ANNEE ACADEMIQUE :2024/2025

**REPUBLIC OF CAMEROON**

**\*\*\*\*\*\*\***

**Peace-Work-Fatherland**

**\*\*\*\*\*\***

**UNIVERSITY OF YAOUNDE I**

**\*\*\*\*\*\***

**FACULTY OF SCIENCE**

**\*\*\*\*\*\***

**DEPARTMENT OF COMPUTER SCIENCE**

**\*\*\*\*\*\***

**DATA SCIENCE**

**\*\*\*\*\*\***

**MASTER 1**

INFO 4258 : UE PROJET-DESCRIPTION

**«APPLICATION POUR LA CLASSIFICATION DES INVERTEBRES MARRINS »**

**GROUPE**

**NOMS ET MATRICULES**

1. NOUBISSI FOPA CHRISTIAN JUNIOR 19M2315
2. ESSUTHI MBANGUE ANGE ARMEL 0000
3. MBELL BONO AURELIEN 0000

|  |
| --- |
|  |

Sous la supervision du PR TSOPZE NORBERT



**REPUBLIQUE DU CAMEROUN**

**\*\*\*\*\*\***

**Paix-Travail-Patrie**

**\*\*\*\*\*\***

**UNIVERSITE DE YAOUNDE I**

**\*\*\*\*\*\***

**FACULTE DES SCIENCES**

**\*\*\*\*\*\***

**DEPARTEMENT D’INFORMATIQUE**

**\*\*\*\*\***

**SCIENCES DE DONNEES**

**\*\*\*\*\***

**MASTER 1**

SOMMAIRE

1. CONTEXTE
2. OBJECTIF
3. PROBLEME A RESOUDRE
4. SOLUTIONS EXISTANTES
5. SOLUTION PROPOSEE
6. Idée de la solution proposée
7. Nature de la solution proposée
8. Méthodologie envisagée
9. Résultats envisagés
10. Algorithme de machine learning
11. Rentabilité de la solution
12. FONCTIONNALITEES
13. PLANNING DE REALISATION
14. EQUIPE PROJET
    1. **CONTEXTE**

La biodiversité marine, et particulièrement celle des invertébrés marins, représente un patrimoine naturel d'une richesse exceptionnelle mais souvent méconnue du grand public. Les invertébrés marins constituent plus de 98% des espèces animales marines connues et jouent un rôle crucial dans les écosystèmes marins. Cependant, leur identification sur le terrain reste un défi majeur pour :

* Les chercheurs et étudiants en biologie marine
* Les gestionnaires d'aires marines protégées
* Les plongeurs amateurs et professionnels
* Les acteurs de la conservation marine
* Le grand public

L'identification précise des invertébrés marins nécessite actuellement :

* Une expertise pointue en taxonomie
* L'accès à des guides spécialisés souvent volumineux et peu pratiques sur le terrain
* Du temps considérable pour la consultation de clés de détermination
* Une formation approfondie pour reconnaître les caractéristiques morphologiques distinctives

L'essor des technologies d'intelligence artificielle, particulièrement dans le domaine de la vision par ordinateur et du Deep Learning, offre de nouvelles perspectives pour :

* Automatiser et accélérer le processus d'identification des espèces
* Démocratiser l'accès à l'expertise taxonomique
* Contribuer à la science participative et à la collecte de données sur la biodiversité marine
* Sensibiliser le grand public à la richesse des écosystèmes marins
  1. **OBJECTIF**

L’objectif de ce projet est de développer une application utilisant des techniques vision par ordinateur pour :

1. Identifier rapidement et précisément les invertébrés marins à partir de photos
2. Fournir des informations détaillées sur les espèces identifiées
3. Contribuer à la collecte de données scientifiques sur la distribution des espèces
4. Servir d'outil pédagogique pour l'apprentissage de la biodiversité marine
   1. **PROBLEME A RESOUDRE**

L'objectif principal de ce projet est de développer une application mobile de classification des invertébrés marins capable de fonctionner hors-ligne, en se concentrant sur la distinction précise entre espèces morphologiquement similaires. De manière spécifique, les problèmes a résoudre sont :

**1. Prétraitement des Images**

* Incluant un processus de nettoyage des images
* Standardisation des données d'entrée pour le modèle d’IA
* Optimisation des images pour l'analyse tout en préservant les détails distinctifs

**2. Classification des Espèces**

* Implémentation d'un modèle de classification précis pour les invertébrés marins
* Développement d'algorithmes capables de distinguer les espèces morphologiquement similaires

**3. Optimisation des Performances**

* Réduction du temps de traitement pour une utilisation mobile
* Optimisation de la consommation des ressources (CPU, mémoire, batterie)
* Adaptation du modèle pour un fonctionnement hors-ligne efficace

**4. Interface Utilisateur**

* Conception d'une interface intuitive et ergonomique
* Développement d'une navigation fluide dans l'application
* Implémentation d'un affichage clair des résultats de classification
* Optimisation des temps de réponse de l'interface
  1. **SOLUTIONS EXISTANTES**

Des solutions dans cette thématique ont déjà été proposées comme :

* **AquaInvertebrates** : Application mobile disposant d'une base de données riche et d'un système de filtres taxonomiques offrant des fiches descriptives détaillées des espèces. Malheureusement, elle ne se limite qu'aux espèces de la Méditerranée.
* **OceanID** : Plateforme utilisant un modèle d'IA assez performant, mais fonctionnant uniquement en ligne et disposant d'une base de données spécialisée sur les vertébrés.
* **Seek by iNaturalist** : Variante de la première fonctionnant en mode hors-ligne. Elle permet une identification rapide mais gère mal les espèces similaires. De plus, elle renferme une base de données assez limitée.
* **CoralNet** : Fournit d'assez bons résultats mais présente une interface complexe pour les non-experts.
* **Outils scientifiques** : Le World Register of Marine Species (WoRMS) ou la Marine Species Identification Portal qui se comportent plus comme des bases de données et non des applications.
* **iNaturalist** : Est une plateforme en ligne qui permet aux utilisateurs de partager des observations de la biodiversité, facilitant ainsi l'identification des espèces et la collecte de données sur la faune et la flore. Son architecture technique repose sur plusieurs composants clés, et elle intègre des modèles d'apprentissage automatique pour améliorer l'expérience utilisateur. Offre une large base de données d'espèces avec une communauté active et une validation collaborative des identifications. Cependant, son utilisation nécessite une connexion internet avec un temps de traitement relativement long pour une précision limitée sur des espèces spécifiques.

**1. Architecture globale de l'application :**

* **Interface utilisateur (UI) :** iNaturalist propose des applications mobiles et un site web permettant aux utilisateurs de télécharger des observations, de les identifier et de les commenter.
* **Serveur backend :** Le backend gère les requêtes des utilisateurs, traite les données et communique avec la base de données et les services d'apprentissage automatique.
* **Base de données :** Elle stocke les observations, les identifications, les commentaires et les informations sur les utilisateurs.
* **Services d'apprentissage automatique :** Ces services fournissent des suggestions d'identification basées sur les images téléchargées par les utilisateurs.

**2. Technologies utilisées :**

* **Frontend :** Les applications mobiles sont développées pour Android et iOS, tandis que le site web utilise des technologies web modernes pour une interface réactive.
* **Backend :** Le serveur est construit avec des frameworks web robustes, assurant une communication efficace entre le frontend et la base de données.
* **Base de données :** Un système de gestion de base de données relationnelle est utilisé pour stocker et organiser les données des utilisateurs et des observations.
* **Apprentissage automatique :** iNaturalist utilise des bibliothèques et frameworks d'apprentissage automatique, tels que TensorFlow, pour entraîner et déployer ses modèles de vision par ordinateur.

**3. Méthodologies :**

* **Collecte de données :** Les utilisateurs téléchargent des photos et des descriptions d'observations, fournissant ainsi des données étiquetées pour l'entraînement des modèles.
* **Prétraitement :** Les images sont normalisées et les métadonnées sont structurées pour assurer la qualité des données.
* **Entraînement du modèle :** Les modèles de vision par ordinateur sont entraînés sur des millions d'images étiquetées pour reconnaître une vaste gamme d'espèces.
* **Évaluation :** Les performances des modèles sont évaluées en utilisant des métriques standard pour assurer leur précision et leur fiabilité.
* **Déploiement :** Les modèles sont intégrés dans les applications iNaturalist pour fournir des suggestions d'identification en temps réel aux utilisateurs.

**4. Fonctionnement précis du modèle de machine learning :**

iNaturalist utilise des réseaux de neurones convolutifs (CNN) pour l'identification des espèces à partir des images. Le processus fonctionne comme suit :

* **Entrée :** L'utilisateur télécharge une image d'un organisme.
* **Prétraitement :** L'image est redimensionnée et normalisée pour correspondre aux exigences du modèle.
* **Extraction de caractéristiques :** Le CNN analyse l'image pour extraire des caractéristiques pertinentes, telles que les formes, les textures et les motifs.
* **Classification :** Les caractéristiques extraites sont utilisées pour attribuer une probabilité à chaque espèce possible.
* **Sortie :** Le système suggère à l'utilisateur les espèces les plus probables en fonction des probabilités calculées.

En résumé, iNaturalist combine une architecture technique solide avec des modèles avancés d'apprentissage automatique pour fournir une plateforme efficace d'identification et de partage des observations de la biodiversité.

* 1. **SOLUTIONS PROPOSEE**

Au vu de tout ce qui précède, nos principaux atouts reposent sur une application mobile fonctionnant hors-ligne, une spécialisation sur les invertébrés marins, des algorithmes optimisés et un temps de réponse relativement bas.

**2. Nature de la Solution Proposée**

L'application comportera :

* Un pipeline pour le prétraitement des photos télécharger
* Un système de classification basé sur l'intelligence artificielle
* Une base de données locale d'espèces d'invertébrés marins
* Une interface utilisateur intuitive
* Un système de synchronisation pour les mises à jour (en mode connecté)

**3. Méthodologie Envisagée**

**Phase 1 : Collecte et Préparation des Données**

* Acquisition de la base de données d'images d'invertébrés marins
* Nettoyage et prétraitement des images

**Phase 2 : Développement du Modèle**

* Sélection et adaptation d'architectures de Deep Learning.
* Entraînement et optimisation du modèle
* Tests et validation des performances

**Phase 3 : Développement Mobile**

* Création de l'interface utilisateur
* Intégration du modèle optimisé
* Implémentation du système de stockage local

**Phase 4 : Tests et Déploiement**

* Tests utilisateurs
* Optimisation des performances

**4. Résultats Envisagés**

* Précision de classification > 90% pour les espèces communes
* Temps de traitement < 5 secondes par image
* Taille de l'application < 100 MB
* Fonctionnement fluide sur des appareils de moyenne gamme

**5. Algorithme de Machine Learning**

**Architecture Proposée**

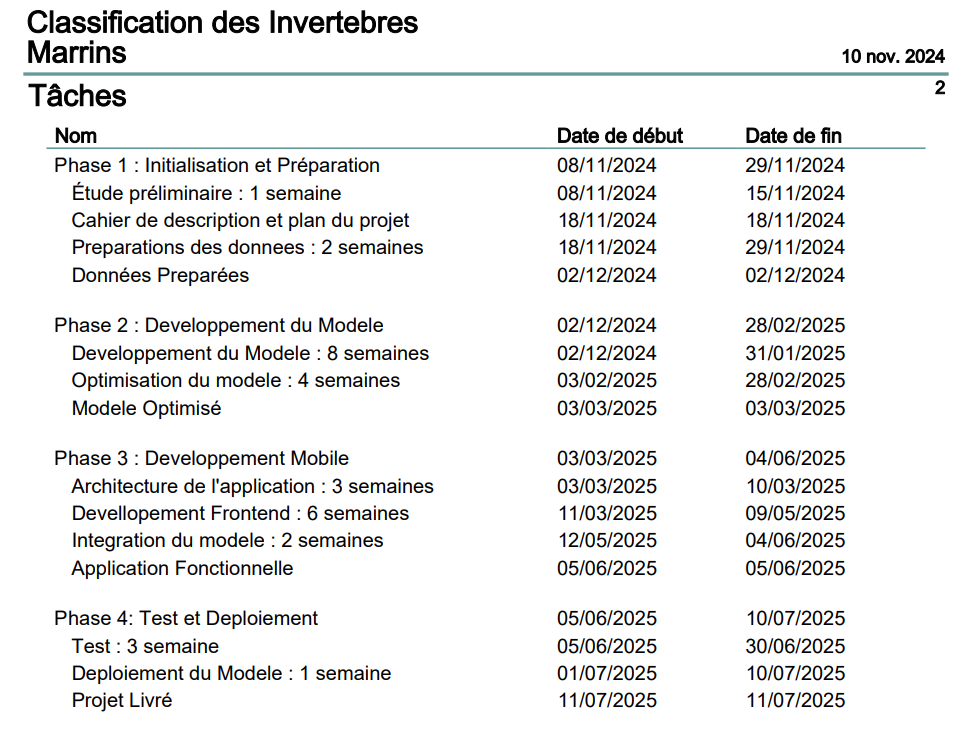
* Réseau de neurones convolutifs (CNN) optimisé pour mobile
* Options envisagées : MobileNetV3, Efficient Net-Lite
* Transfer Learning à partir de modèles pré-entraînés

**Pipeline de Traitement**

1. Prétraitement des images
2. Détection et segmentation
3. Extraction des caractéristiques
4. Classification
   1. **FONCTIONNALITES**

La fonction principale de notre application est d’identifier les invertébrés marins à partir d’image. Elle fonctionnera comme suit :

* Prise de photo
* Téléchargement et soumission des photos à l’application
* Analyse par IA : une fois que photo a été soumise, l’application utilisera des algorithmes d’apprentissage automatique pour analyser les caractéristiques visuelles de l’espèce, ensuite elle compare ces caractéristiques avec celles des espèces présente dans la base de données
* Ensuite à partir d’un algorithme de reconnaissance elle fait une correspondance entre les caractéristiques uniques de l’espèce et les caractéristiques des espèces disponibles dans la base de données
* Ensuite elle affiche l’espèce possible ou une liste d’espèce possible avec leurs différentes probabilités
  1. **PLANNING DE REALISATION**

****

* 1. **EQUIPE PROJET**

|  |  |  |  |
| --- | --- | --- | --- |
| Nom | Identifiant | Rôle | Contact |
| Pr Tsopze Norbert |  | Superviseur | norbert.tsopze@facsciences-uy1.cm |
| Noubissi Fopa Christian Junior | 19M2315 | Membre | christian.noubissi@facsiences-uy1.cm |
| Essuthi Mbangue Ange Armel | 0000 | Membre | essuthiange@gmail.com |