

Note méthodologique

Preuve de concept pour la classification automatique d'images

Table des matières

1. Dataset retenu

- 1.1. Fichier de données
- 1.2. Dossier "Images"
- 1.3. Catégorisation des articles

2. Concepts des algorithmes utilisés : VGG16 et EfficientNet-B7

- 2.1. VGG16
- 2.2. EfficientNet-B7

3. Modélisation

- 3.1. Méthodologie de modélisation
- 3.2. Métriques d'évaluation

4. Synthèse des résultats

- 4.1. Résultats VGG16
- 4.2. Résultats EfficientNet-B7
- 4.3. Conclusion

5. Analyse de la Feature importance

6. Limites et améliorations possibles

- 6.1. Limites du modèle
- 6.2. Perspectives d'amélioration

1. Dataset retenu

L'objectif principal de cette preuve de concept est de démontrer la capacité de classification automatique d'articles à partir d'images. Le dataset utilisé contient des images d'articles classés en **sept** catégories principales, fournies par l'entreprise "Place de Marché". Ce processus est actuellement réalisé manuellement par les vendeurs, ce qui justifie la nécessité d'une automatisation.

1.1. Fichier de données

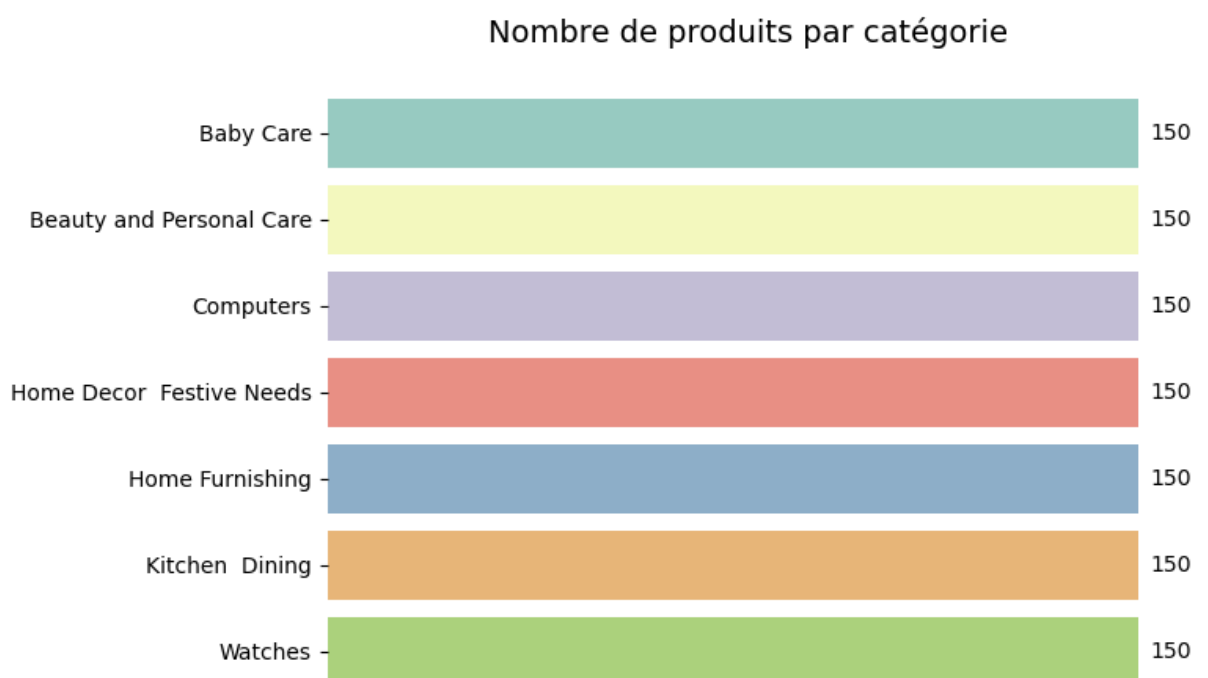
- Fichier CSV contenant les détails des articles, comme le nom, la catégorie et les images associées.
- 1050 lignes et 15 colonnes comprenant les principales informations des produits.

1.2. Dossier "Images"

- 1050 images associées aux articles dans un répertoire dédié.
- Chaque image correspond à une ligne dans le fichier CSV, permettant une classification basée sur l'image associée.

1.3. Catégorisation des articles

Les articles sont répartis en sept catégories principales : Baby Care, Beauty and Personal Care, Computers, Home Décor Festive Needs, Home Furnishing, Kitchen Dining, et Watches. Le but est d'automatiser cette attribution de catégories via un modèle de classification d'images.



2. Concepts des algorithmes utilisés : VGG16 et EfficientNet-B7

Les réseaux de neurones convolutifs (ConvNets ou CNN) sont largement utilisés pour la reconnaissance et la classification d'images. Traditionnellement, leur taille a été augmentée pour une meilleure précision, dépassant souvent les limitations des ressources informatiques.

2.1. VGG16

Le modèle VGG16 est un réseau de neurones convolutifs populaire, utilisé pour la classification d'images. Ce modèle utilise des couches convolutives suivies de couches entièrement connectées pour effectuer la classification.

2.2. EfficientNet-B7

EfficientNet est un réseau de neurones avancé qui applique une mise à l'échelle composée, équilibrant profondeur, largeur et résolution de manière optimisée. EfficientNet-B7, une version plus large du modèle, est sélectionné pour sa précision et son efficacité, surpassant de nombreux autres modèles de CNN, avec un nombre de paramètres et une consommation de ressources inférieure à celles de VGG16.

Ce que fait EfficientNet :

1. **Mise à l'échelle équilibrée (Compound Scaling)** : Contrairement aux autres modèles classiques, EfficientNet ajuste simultanément trois dimensions du réseau :
 - **La profondeur** (nombre de couches dans le réseau),
 - **La largeur** (nombre de filtres ou de canaux par couche),
 - **La résolution** des images d'entrée.

Cette approche permet au modèle de gérer des images de haute résolution avec plus de couches et de détails, tout en restant plus efficace qu'un simple agrandissement des réseaux.

2. **Efficacité optimisée** : EfficientNet a été conçu pour atteindre une précision élevée avec beaucoup moins de paramètres et de calculs que les modèles comme VGG ou ResNet. Cela signifie qu'il peut obtenir des résultats comparables, voire meilleurs, avec un modèle plus petit et plus rapide.

3. Modélisation

3.1. Méthodologie de modélisation

Deux modèles ont été utilisés pour cette preuve de concept :

- **VGG16** : Un modèle pré-entraîné avec ImageNet, utilisé pour la classification initiale.
- **EfficientNet-B7** : Un modèle plus avancé, également pré-entraîné sur ImageNet.

La méthodologie est la suivante :

- Les images sont redimensionnées à 224x224 pour correspondre aux exigences des modèles pré-entraînés.
- Le dataset est divisé en trois parties : 75% pour l'entraînement, 25% pour la validation, et une partie dédiée au test.
- Apprentissage par transfert : Les couches finales des deux modèles sont remplacées par des couches entièrement connectées adaptées au dataset.

3.2. Métriques d'évaluation

L'**Accuracy** a été choisie comme principale métrique pour évaluer les performances des modèles. Elle mesure le pourcentage d'images correctement classées.

Des techniques telles que l'arrêt précoce (EarlyStopping) et la sauvegarde des meilleurs poids (Checkpoint) ont été utilisées pour éviter le surapprentissage.

4. Synthèse des résultats

4.1. Résultats VGG16

Après entraînement, VGG16 a atteint une précision de **0.93%** sur l'ensemble de validation et de **84.75%** sur l'ensemble de test.

Bien que performant, ce modèle montre des signes de surajustement, avec une baisse notable de l'Accuracy sur le jeu de test.

4.2. Résultats EfficientNet-B7

Le modèle EfficientNet-B7 a montré de meilleures performances avec une **Accuracy de validation de 98.17%** et une **Accuracy de test de 85.59%**.

Cela montre une meilleure capacité de généralisation comparée à VGG16.

4.3. Conclusion

En comparaison avec VGG16, EfficientNet-B7 a montré des performances globalement meilleures, tant sur l'ensemble de validation que sur celui de test. Le modèle EfficientNet, grâce à sa capacité de mise à l'échelle, parvient à maintenir une précision élevée tout en évitant le surajustement.

5. Analyse de la Feature importance du nouveau modèle

Pour interpréter les prédictions, nous avons utilisé la méthode SHAP afin de visualiser l'importance des différentes parties d'une image dans la décision du modèle. Cela nous permet de mieux comprendre pourquoi le modèle fait certaines prédictions.

6. Limites et améliorations possibles

6.1. Limites du modèle

- **Interprétabilité** : Malgré l'utilisation de SHAP, l'interprétation des modèles de deep learning reste complexe.
- **Biais des données** : Le modèle dépend fortement des données d'entraînement.

6.2. Perspectives d'amélioration

- **Amélioration des données** : L'ajout de nouvelles images et une plus grande diversité des données aideraient à mieux généraliser.
- **Optimisation des hyperparamètres** : Une exploration plus poussée des hyperparamètres pourrait améliorer davantage les résultats du modèle.

En conclusion, l'utilisation d'EfficientNet-B7 apporte des améliorations significatives par rapport à VGG16 en termes de précision et de capacité à généraliser les résultats, tout en maintenant une efficacité computationnelle.