 UNIVERSITÉ DE SHERBROOKE	PROJET UdeS MELCC	DOCUMENT NO. 20210930	DATE 2021-09-30	VERSION 3
	TITRE Système collaboratif d'aide à l'annotation des espèces de végétation par apprentissage actif	PAGE 1		PAGES 8

Plan de travail


**Université de Sherbrooke
Département de géomatique appliquée**

**Système collaboratif d'aide à l'annotation des espèces de végétation
par apprentissage actif**

Suivi des réalisations et des ajouts demandés (septembre 2021)

**Professeur
Mickaël Germain**

**Date
30 septembre 2021**

 UNIVERSITÉ DE SHERBROOKE	PROJET UdeS MELCC	DOCUMENT NO. 20210930	DATE 2021-09-30	VERSION 3
	TITRE Système collaboratif d'aide à l'annotation des espèces de végétation par apprentissage actif	PAGE 2		PAGES 8

Suivi des changements

*A – Ajouté M – Modifié S – Supprimé

NUMÉRO DE VERSION	DATE aaaa/mm/jj	INFORMATION	A* M S	BRÈVE DESCRIPTION DU CHANGEMENT	NUMÉRO DE DEMANDE CHANGEMENT
1.0	2021/05/10	Proposition initiale			
2.0	2021/05/14	Modification	M	Correction	
3.0	2021/09/30	Mise à jour			

Contact - Université de Sherbrooke

Mickaël Germain, PhD
 Professeur adjoint
 Département de géomatique appliquée
 Université de Sherbrooke
 Tél. : 1 (819) 821-8000 poste 65687
 Courriel : mickael.germain@usherbrooke.ca



 UNIVERSITÉ DE SHERBROOKE	PROJET UdeS MELCC	DOCUMENT NO. 20210930	DATE 2021-09-30	VERSION 3
	TITRE Système collaboratif d'aide à l'annotation des espèces de végétation par apprentissage actif	PAGE 3		PAGES 8

TABLE DES MATIÈRES

1.	Mise en contexte	4
2.	Plan de travail.....	4
2.1	Étape n° 1 – installation du logiciel et des outils de développement.....	4
2.2	Étape n° 2 – test sur des images de microdrones	4
2.3	Étape n° 3 – définition d'un modèle d'apprentissage profond.....	5
2.4	Étape n° 4 – développement d'un système multi-utilisateur	5
2.5	Étape n° 5 – amélioration de l'interface des modèles IA et entraînement des modèles	5
2.6	Étape n° 6 – amélioration des interfaces d'annotations.....	5
2.7	Étape n° 7 – système de gestion des données géospatiales et interopérabilité du système.....	6
2.8	Étape n° 8 – personnalisation du système et test avec les partenaires	6
3.	Calendrier prévisionnel des travaux (mai 2021)	7
4.	Calendrier de la mise à jour des travaux (septembre 2021)	7

 UNIVERSITÉ DE SHERBROOKE	PROJET UdeS MELCC	DOCUMENT NO. 20210930	DATE 2021-09-30	VERSION 3
	TITRE Système collaboratif d'aide à l'annotation des espèces de végétation par apprentissage actif	PAGE 4		PAGES 8

1. MISE EN CONTEXTE

Le document présente un plan de travail détaillé du développement d'un système collaboratif d'aide à l'annotation des espèces de végétation. Le projet vise à utiliser intelligence artificielle (IA) sur des données de microdrones, et plus particulièrement par l'utilisation d'un protocole d'annotation des jeux de données appelé « apprentissage actif » afin de sélectionner un jeu de données pour l'apprentissage le plus complet et le plus efficace possible, tout en réduisant la dimensionnalité de ce dernier. Le modèle IA généré sera optimisé de façon dynamique avec l'ajout des nouvelles annotations.

Le système proposé s'appuiera sur les bases d'un système *Open Source* déjà élaboré pour l'annotation des espèces animales par apprentissage actif : AIDE (*Accelerating image-bases ecological surveys interactive machine learning*) développé par Benjamin Kellenger avec des fonds du programme de *Microsoft AI for earth*. L'outil est une plateforme web collaborative qui en est à ses premiers développements, et qui nécessite des améliorations pour une utilisation dans le cadre de ce projet.

2. PLAN DE TRAVAIL

Le plan de travail reprend les éléments de la proposition sous la forme d'étapes de réalisation et d'un échéancier prévisionnel.

2.1 Étape n° 1 – installation du logiciel et des outils de développement


Le système AIDE actuel est disponible sur un dépôt de logiciel (Github) pour le téléchargement et l'installation sur un serveur. Il est ouvert et libre (*Open Source*) avec une documentation sommaire pour son utilisation.

La première étape est d'installer le logiciel sur un serveur de Calcul Canada de l'Université de Sherbrooke, ainsi que les outils de développement pour l'amélioration du système dans le cadre du projet. Un professionnel de recherche travaillera sur le projet, et installera l'ensemble des outils pour modifier et améliorer les fonctionnalités du système. Une copie du système pourra être installée sur un ordinateur local pour faciliter les tests et le développement.

Étape terminée.

2.2 Étape n° 2 – test sur des images de microdrones

Une équipe du Pr Étienne Laliberté de l'Université de Montréal travaille sur l'acquisition d'images par microdrones qui seront utilisées dans le système. Cette étape vise à intégrer un ensemble d'images dans le système et à proposer une procédure d'utilisation pour l'équipe de l'UdeM. Ce premier test d'utilisation devra vérifier le fonctionnement de l'interface d'annotations web, et de définir les limitations actuelles qui devront être améliorées.

 UNIVERSITÉ DE SHERBROOKE	PROJET UdeS MELCC	DOCUMENT NO. 20210930	DATE 2021-09-30	VERSION 3
	TITRE Système collaboratif d'aide à l'annotation des espèces de végétation par apprentissage actif	PAGE 5		PAGES 8

2.3 Étape n° 3 – définition d'un modèle d'apprentissage profond

Les annotations définies par l'interface web seront utilisées pour alimenter un modèle d'apprentissage profond afin d'améliorer la classification des données sur les images. Une revue de littérature sera proposée pour l'utilisation des modèles. Actuellement, le choix d'un modèle d'apprentissage profond est limité (ResNet pour la classification, RetinaNet pour la détection d'objets et U-Net pour la segmentation). D'autres modèles seront proposés.

L'architecture pour la création/utilisation d'un modèle d'apprentissage profond est basée sur une architecture de classification multiétiquettes. C'est un changement par rapport à l'idée initiale de travailler avec l'interface d'annotation originale en mode « segmentation ».

2.4 Étape n° 4 – développement d'un système multi-utilisateur

Le système de base actuel est limité à un utilisateur et ne facilite pas la collaboration active pour l'amélioration des annotations. L'interface web devra intégrer plusieurs collaborateurs pour l'aide à l'annotation. Une solution simple et efficace sera proposée. Elle sera testée avec plusieurs utilisateurs sur des images de microdrones.

Étape terminée.

2.5 Étape n° 5 – amélioration de l'interface des modèles IA et entraînement des modèles

L'interface du choix des modèles et la gestion des paramètres n'est pas conviviale : une information sous la forme d'un fichier texte; et certaines fonctionnalités ne sont pas disponibles. De plus, l'entraînement des modèles avec les nouvelles annotations n'est pas configuré.

Une étude sera réalisée sur la vitesse de l'entraînement et de l'inférence sur les serveurs de Calcul Canada. L'interface devra proposer l'activation du GPU (*Graphic Processing Unit*) pour accélérer l'entraînement des modèles.


La configuration d'un serveur avec un GPU est en cours.

2.6 Étape n° 6 – amélioration des interfaces d'annotations

Les outils d'annotations sont les outils principaux pour classer, délimiter ou segmenter les données sur les images. Une attention particulière sera mise sur l'amélioration des outils actuels qui sont simples et pas assez précis pour les annotations du projet. Par exemple, un outil pour définir les polygones devra être le plus précis possible dans un contexte de segmentation.

Modification majeure de cette étape de travail. Une nouvelle interface d'annotations a été créée pour prendre en compte l'ensemble de l'image et les annotations multiétiquettes sous la forme de tuiles (image) pour faciliter la mise en contexte des annotations. Une taxonomie complète du MELCC a été intégrée dans l'interface web.

Une nouvelle demande du MELCC : gérer l'historique des étiquettes dans l'interface d'annotation.

 UNIVERSITÉ DE SHERBROOKE	PROJET UdeS MELCC	DOCUMENT NO. 20210930	DATE 2021-09-30	VERSION 3
	TITRE Système collaboratif d'aide à l'annotation des espèces de végétation par apprentissage actif	PAGE 6		PAGES 8

2.7 Étape n° 7 – système de gestion des données géospatiales et interopérabilité du système

Les données géoréférencées utilisées devront être sauvegardées dans un système de gestion de bases de données géospatiales comme le système PostgreSQL/PostGIS ou structurées dans des répertoires du disque dur du serveur. Un serveur cartographique sera installé et configuré pour diffuser les données sur Internet dans les principaux services de l'OGC (*Open Geospatial Consortium*) : WMS (*Web Map Service*) et WFS (*Web Feature Service*). L'interopérabilité des données assurera l'affichage des informations géoréférencées dans les logiciels cartographiques comme QGIS et ArcMap.


Un travail est en cours pour la nouvelle demande MELCC pour intégrer la gestion de l'historique de la taxonomie (les changements de nom et les différents noms pour une même espèce)

Une nouvelle demande du MELCC : outil de récupération des annotations. Test en cours avec un service WFS.

2.8 Étape n° 8 – personnalisation du système et test avec les partenaires

Une personnalisation complète des interfaces web sera proposée avec la thématique de travail et les partenaires du projet (MELCC et l'UdeM). Une version complète sera mise à disposition pour le MELCC pour les tests avec un guide d'utilisation sous la forme de tutoriel vidéo.

Guide de démarrage préparé pour l'annotation. Création des tutoriels sous la forme de capsule vidéo après les commentaires et les tests.


 UNIVERSITÉ DE SHERBROOKE	PROJET UdeS MELCC	DOCUMENT NO. 20210930	DATE 2021-09-30	VERSION 3
	TITRE Système collaboratif d'aide à l'annotation des espèces de végétation par apprentissage actif		PAGE 7	PAGES 8

3. CALENDRIER PRÉVISIONNEL DES TRAVAUX (MAI 2021)

<u>Mois</u>		Mai				Juin					Juillet				Août			
<u>Semaines</u>		1	2	3	4	1	2	3	4	5	1	2	3	4	1	2	3	4
Étapes	Délai sem.																	
2.1	1.5																	
2.2	1.5																	
2.3	1																	
2.4	1																	
2.5	3																	
2.6	3																	
2.7	1																	
2.8	4																	
Nombre total de semaines	16																	

4. CALENDRIER DE LA MISE À JOUR DES TRAVAUX (SEPTEMBRE 2021)

<u>Mois</u>		Mai				Juin					Juillet				Août			
<u>Semaines</u>		1	2	3	4	1	2	3	4	5	1	2	3	4	1	2	3	4
Étapes	Délai sem.																	
2.1	1.5																	
2.2	1.5																	
2.3	1																	
2.4	1																	
2.5	0																	
2.6	11																	
2.7	0																	
2.8	0																	
Nombre total de semaines	16																	

 UNIVERSITÉ DE SHERBROOKE	PROJET UdeS MELCC	DOCUMENT NO. 20210930	DATE 2021-09-30	VERSION 3
	TITRE Système collaboratif d'aide à l'annotation des espèces de végétation par apprentissage actif		PAGE 8	PAGES 8

		Mois septembre				Mois Octobre				
		Semaines				Semaines				
Étapes	Délai sem.	1	2	3	4	1	2	3	4	5
2.1	terminé									
2.2	terminé									
2.3	terminé									
2.4	terminé									
2.5	2									
2.6	6									
2.7	2									
2.8	2									