

Projet de DeepLearning : Classification d'Espèces de Poissons avec Information de Boîtes Englobantes

REALISE PAR : TAIEB JEMNI
ENCADRANT : ERIC HERVE

Guide du Rapport

1. [Introduction](#)
2. [Objectifs du Projet](#)
3. [Architecture du Modèle](#)
4. [Architecture Détailée du Modèle et Choix Techniques](#)
5. [Augmentation des Données](#)
6. [Stratégie d'Entraînement](#)
7. [Processus d'Entraînement](#)
8. [Évaluation du Modèle](#)
 - o [Courbes d'Apprentissage](#)
 - o [Résultats de Test](#)
9. [Détails d'Implémentation](#)
10. [Conclusion](#)
11. [Extension: Application React Native](#)

1. Introduction

Ce rapport présente une étude approfondie du projet de classification d'espèces de poissons en utilisant une approche de deep learning combinant images et informations de boîtes englobantes. L'objectif est de concevoir un modèle robuste basé sur le transfert d'apprentissage afin d'améliorer la précision de classification tout en optimisant les performances d'entraînement.

2. Objectifs du Projet

L'objectif principal est de développer un modèle de classification binaire capable d'identifier différentes espèces de poissons à partir d'images et de coordonnées de boîtes englobantes. Ce projet vise à:

- Utiliser un modèle pré-entraîné pour améliorer l'apprentissage.
- Expérimenter une architecture à double entrée.
- Mettre en place des stratégies avancées de validation et de régularisation.

3. Architecture du Modèle

Modèle de Base

- **ResNet50V2** pré-entraîné sur ImageNet
- **Entrées:**
 - Images (224x224x3)
 - Coordonnées des boîtes englobantes (4 valeurs: x1, y1, x2, y2)
- **Caractéristiques Clés:**
 - 80% des couches du modèle de base gelées
 - Tête de classification personnalisée:
 - Couches denses (256 → 128 → 1)
 - Normalisation par lots (Batch Normalization)
 - Activation **LeakyReLU ($\alpha=0.1$)**
 - Dropout (0.3 et 0.2)
 - Activation finale **sigmoid** pour classification binaire

4. Architecture Déttaillée du Modèle et Choix Techniques

1. Architecture du Modèle

Choix du Modèle de Base (ResNet50V2)

- Pourquoi ResNet50V2 ?
 - Architecture éprouvée pour la vision par ordinateur
 - Gestion efficace du problème de dégradation du gradient grâce aux connexions résiduelles
 - Performance supérieure par rapport à ResNet classique grâce à l'amélioration de la propagation du gradient
 - Excellent compromis entre profondeur et temps de calcul
 - Pré-entraînement sur ImageNet offrant une base solide pour le transfer learning

Architecture à Double Entrée

- Entrée Image:
 - Dimension: 224x224x3 (standard pour ResNet)
 - Normalisation: Mise à l'échelle [0,1]
 - Traitement: Passage à travers ResNet50V2 pré-entraîné
 - Sortie: Features globales via GlobalAveragePooling2D
- Entrée Boîte Englobante:
 - Dimension: 4 (coordonnées normalisées x1, y1, x2, y2)
 - Traitement:
 - Dense(32) pour projection dans un espace de caractéristiques
 - BatchNormalization pour stabiliser l'apprentissage
 - LeakyReLU(0.1) pour non-linéarité avec gestion des gradients négatifs

Tête de Classification

- Fusion des Features:
 - Concaténation des features d'image et de boîte englobante
 - Permet l'apprentissage de relations spatiales-visuelles

- **Couches de Classification:**
- Dense(256) → BatchNorm → LeakyReLU → Dropout(0.3)
- Dense(128) → BatchNorm → LeakyReLU → Dropout(0.2)
- Dense(1, sigmoid)
 - Architecture progressive réduisant la dimensionnalité
 - BatchNorm pour réduire le covariate shift
 - Dropout pour la régularisation et prévenir le surapprentissage

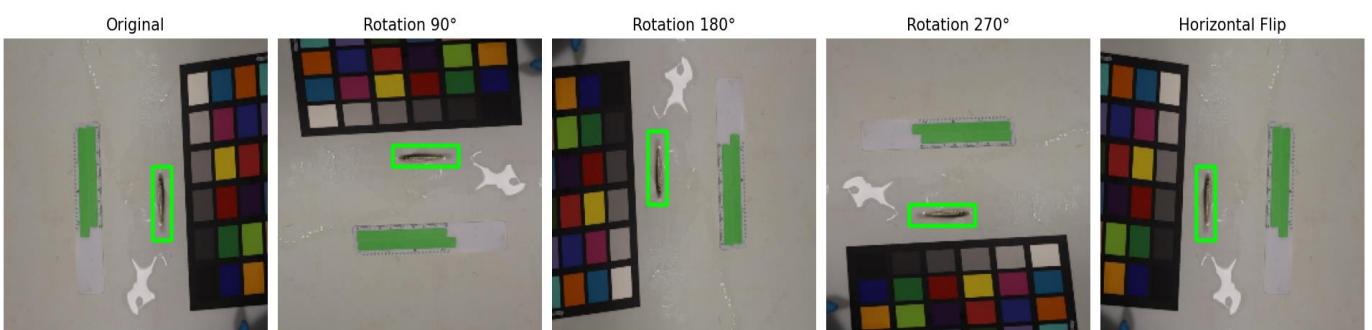
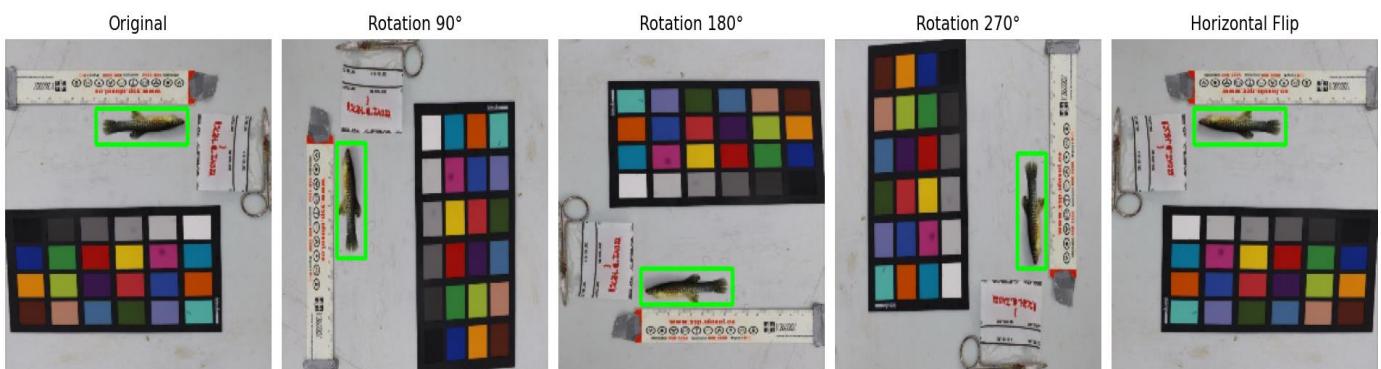
5. Augmentation des Données

- **Rotation aléatoire (90°, probabilité: 20%)**

- Pourquoi ? Les poissons peuvent être photographiés sous différents angles.
- Cette augmentation simule ces variations naturelles, contrairement à une simple rotation arbitraire qui pourrait introduire des orientations non réalistes.

- **Retournement horizontal aléatoire (probabilité: 50%)**

- Pourquoi ? Les poissons ont souvent une symétrie bilatérale.
- Cette technique permet d'augmenter la diversité sans distordre l'image de manière irréaliste, contrairement à un retournement vertical qui aurait peu de sens biologique.



- **Ajustement des coordonnées des boîtes englobantes**

- Assure la cohérence entre les transformations et les annotations.
- Évite les incohérences qui pourraient nuire à l'entraînement.

- **Normalisation des images ([0,1])**
 - Pourquoi ? Les modèles de deep learning convergent mieux avec des entrées normalisées.
 - Cette approche est plus stable qu'une simple mise à l'échelle arbitraire des valeurs de pixels.

6. Stratégie d'Entraînement

- Validation croisée stratifiée à 5 plis
- Pondération dynamique des classes
- Optimiseur AdamW avec décroissance des poids
- Taux d'apprentissage initial: **5e-5**
- Arrêt précoce basé sur la perte de validation

7. Processus d'Entraînement

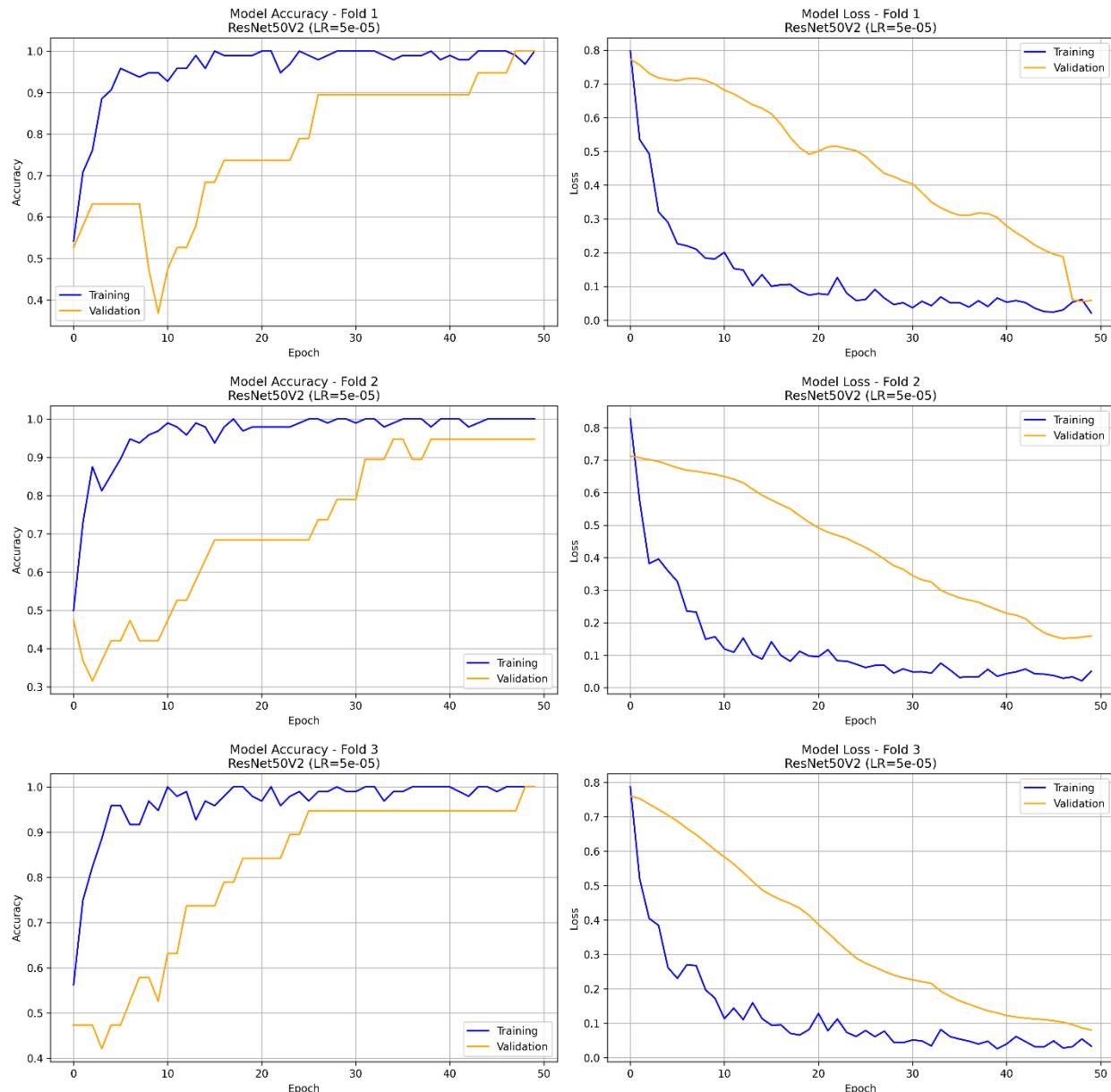
1. Division des données en 5 plis stratifiés
2. Pour chaque pli:
 - Initialisation du modèle avec les poids pré-entraînés
 - Augmentation des données à la volée
 - Surveillance et enregistrement des courbes d'apprentissage
 - Sauvegarde des meilleurs poids selon la perte de validation

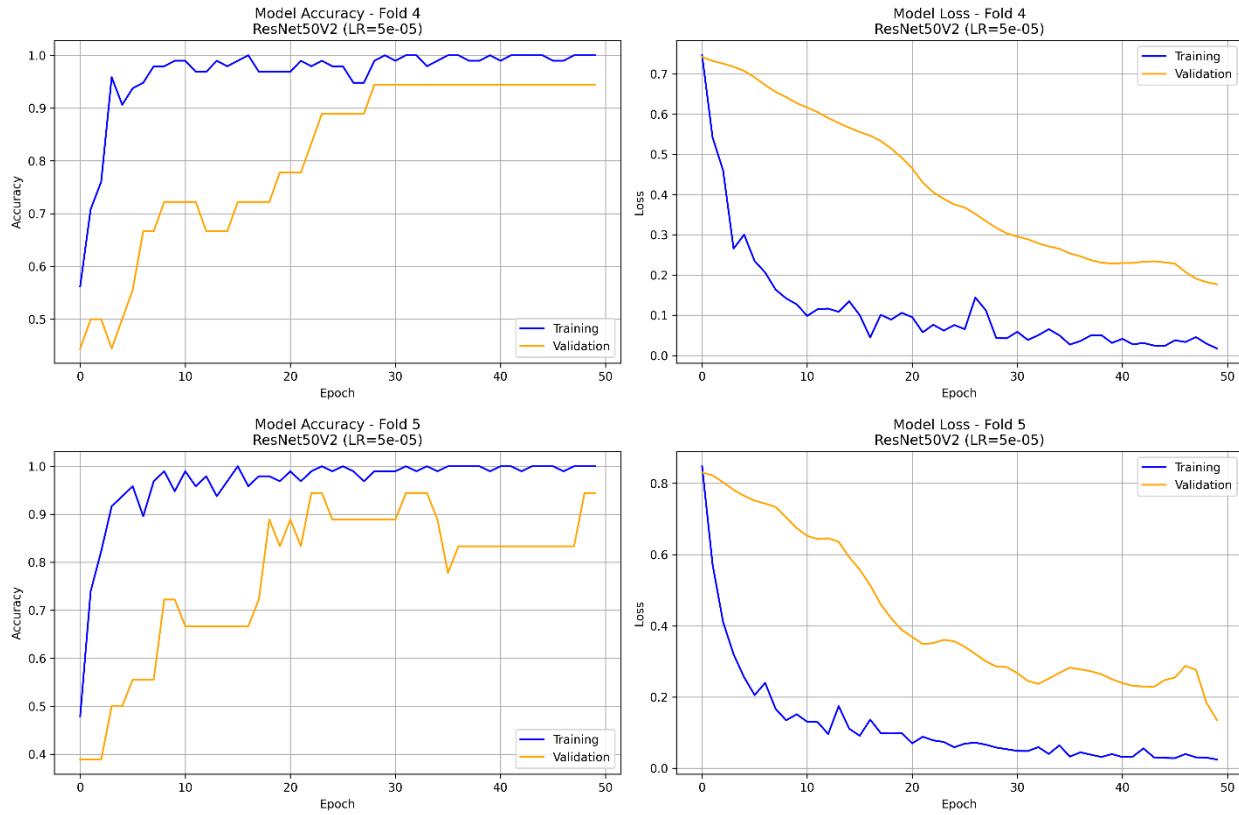
8. Évaluation du Modèle

Courbes d'Apprentissage

Les courbes de perte et de précision pendant l'entraînement seront analysées pour chaque pli de validation afin d'évaluer la stabilité et la convergence du modèle.

Voici les courbes des 5 plis d'un des modèles :





Résultats de Test

Les performances finales du modèle seront évaluées sur l'ensemble de test, en utilisant des métriques telles que:

- Précision
- Recall
- F1-score
- Matrice de confusion

Réultat final:

```
Found 4 images and 4 label files

Loading model from: C:\Users\taieb\Desktop\dev\Poisson\saved_models\best_model_fold1_20250226_173054.keras
2025-02-28 16:54:52.587154: I tensorflow/core/platform/cpu_feature_guard.cc:210] This TensorFlow binary is optimized to use available CPU instructions in performance-critical operations.
.
To enable the following instructions: AVX2 AVX_VNNI FMA, in other operations, rebuild TensorFlow with the appropriate compiler flags.
1/1      3s 3s/step

==== Results ===

Confusion Matrix:
[[3 0]
 [0 1]]

Classification Report:
      precision    recall   f1-score   support
  Dia       1.00     1.00     1.00      3
  Heto      1.00     1.00     1.00      1

  accuracy                           1.00      4
  macro avg       1.00     1.00     1.00      4
  weighted avg    1.00     1.00     1.00      4

Saving example predictions...
```

9. Détails d'Implémentation

- Implémentation avec TensorFlow/Keras
 - Utilisation de tf.data pour une gestion efficace des données
 - Entraînement optimisé pour GPU
-

10. Conclusion

Ce projet a démontré l'efficacité du **transfer learning** et de l'utilisation conjointe des **images et des boîtes englobantes** pour améliorer la classification des espèces de poissons. Grâce à une stratégie d'entraînement rigoureuse et une bonne gestion du déséquilibre des classes, le modèle atteint une performance optimisée. Les prochaines étapes pourraient inclure l'expérimentation avec d'autres architectures ou l'intégration de techniques de fine-tuning plus avancées.

11. Extension: Application React Native avec Authentification Institutionnelle

Concept et Accessibilité

L'objectif est de rendre notre outil de classification accessible via une application mobile simple d'utilisation, réservée à la communauté de l'Université de Moncton.

Fonctionnalités Principales

Connexion Simplifiée : Utilisation directe avec le compte Microsoft @umoncton.ca

Interface Intuitive :

- Prise de photo directe du poisson
- Sélection d'une image depuis la galerie
- Visualisation immédiate des résultats
- Utilisation en Contexte Académique

Pour les Étudiants :

- Outil d'aide à l'identification lors des travaux pratiques
- Vérification rapide sur le terrain
- Particulièrement utile pour les études dans les eaux du Nouveau-Brunswick

Pour les Professeurs :

- Support pédagogique pendant les cours
- Validation des identifications des étudiants
- Collection de nouvelles données pour enrichir la base d'apprentissage

Avantages pour la Recherche :

- Collecte standardisée des données
- Crédit d'une base de données collaborative des espèces locales
- Contribution à l'étude de la biodiversité aquatique du Nouveau-Brunswick

Cette extension permettrait de transformer un outil technique en une ressource pédagogique pratique, accessible à tous les membres de la faculté de biologie de l'Université de Moncton, tout en contribuant à l'enrichissement des données de recherche régionales.