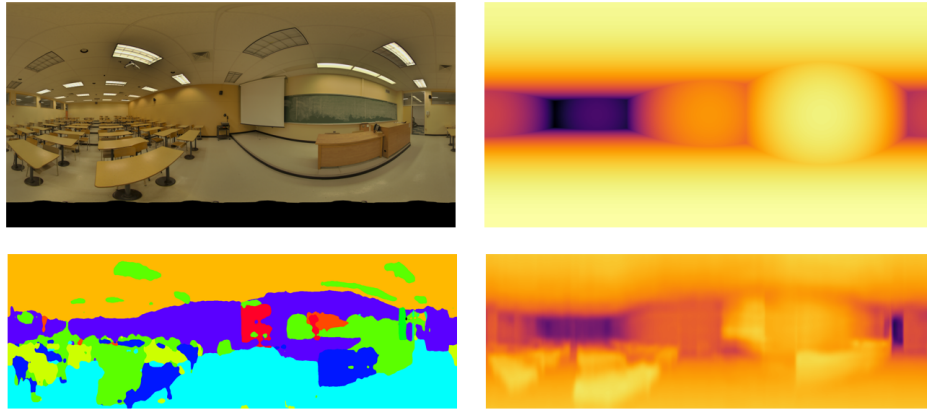


FIGURE 7 – Image originale 9C4A9529 avec les trois résultats

3.8 Comparaison

Au début du stage, l'étudiante espérait avoir le temps de faire le projet dans son ensemble afin d'avoir une idée du métier de la recherche dans le domaine de la vision numérique sur les photos et pouvoir se décider pour continuer ses études à la maîtrise ou bien rentrer sur le marché du travail. Celle-ci avait un intérêt pour poursuivre ses études au niveau supérieur. Ce second stage lui permet de confirmer son fort désir pour le milieu de la recherche et son profond intérêt pour le domaine de la vision numérique.

Comparativement au premier stage, l'étudiante pu réaliser un projet sur la vision numérique avec les étapes nécessaires pour y parvenir. Dans un projet scolaire, les étudiants n'ont pas l'occasion ni le temps de tester plusieurs modèles sur plusieurs jeu de données et la phase de recherche ne peut pas être aussi poussée par manque de temps.

Après ces deux stages dans le milieu de la recherche, l'étudiante se sent confiante pour rentrer sur le marché du travail dans le milieu de la recherche et dans le domaine de la vision numérique. Ils lui ont permis de parfaire ses connaissances autant théorique que pratique. Celle-ci prendra le temps de développer son expertise dans ce domaine pour éventuellement faire une maîtrise dans quelques années.

4 Développement et renforcement des compétences

4.1 Techniques

Le stagiaire a pu développer ses habiletés avec plusieurs technologies. Certaines d'entre elles étaient nouvelles comme la communication avec des serveurs SSH, le langage de programmation Lua et quelques logiciels pour transformer les données. Tandis que d'autres ont pu être améliorées comme les commandes Linux, la programmation en Python et l'apprentissage profond.

4.2 Scientifiques

Le stagiaire a pu également renforcer ses habiletés de recherche, de lecture, d'analyse, de synthèse et de pensées critiques mais aussi la gestion de projet.

4.3 Communication

Des compétences en communication ont été requises afin de rendre compte des avancements du stagiaire chaque semaine avec son co-superviseur. Elle a dû partager ses résultats de recherches, d'analyse puis des tests devant son superviseur et l'étudiant à la maîtrise. Des interactions avec différents membres du laboratoire ont été nécessaires afin de faire avancer le projet. Elle a également pu pratiquer son anglais avec les groupes de lecture, les communication sur le Slack du laboratoire.

4.4 Réflexion sur la formation pratique

La stagiaire étant intéressée dans le domaine de la recherche, l'apprentissage machine et la vision numérique, a effectué ce stage pour en apprendre davantage. Il lui a permis d'en apprendre beaucoup plus sur la segmentation sémantique et l'apprentissage profond, un domaine qui l'intéresse. Après cette expérience enrichissante, cette dernière a confirmé son intérêt dans la recherche et continuera éventuellement au niveau de la maîtrise.

4.5 Réflexion sur la formation théorique

Être sous la supervision d'un professeur a de grands avantages, car celui-ci a pu transférer son expertise à la stagiaire. Celle-ci a eu l'occasion de voir la théorie sur la vision numérique et la photographie dans le cours de photographie algorithmique (GIF-4105) vu en accéléré avec son superviseur afin de se mettre à niveau dans le domaine de la vision numérique appliqué aux photographies. Durant son parcours universitaire, la stagiaire a eu l'occasion de voir la théorie sur l'apprentissage profond dans le cours d'apprentissage par réseaux de neurones profonds (GLO-4030) ce qui lui a donné une base de connaissance et lui a permis une meilleure compréhension lors de son stage.

4.6 Bilan des acquis

Au début du stage, l'étudiante espérait développer des compétences sur le milieu de la recherche mais aussi renforcer ses compétences avec l'apprentissage profond et la vision numérique. Elle a eu le temps de développer ses compétences dans le domaine de l'informatique mais aussi dans le domaine de la vision numérique. En peu de temps, elle a pu acquérir de nombreuses connaissances dont certaines notions encore inconnues jusque-là. Dans le domaine de la vision numérique, elle acquit des connaissances en segmentation sémantique. Elle a pu également acquérir de nouvelles connaissances professionnelles, se faire connaître et apprendre à intégrer une équipe de travail bilingue.

5 Conclusion

Durant ce second stage en recherche en informatique, celui-ci a été l'occasion de mettre en pratique les différentes connaissances théoriques en intelligence artificielle acquises durant son parcours académique.

Avec ce mandat, celle-ci a pu se familiariser avec différentes méthodes et outils utilisées en segmentation sémantique et en apprentissage en profondeur dans le domaine de la vision par ordinateur qu'elle pourra réutiliser dans son parcours professionnel.

Les compétences acquises durant ce stage et le précédent permettront au stagiaire de travailler dans le domaine de la vision numérique.

Ce stage lui a permis de perfectionner ses connaissances théoriques et pratiques tout en favorisant l'intégration de ses apprentissages. Il lui permet également de développer une attitude professionnelle au sein d'un laboratoire de recherche.

6 Référence

- [1] Site web du laboratoire de Vision et Systèmes Numériques de l'Université Laval, LSVN, <http://vision.gel.ulaval.ca/>
- [2] Projet Sentinelle Nord sur la lumière HDR, <https://sentinellenord.ulaval.ca/projets-de-recherche/design-biophilique-dans-larctique-co-creation-communautaire/>
- [3] Liste des membres du LSVN, <http://vision.gel.ulaval.ca/fr/people/index.php>
- [4] Site web de l'Institut Intelligence des Données, IID, <https://iid.ulaval.ca/>
- [5] Liste des étudiants de Jean-François Lalonde, <http://vision.gel.ulaval.ca/~jflalonde/students/>
- [6] Site web de Jean-François Lalonde, <http://vision.gel.ulaval.ca/~jflalonde/>
- [7] Matterport3D : apprendre des données RVB-D dans les environnements intérieurs, <https://niessner.github.io/Matterport/>
- [8] Regroupement pour l'étude des environnements partagés intelligents répartis (REPARTI), <https://frq.gouv.qc.ca/regroupement/reparti-systemes-cyberphysiques-et-intelligence-machine-materialisee/>
- [9] Le Fonds de recherche du Québec – Nature et technologies (FRQNT), <https://frq.gouv.qc.ca/nature-et-technologies/mission-et-domaines-de-recherche/>
- [10] Site web de la plateforme de communication collaborative Slack, <https://slack.com/intl/fr-ca/>
- [11] Site web du logiciel de visioconférence Zoom, <https://zoom.us/>
- [12] Site web du logiciel de présentation PowerPoint de Microsoft, <https://office.live.com/start/powerpoint.aspx>
- [13] Formation Bootcamp 2022 de IID, <https://iid.ulaval.ca/evenements/formation-bootcamp-iid-2022/>
- [14] PyTorch est une bibliothèque logicielle Python open source d'apprentissage machine, <https://pytorch.org/>
- [15] Site web de Calcul Canada, <https://www.computecanada.ca/>
- [16] Site du cours GIF-4105 photographie algorithmique de Jean-François Lalonde de l'hiver 2021, <http://wcours.gel.ulaval.ca/GIF4105/index-modules.html>

- [17] Livre *Computer Vision : Algorithms and Applications* de Richard Szeliski, Seconde Édition Springer de 2022, <https://szeliski.org/Book>
- [18] Travail pratique sur la génération d'images couleur d'après la méthode de Sergei Mikhailovich Prokudin-Gorskii datant de 1907, <http://wcours.gel.ulaval.ca/GIF4105/tps/tp1/index.html>
- [19] Article sur la création d'images hybrides, *Hybrid images* écrit par Oliva, Torralba et Schyns en 2006, https://stanford.edu/class/ee367/reading/OlivaTorralba_Hybrid_Siggraph06.pdf
- [20] Kaggle est une plateforme web organisant des compétitions en science des données appartenant à Google, <https://www.kaggle.com/>
- [21] TensorFlow est un outil open source d'apprentissage automatique développé par Google, <https://www.tensorflow.org/?hl=fr>
- [22] Keras permet une expérimentation rapide avec les réseaux de neurones profonds de Tensorflow, <https://keras.io/>
- [23] DeepLab est un model de segmentation sémantique d'instance, <https://arxiv.org/abs/1412.7062v4>
- [24] HoHoNet (ResNet-101) est un model pour la compréhension holistique d'intérieure, <https://arxiv.org/pdf/2011.11498v3.pdf>
- [25] Base de données MS COCO Detection , <https://paperswithcode.com/dataset/coco>
- [26] Ensemble de données sémantiques 2D-3D de Stanford (2D-3D-S), <http://buildingparser.stanford.edu/dataset.html>
- [27] Lua est un langage de script libre, réflexif et impératif, <https://www.lua.org/>
- [28] Article *Hierarchical Text-Conditional Image Generation with CLIP*, <https://arxiv.org/abs/2204.06125>
- [29] Article *Gauge Equivariant Convolutional Networks and the Icosahedral CNN*, <https://arxiv.org/abs/1902.04615>
- [30] Article *Learning to Factorize and Relight a City*, <https://arxiv.org/abs/2008.02796>
- [31] Article *Block-NeRF : Scalable Large Scene Neural View Synthesis*, <https://arxiv.org/abs/2202.05263>
- [32] Article *Spatial-Temporal Transformer for Dynamic Scene Graph Generation*, <https://arxiv.org/abs/2107.12309>
- [33] Article *Vision Transformer with Deformable Attention*, <https://arxiv.org/abs/2201.00520>

- [34] Article *Denoising Diffusion Probabilistic Models*, <https://arxiv.org/abs/2006.11239>
- [35] Article *Instant Neural Graphics Primitives with a Multiresolution Hash Encoding*, <https://nvlabs.github.io/instant-ngp/>
- [36] Article *Scene Representation Transformer : Geometry-Free Novel View Synthesis Through ...*, <https://arxiv.org/abs/2111.13152>
- [37] Article *Partial success in closing the gap between human and machine vision*, <https://arxiv.org/abs/2106.07411>
- [38] Article *Diffusion Models Beat GANs on Image Synthesis*, <https://arxiv.org/abs/2105.05233>
- [39] Article *Deep Unsupervised Learning using Nonequilibrium Thermodynamics*, <https://arxiv.org/abs/1503.03585>
- [40] Article *Paint By Word*, <https://arxiv.org/abs/2103.10951>
- [41] Article *Dynamic Routing Between Capsules*, <https://arxiv.org/abs/1710.09829>
- [42] Article *Leaning Invariances in Neural Networks*, <https://arxiv.org/abs/2010.11882>
- [43] Article *Learning Transferable Visual Models From Natural Language Supervision*, <https://arxiv.org/abs/2103.00020>