



Assurez une veille technique

Projet 8 – Parcours Data Scientist

Sommaire



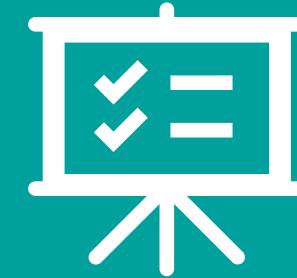
CONTEXTE

VEILLE TECHNIQUE : CLIP

Contrastive
Language – Image
Pre-training



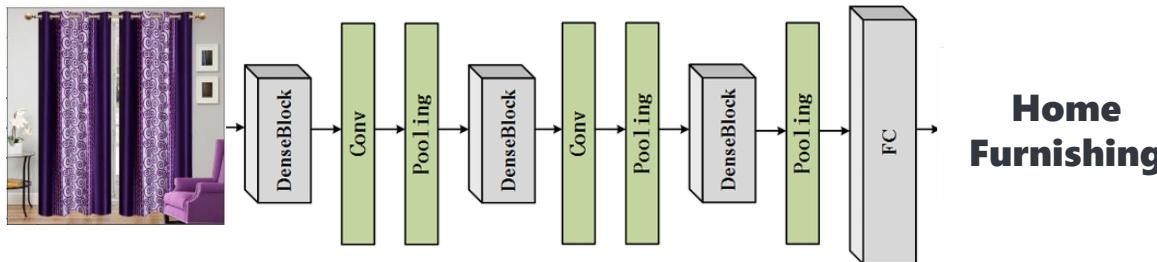
CONCLUSION
&
PERSPECTIVES



Veille technique sur un modèle multimodal (Texte + Image)

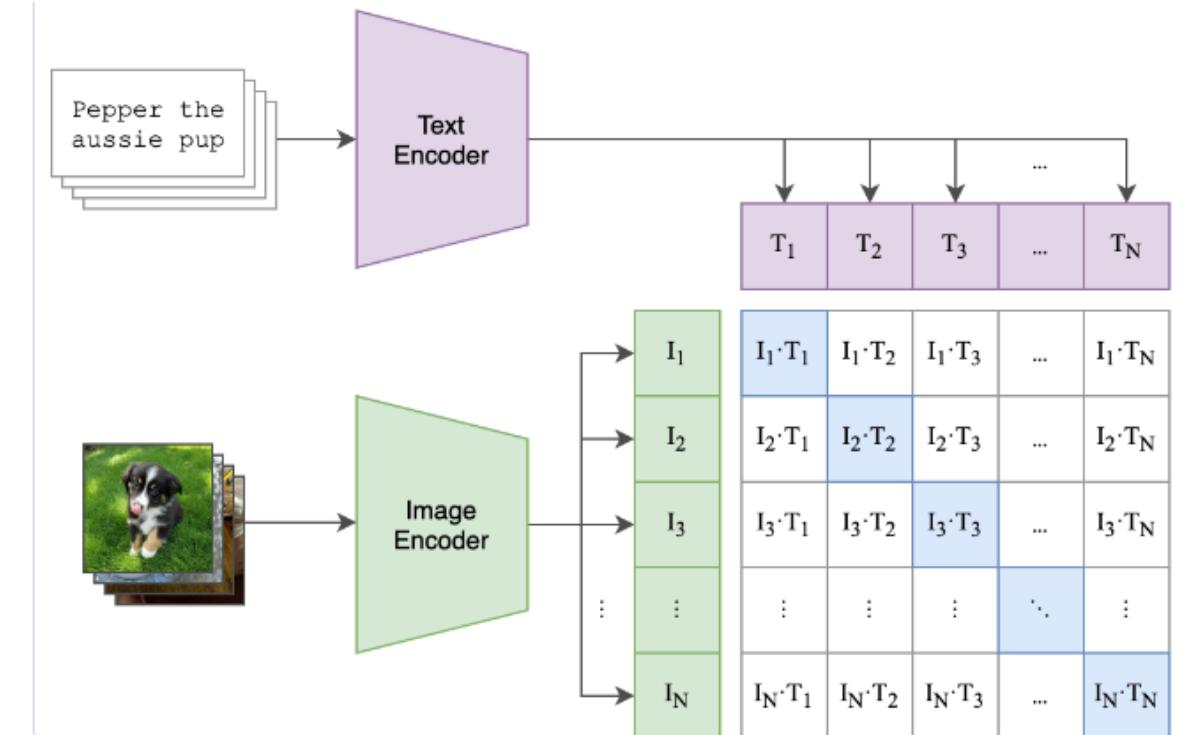
Approche classique : DenseNet201 avec Data Augmentation

Supervised Learning with Data Augmentation



Modèle multimodal : CLIP (Contrastive Language-Image Pre-training)

Zero-shot learning

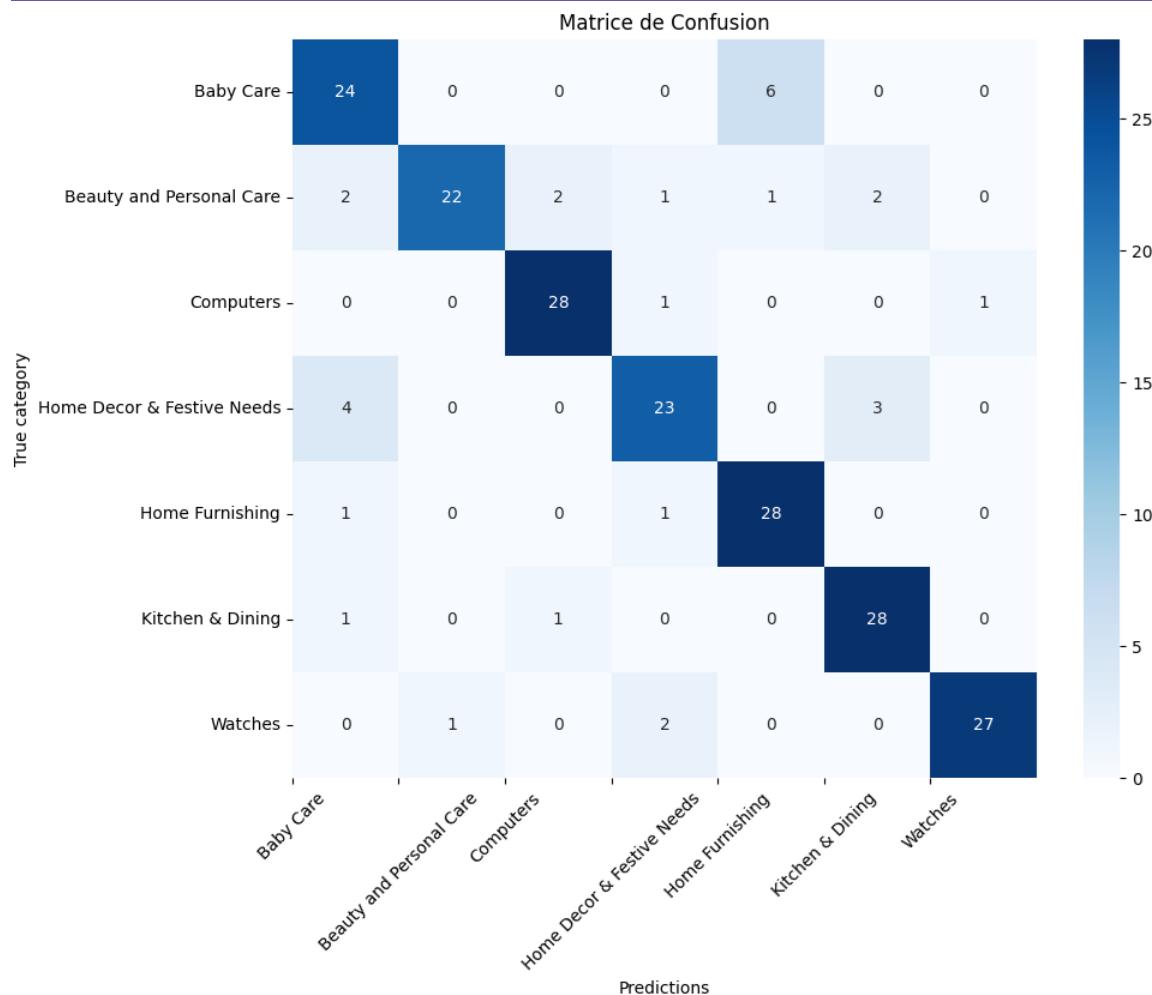


Jeu de données biens de consommation

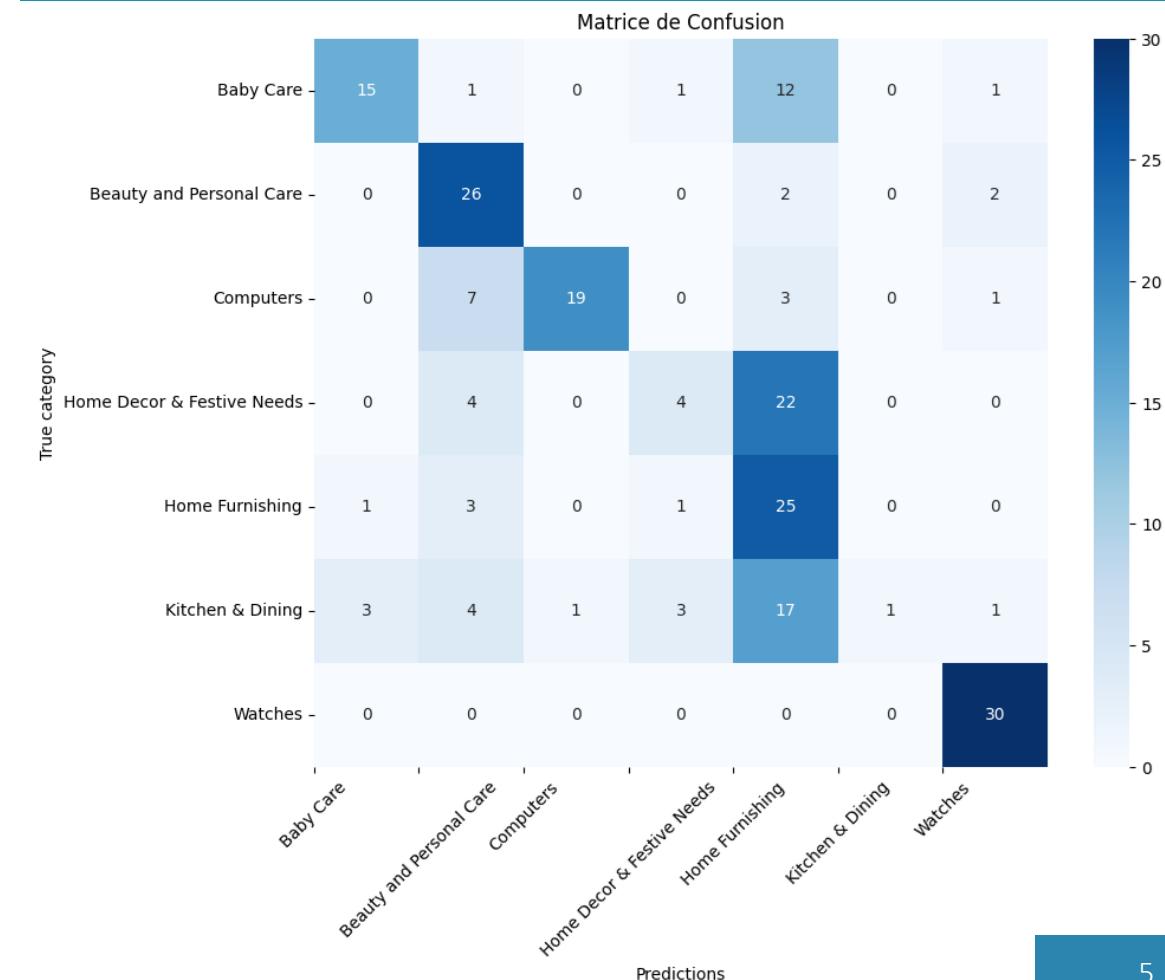


DenseNET vs CLIP avec noms des catégories

DenseNet201 avec Data Augmentation



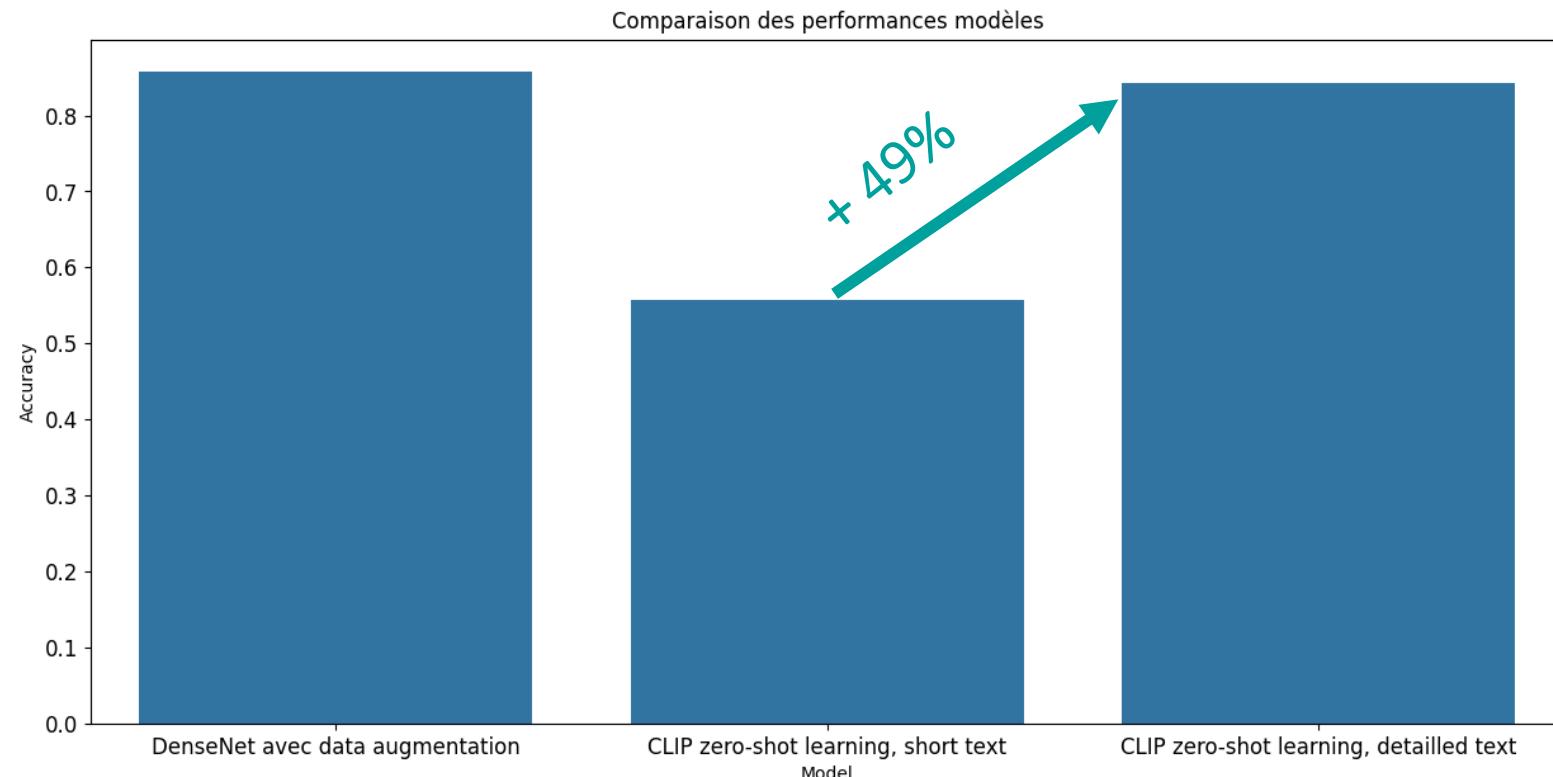
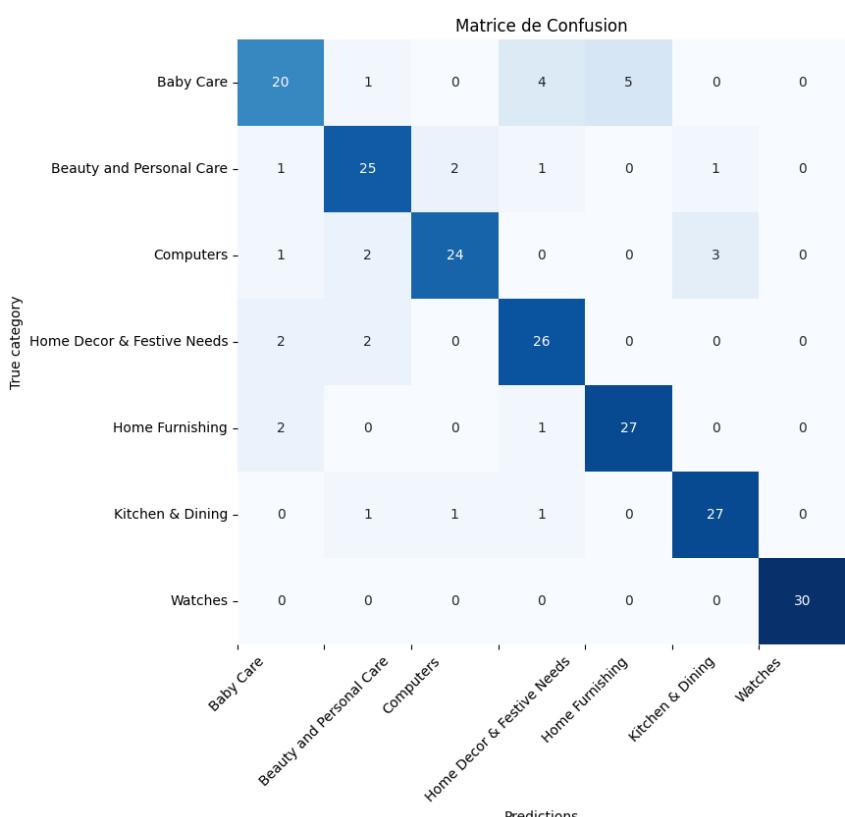
CLIP Zero-shot avec noms des catégories



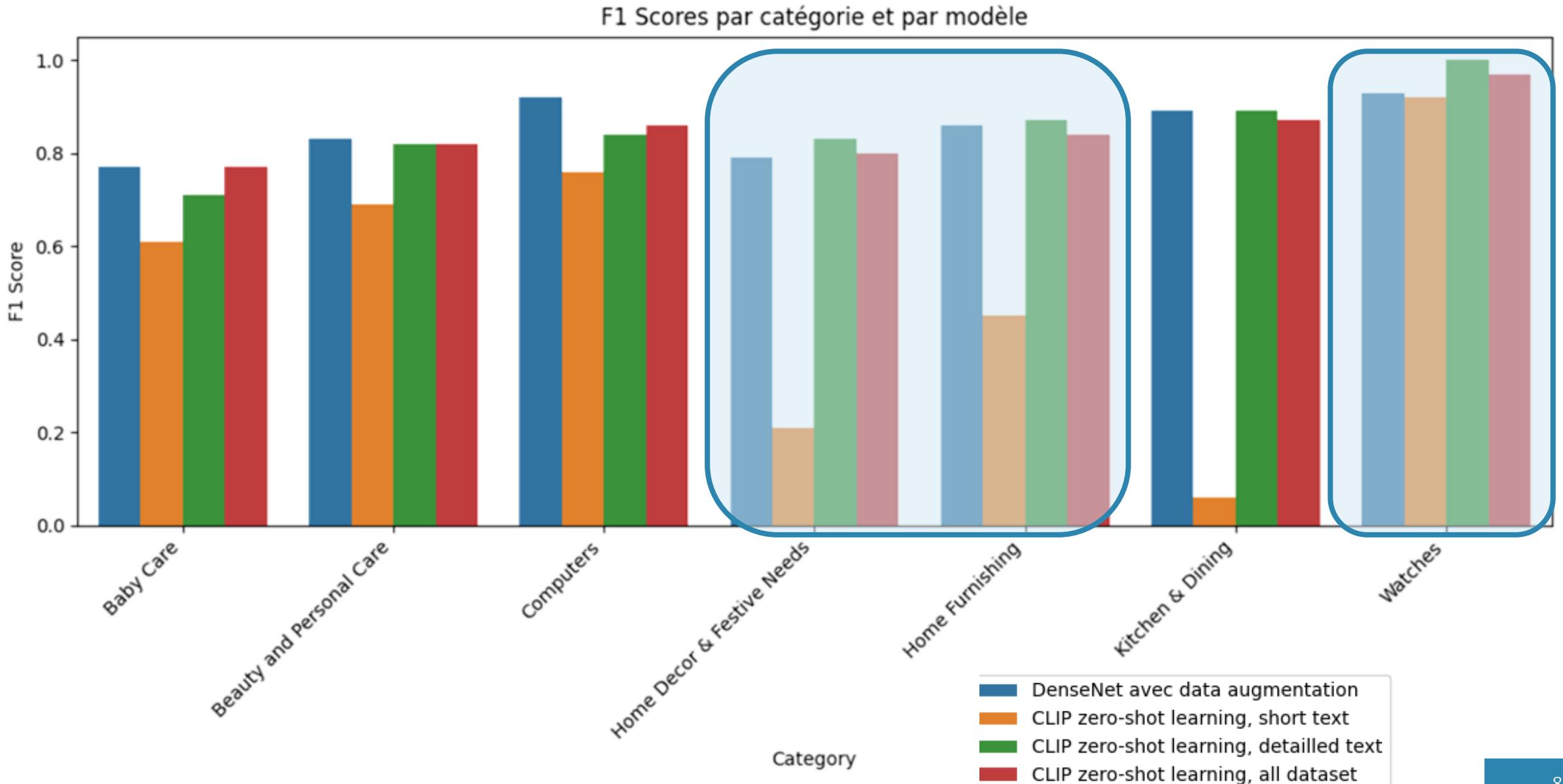
CLIP : enrichir l'entrée texte avec plus de détails

Nom des catégories	Description détaillée
Baby Care	Clothes, toys, bed, furniture for infant, baby boys and girls
Beauty and Personal Care	Fragrances, Products and accessories for hair, body, skin, eye care for women and men
Computers	Components and accessories for computer, laptop, tablet
Home Decor & Festive Needs	Home decor like showpieces, handicrafts, lamps, vases
Home Furnishing	Curtains, carpets, linen for bed, bath and dining, pillows and others home furnishing
Kitchen & Dining	Coffee mugs, Cookware, Kitchen Tools, Bottles, Lighting and dining accessories
Watches	Wrist Watches

Forte amélioration des performances avec le texte enrichi



CLIP Zero-shot Learning : des résultats très prometteurs

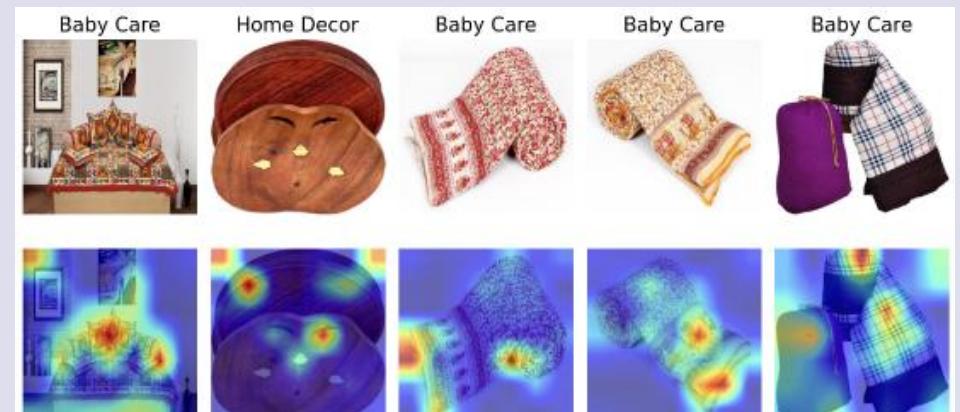
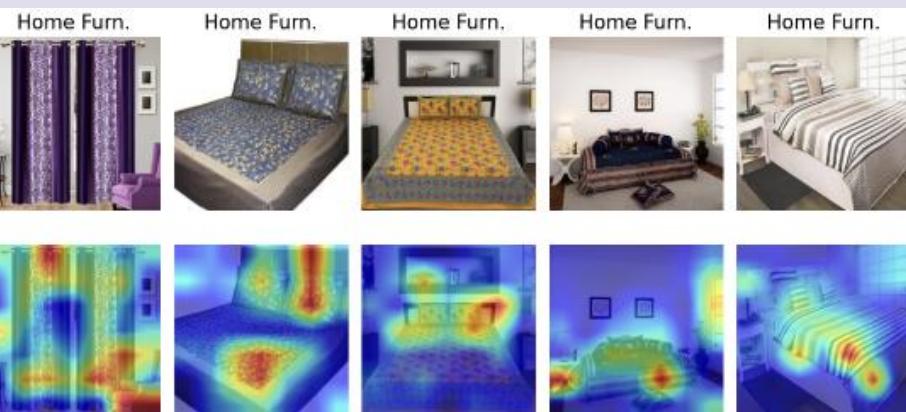
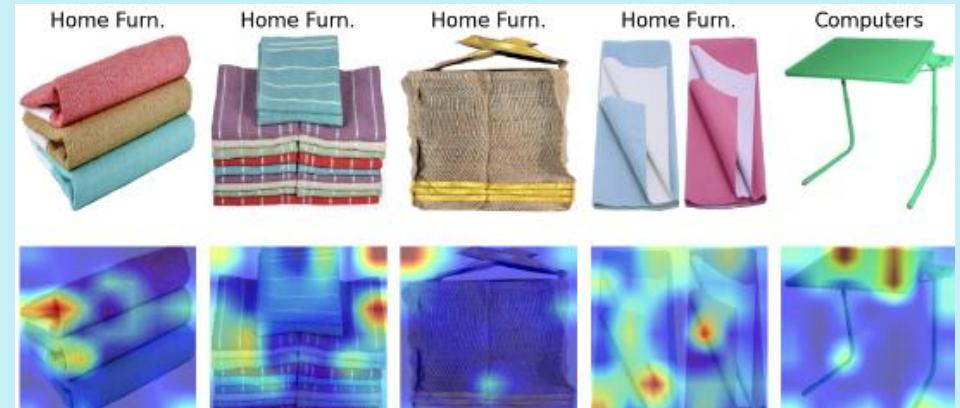


Explicabilité du modèle : Cartes d'attention

Prédictions correctes

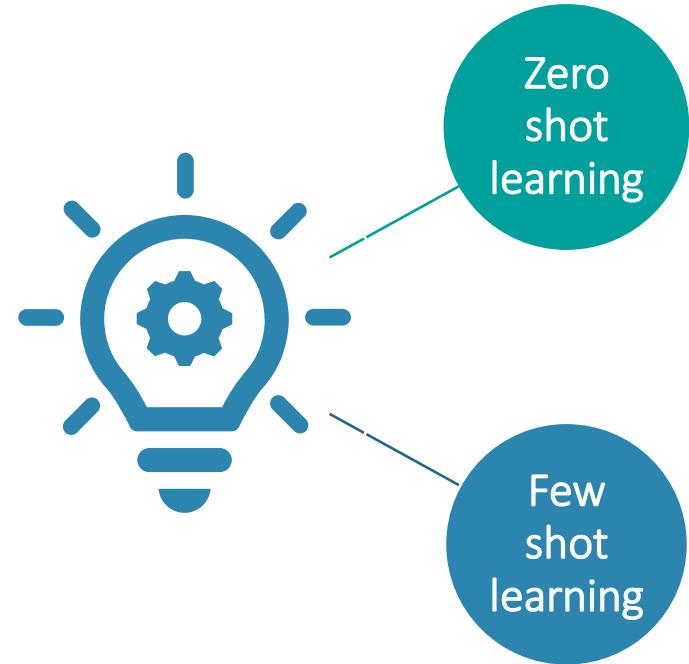
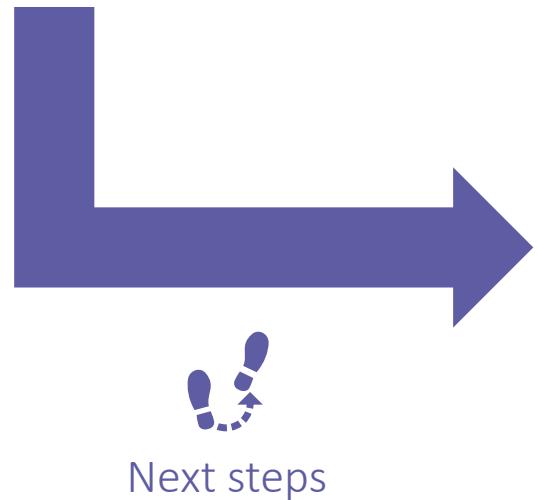


Prédictions incorrectes



Conclusion & Perspectives

CLIP en zero-shot learning montre une bonne capacité de généralisation sans entraînement spécifique



Description plus précise basée sur l'analyse des cartes d'attention

Quelques exemples par sous-catégories

MERCI POUR VOTRE ATTENTION

Back-up slides

Des articles très différents au sein d'une même catégorie

Baby Care



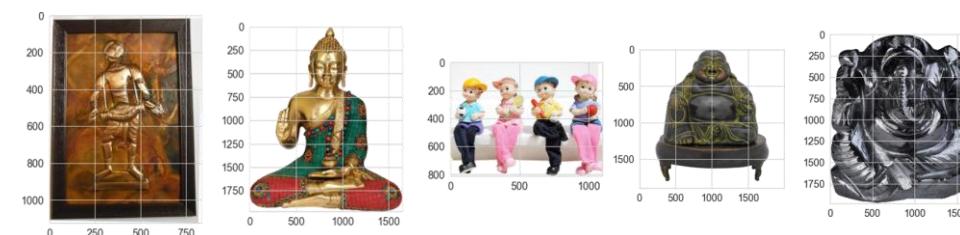
Beauty and Personal Care



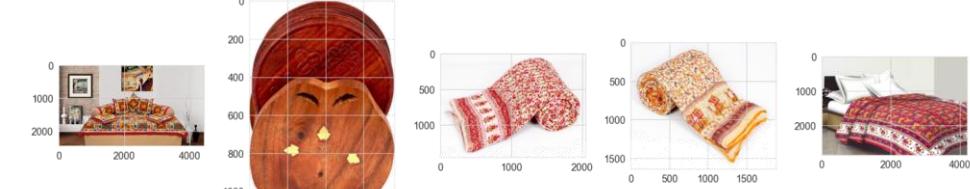
Computers



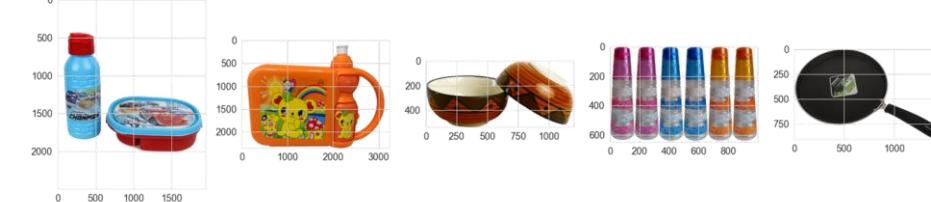
Home Decor & Festive Needs



Home Furnishing



Kitchen & Dining

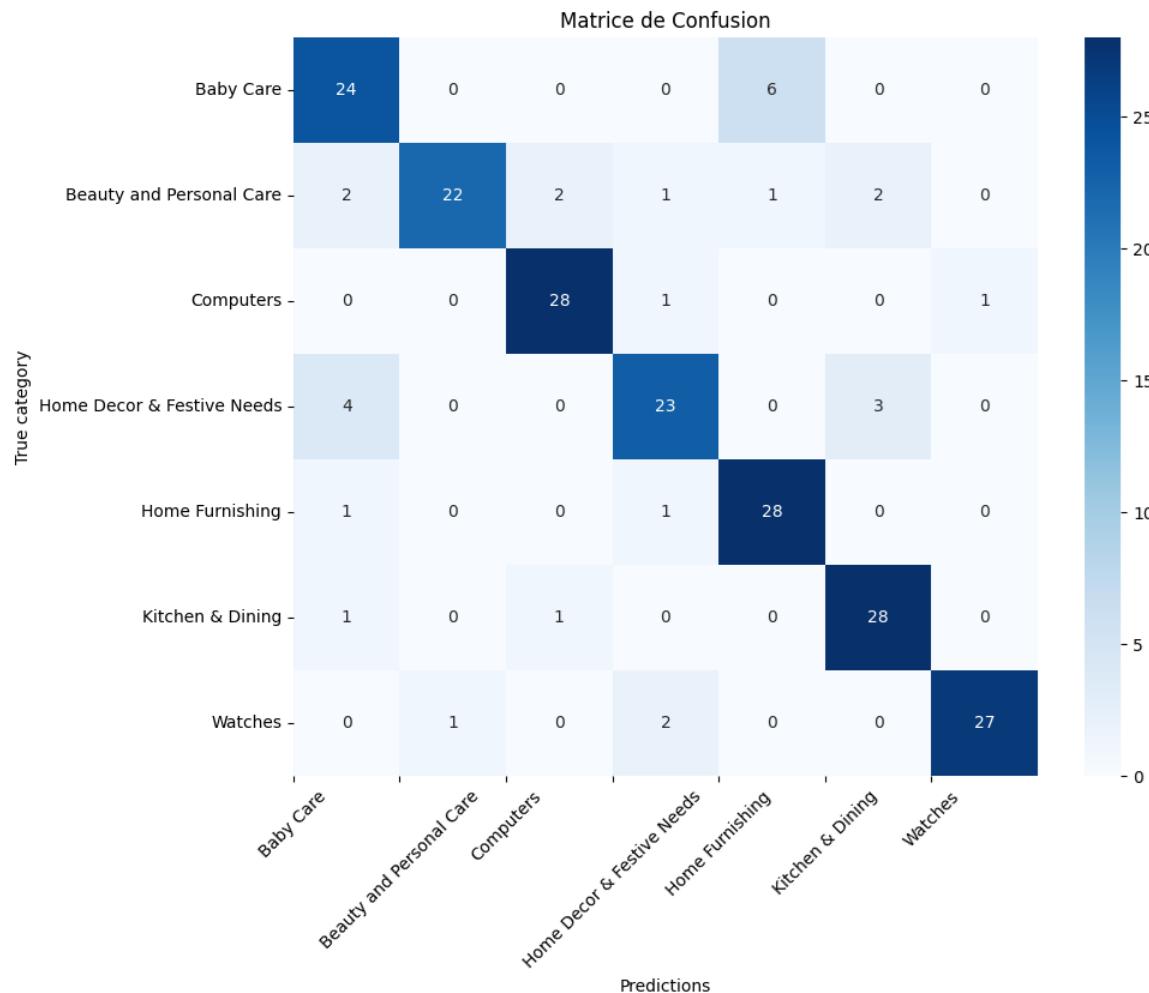


Watches

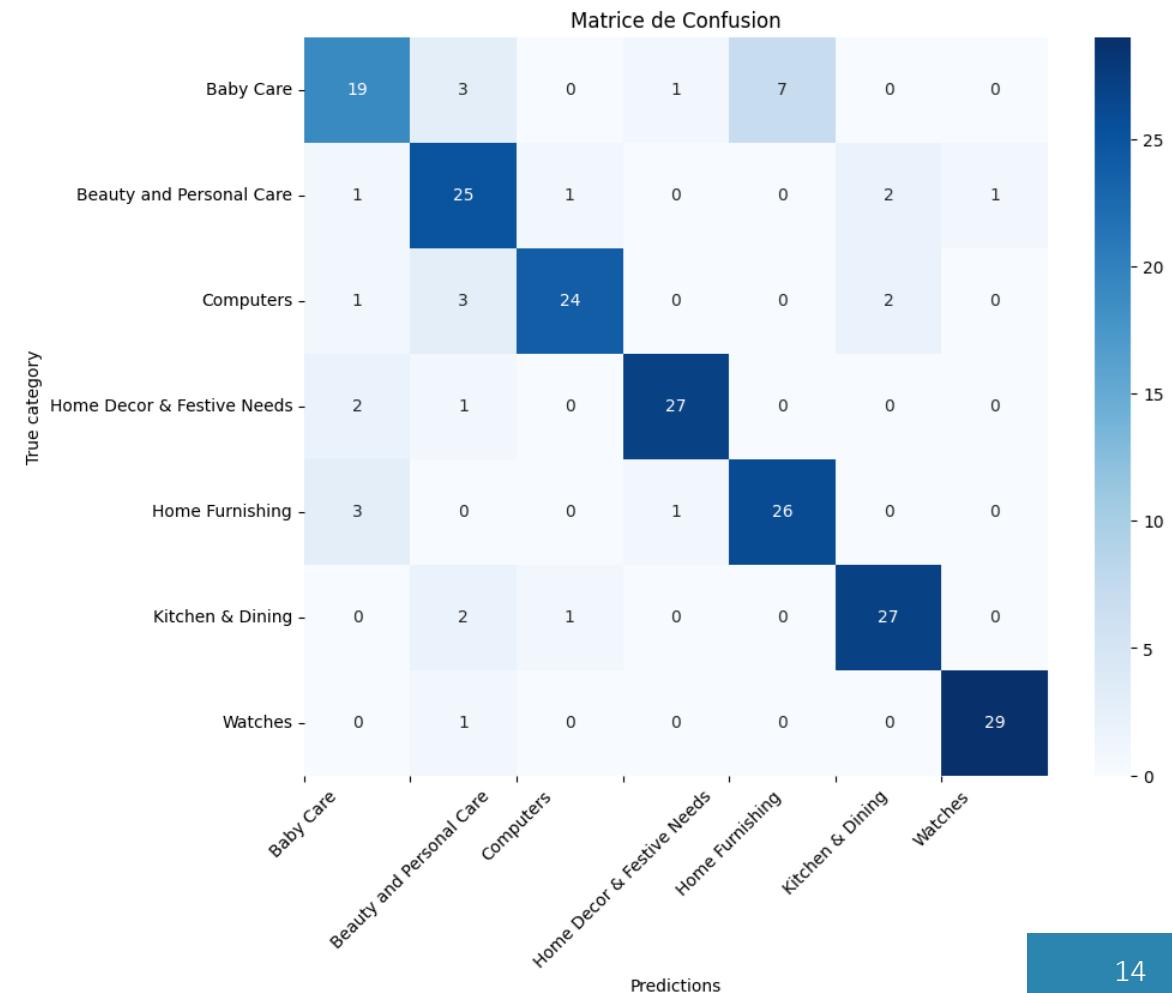


CLIP Zero-Shot : des résultats très proches de DenseNet

DenseNet201 avec Data Augmentation



CLIP avec description des catégories



	Model	Precision / recall - Baby Care	Precision / recall - Beauty and Personal Care	Precision / recall - Computers	Precision / recall - Home Decor & Festive Needs	Precision / recall - Home Furnishing	Precision / recall - Kitchen & Dining	Precision / recall - Watches	
0	DenseNet avec data augmentation	0.75 / 0.80	0.96 / 0.73	0.90 / 0.93	0.82 / 0.77	0.80 / 0.93	0.85 / 0.93	0.96 / 0.90	
1	CLIP zero-shot learning, short text	0.81 / 0.43	0.46 / 0.87	0.95 / 0.60	0.38 / 0.10	0.34 / 0.83	1.00 / 0.07	0.83 / 1.00	
2	CLIP zero-shot learning, detailed text	0.73 / 0.63	0.71 / 0.83	0.92 / 0.80	0.93 / 0.90	0.79 / 0.87	0.87 / 0.90	0.97 / 0.97	
	Model	Accuracy	f1score - Baby Care	f1score - Beauty and Personal Care	f1score - Computers	f1score - Home Decor & Festive Needs	f1score - Home Furnishing	f1score - Kitchen & Dining	
0	DenseNet avec data augmentation	0.857	0.77	0.83	0.92	0.79	0.86	0.89	0.93
1	CLIP zero-shot learning, short text	0.557	0.57	0.60	0.73	0.16	0.49	0.12	0.91
2	CLIP zero-shot learning, detailed text	0.843	0.68	0.77	0.86	0.92	0.83	0.89	0.97