

Reconnaissance des émotions faciales avec ResNet50

Découvrez comment le Deep Learning et le Transfer Learning révolutionnent l'analyse des émotions.

Membre du groupe :

Kevin TCHIAKAY

Erwann HILLION



Objectifs du projet



Construire un modèle de Deep Learning

Capable de reconnaître avec précision les émotions faciales.



Utiliser un réseau pré-entraîné (ResNet50)

Pour capitaliser sur des connaissances préexistantes et améliorer les performances.



Évaluer la performance du modèle

À l'aide de métriques clés comme l'accuracy, les courbes d'apprentissage et une matrice de confusion.

Ce projet explore les défis et les solutions pour l'analyse automatisée des expressions humaines.

Les défis de la reconnaissance d'émotions



Expressions subtiles et variabilité

Les émotions peuvent être nuancées, et la diversité des visages (lumière, angle, âge) complexifie l'apprentissage.

Classes déséquilibrées

Certaines émotions sont moins fréquentes (ex: "dégoût") que d'autres ("joie"), posant des problèmes d'apprentissage.

Ambiguïté visuelle

Des émotions différentes peuvent avoir des manifestations visuelles très similaires, rendant la distinction difficile même pour l'œil humain.

Données et prétraitement

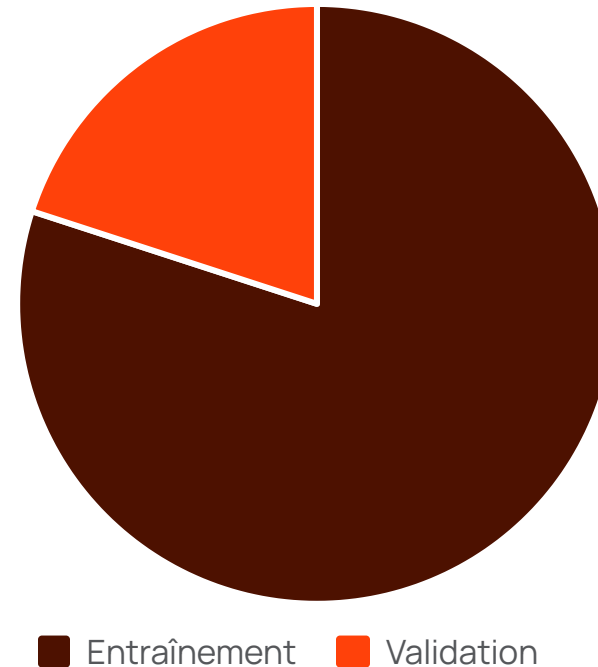
Répartition des données

Dataset

- Images de visages classées par émotions.
- Organisation : dossiers `train` / `validation` / `test`.

Prétraitement

- Redimensionnement : `224 × 224` pixels.
- Normalisation : Utilisation de `preprocess_input` spécifique à ResNet50.



Des optimisations telles que `tf.data`, le cache et le prefetch sont utilisées pour un entraînement rapide avec une taille de lot de 32.

Qu'est-ce que ResNet50 ?

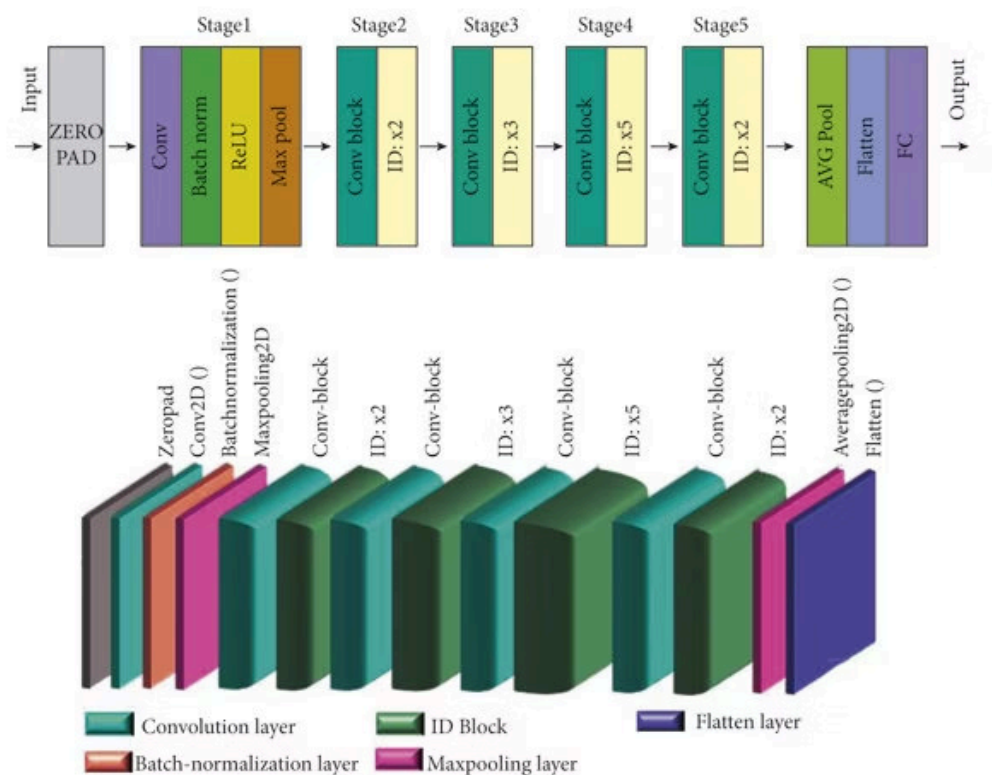
ResNet50 est un réseau de neurones convolutionnel profond (50 couches), pré-entraîné sur l'immense base de données ImageNet. Son innovation majeure réside dans les **skip connections**.

Le défi du Vanishing Gradient

Les réseaux profonds souffrent souvent du problème de la disparition du gradient, qui empêche l'apprentissage des couches initiales.

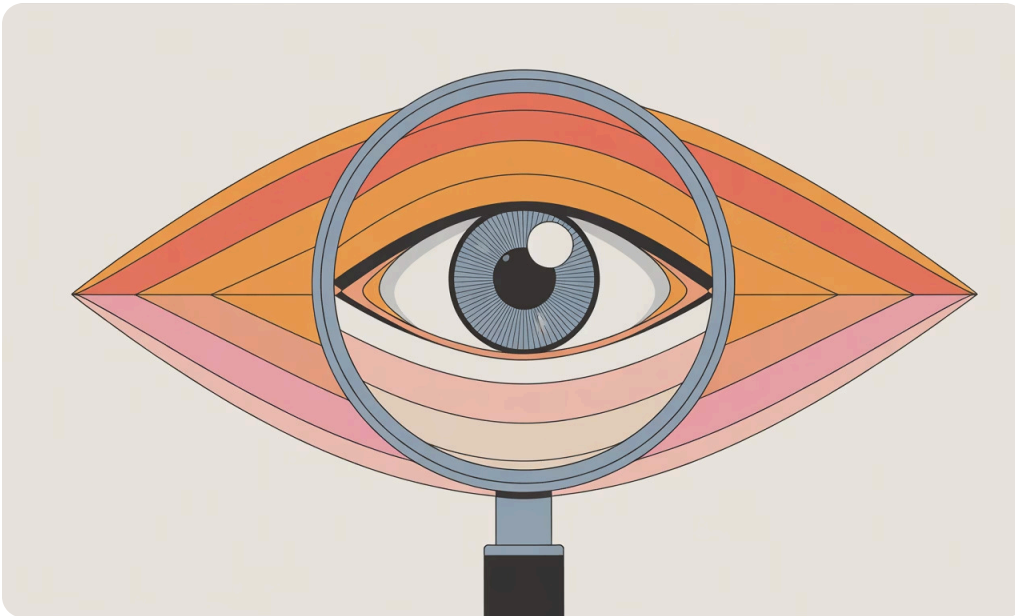
La solution des connexions résiduelles

ResNet ajoute des connexions qui contournent une ou plusieurs couches (skip connections), permettant au gradient de se propager plus facilement.



Ce mécanisme permet d'entraîner des réseaux très profonds, offrant de meilleures performances et une convergence plus stable.

Pourquoi le Transfer Learning avec ResNet50 ?



1 Connaissances pré-acquises

ResNet50 a déjà "appris" à détecter des caractéristiques visuelles fondamentales comme les contours, les textures et les formes générales sur des millions d'images diverses.

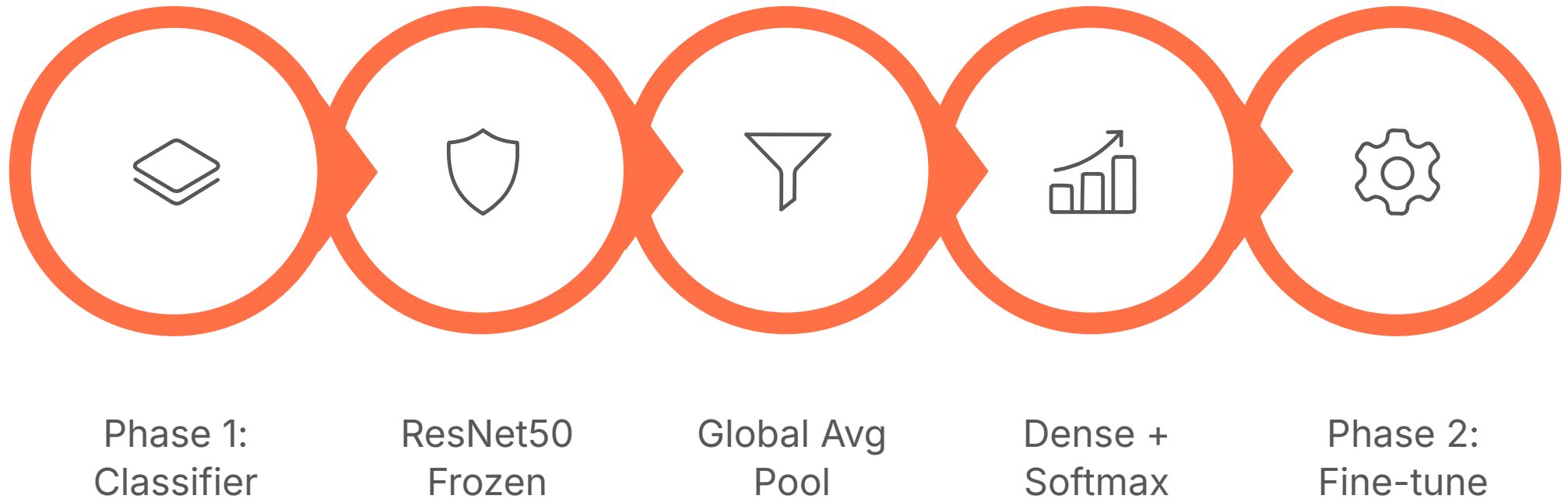
2 Réutilisation des connaissances

Nous réutilisons ces connaissances génériques pour les appliquer à la reconnaissance des visages et des émotions, accélérant ainsi le processus d'apprentissage.

📌 Avantages clés

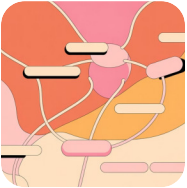
- Moins de données nécessaires pour l'entraînement.
- Entraînement beaucoup plus rapide.
- Meilleure généralisation du modèle sur de nouvelles données.

Architecture du modèle



L'architecture combine un puissant **Backbone ResNet50** pré-entraîné sur ImageNet avec des couches finales adaptées à notre tâche de classification d'émotions.

Stratégies d'optimisation avancées



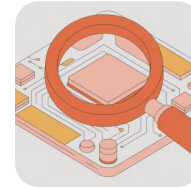
Label Smoothing (0.1)

Aide à éviter que le modèle ne devienne trop confiant dans ses prédictions, améliorant ainsi la généralisation.



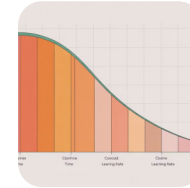
Class Weights

Corrige le déséquilibre des classes en attribuant un poids plus important aux classes sous-représentées.



Fine-tuning

Adaptation des dernières couches de ResNet50 aux caractéristiques spécifiques des émotions faciales.



Cosine Decay Learning Rate

Permet un ajustement dynamique du taux d'apprentissage pour une convergence plus efficace et stable.



Early Stopping

Arrête l'entraînement lorsque la performance sur l'ensemble de validation cesse de s'améliorer, prévenant le surapprentissage.

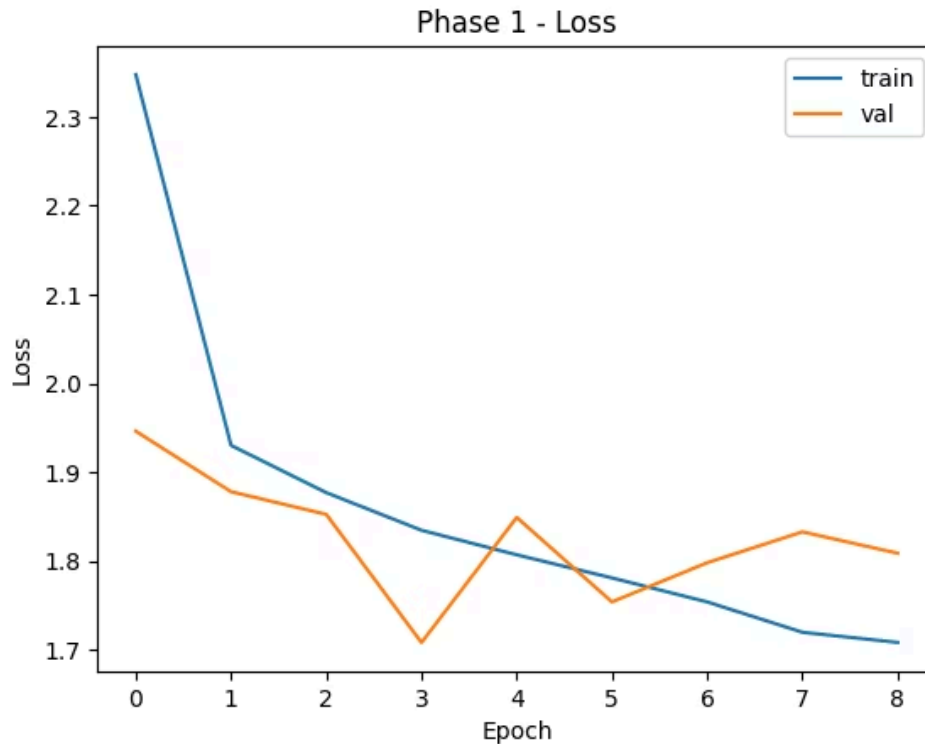


Mixed Precision (GPU)

Utilise des calculs en précision mixte pour accélérer l'entraînement sur GPU tout en maintenant la précision.

Courbes d'apprentissage et résultats

(Image illustrative,



vos graphiques réels doivent être insérés ici)

Analyse des courbes

- Les courbes de perte (loss) d'entraînement et de validation diminuent de manière constante.
- L'écart entre les courbes reste minime, indiquant l'absence de surapprentissage majeur.
- Une amélioration notable est observée après la phase de fine-tuning, où le modèle affine sa compréhension des émotions.

Ces indicateurs suggèrent que le modèle généralise correctement et apprend efficacement à partir des données fournies.

Accuracy obtenue

Accuracy globale sur l'ensemble de test : 48%

Le fine-tuning a démontré une amélioration significative des performances, confirmant son importance dans notre approche.

Matrice de confusion



ive, votre heatmap réelle doit être insérée ici)

La matrice de confusion fournit une vue détaillée des performances du modèle par classe, révélant les points forts et les axes d'amélioration.

Observations clés

Prédictions robustes

Excellentes prédictions pour les émotions "heureux" et "neutre", qui sont souvent plus distinctes visuellement.

Confusions fréquentes

Des confusions sont observées entre "peur" et "surprise", ainsi qu'entre "tristesse" et "neutre".

Interprétation

Ces confusions sont logiques, car ces paires d'émotions peuvent partager des caractéristiques faciales subtiles et être difficiles à distinguer même pour l'œil humain.