

Présentation de l'appel à projets

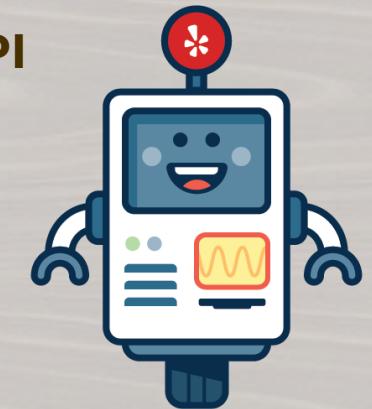
Téléchargement des données



Bases de données volumineuses



API



Labelliser automatiquement les photos postées

Environnement de travail



Librairies python spécialisées importées :

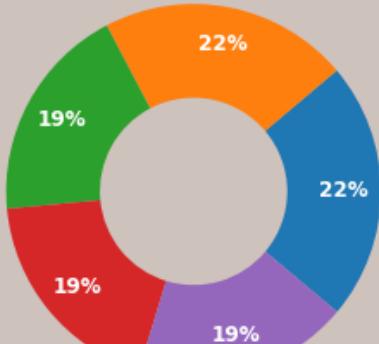


- Scikit-learn
- OpenCV
- SIFT
- Keras
- TensorFlow

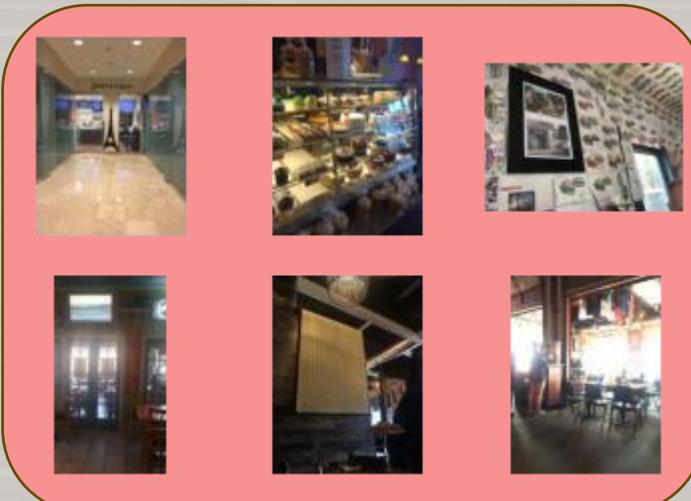
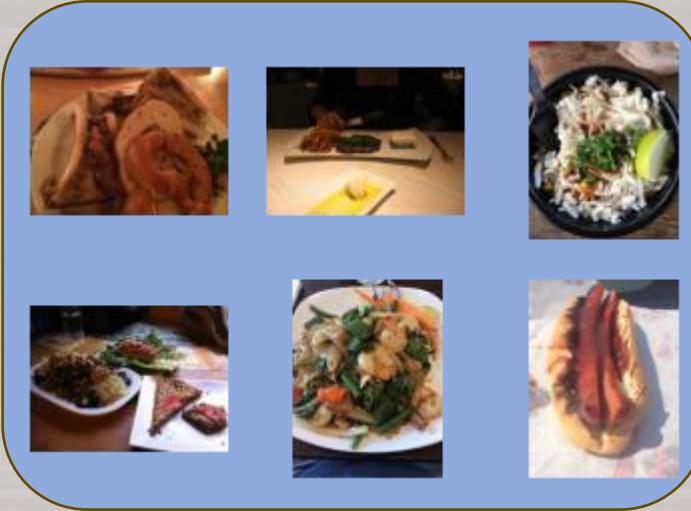


Classification des photos par label

Répartition des labels au sein de la base de données



food
menu
outside
interior
drink



- Répartition des labels légèrement déséquilibrée → Rééquilibrage
- Dimension, exposition, contraste différents → Pré-traitement des images

Pré-traitement des photos : Transformations

Photo brute



Format de la photo : RGB

Dimensions de la photo :

L : 600 px, H : 358 px

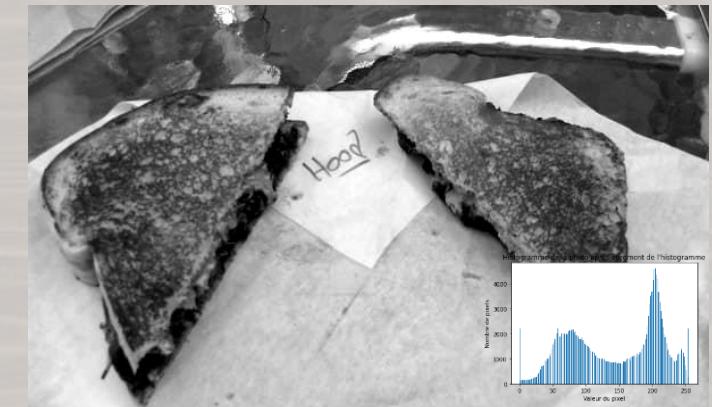
Conversion de l'image



Format de la photo :

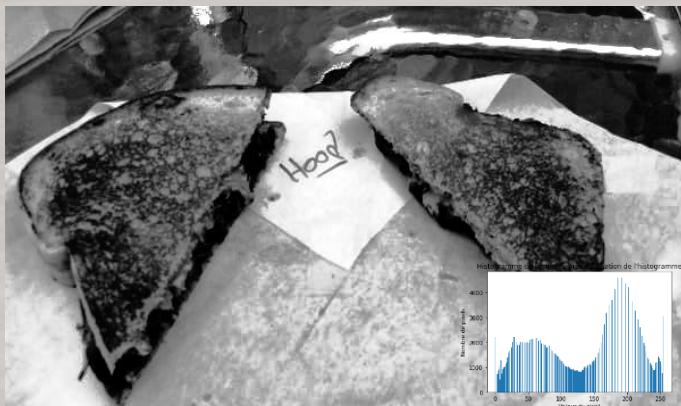
L (niveaux de gris)

Etirement de l'histogramme



Correction de l'exposition

Egalisation de l'histogramme



Correction du contraste

Atténuation du bruit additif



Filtre linéaire : Gaussien

Atténuation du bruit impulsionnel

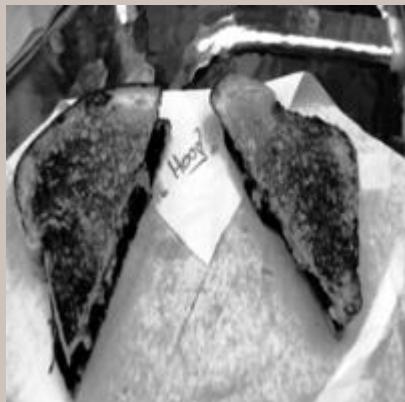


Filtre non-linéaire : Médian

Pré-traitement des photos : Pipeline



Pipeline
Transformation
Redimensionnement



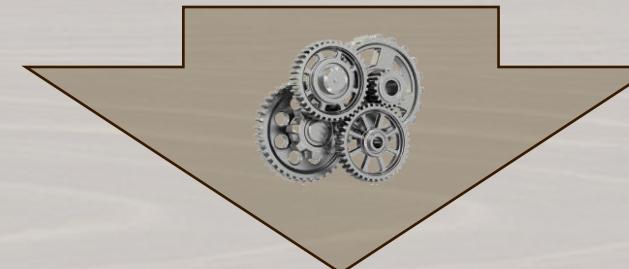
Format de la photo : L (niveaux de gris)

Dimensions de la photo :

L : 200 px, H : 200 px

Matrice 200 * 200 !!

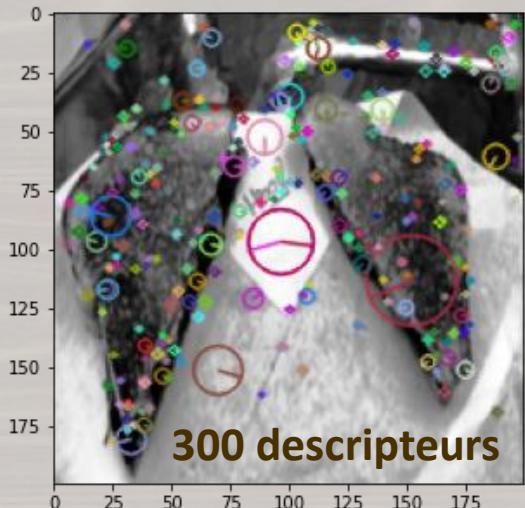
```
[[144, 142, 140, ..., 85, 39, 31],  
 [145, 144, 142, ..., 88, 45, 32],  
 [147, 146, 144, ..., 83, 57, 33],  
 ...  
 [162, 168, 177, ..., 171, 170, 163],  
 [161, 160, 168, ..., 170, 162, 156],  
 [167, 161, 165, ..., 173, 159, 153]]
```



Ensemble de vecteurs

Pré-traitement des photos : SIFT

Recherche des points clés et ses descripteurs

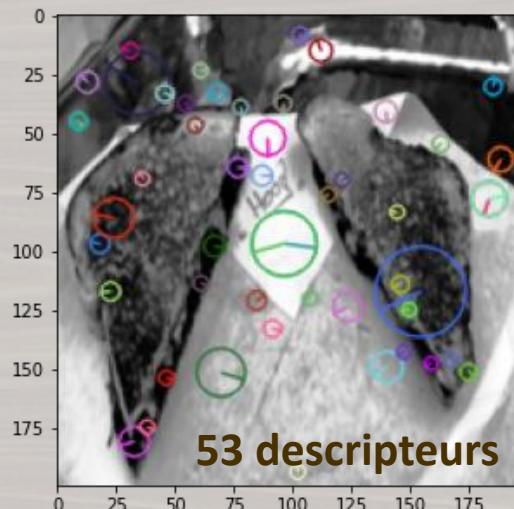


descripteurs
(vecteurs longueur 128 :
16 histogrammes de 8 valeurs)

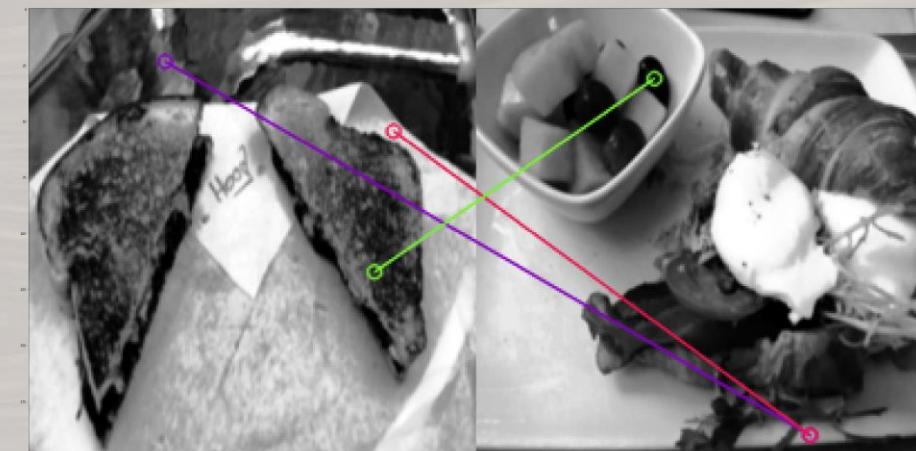
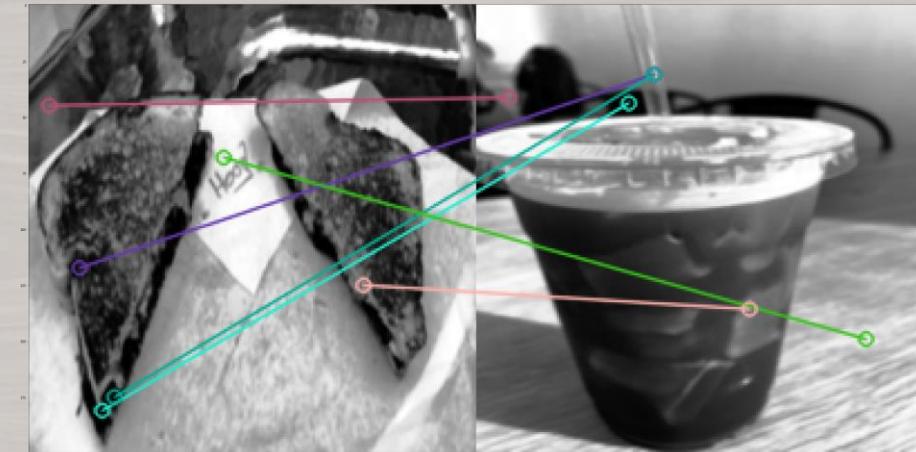
[[0., 0., 0., ..., 0., 23., 140.],
[1., 0., 0., ..., 10., 34., 139.],
...,
[27., 0., 0., ..., 2., 0., 0.],
[11., 12., 4., ..., 1., 0., 1.]]

↓ nombre descripteurs

- Filtre gaussien de variance σ estompe les détails $< \sigma$
- Choix du nombre des descripteurs SIFT_create(nb_desc)



Comparaison des descripteurs



Pré-traitement des photos : Bags of visual words et histogrammes

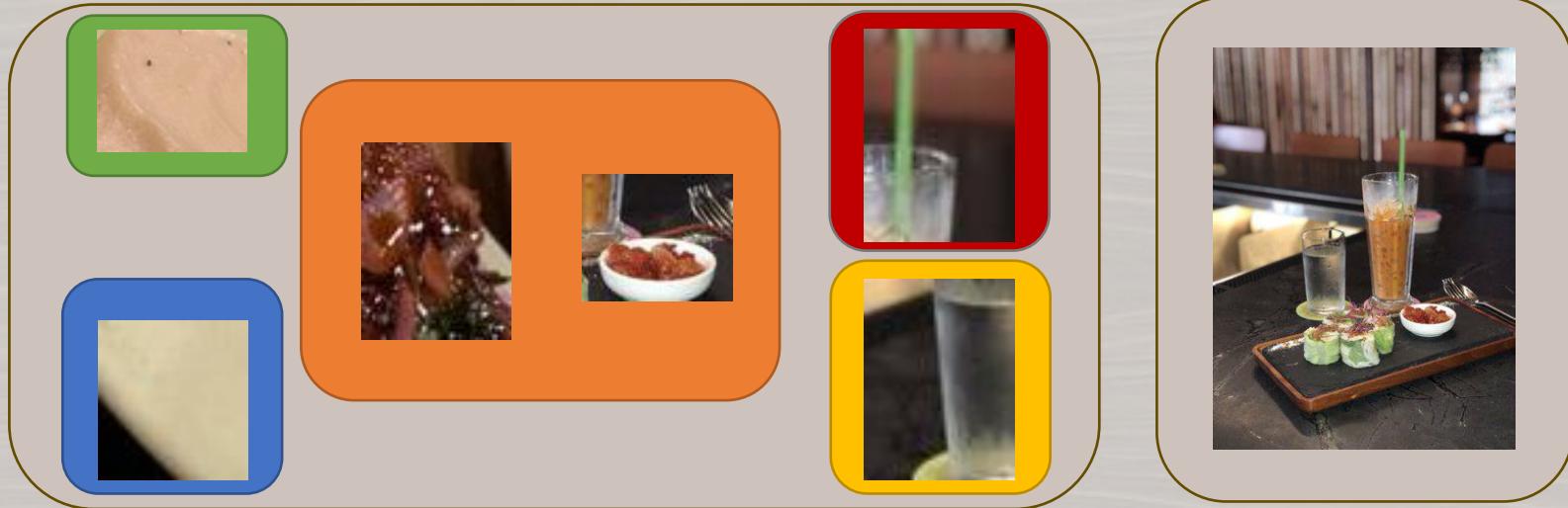
Liste des descripteurs pour 1 image

≈ 50 descripteurs par image



Liste de tous les descripteurs

≈ 815 (photos) * 50 descripteurs ≈ 40 750 descripteurs

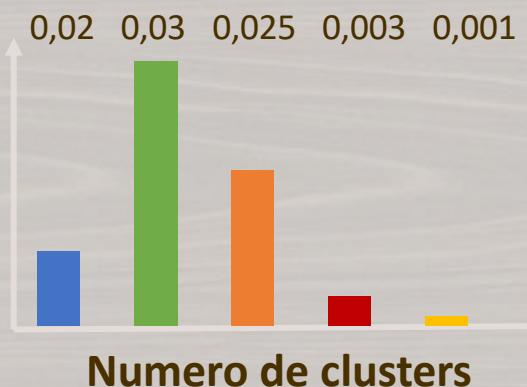


Segmentation en 5 clusters de descripteurs

1 cluster = groupe de descripteurs qui se ressemblent = 1 **BAG OF VISUAL WORDS**

Pour chaque image : pour chaque descripteur → prédition du cluster (K-means)

Pour chaque image → Histogramme



[[0., 0., 0., ..., 0., 23., 140.], ... [11., 12., 4., ..., 1., 0., 1.]]

$$\frac{\text{Nb descripteurs dans le cluster}}{\text{Nb total de descripteurs}}$$

[0,02, ... 0,001]

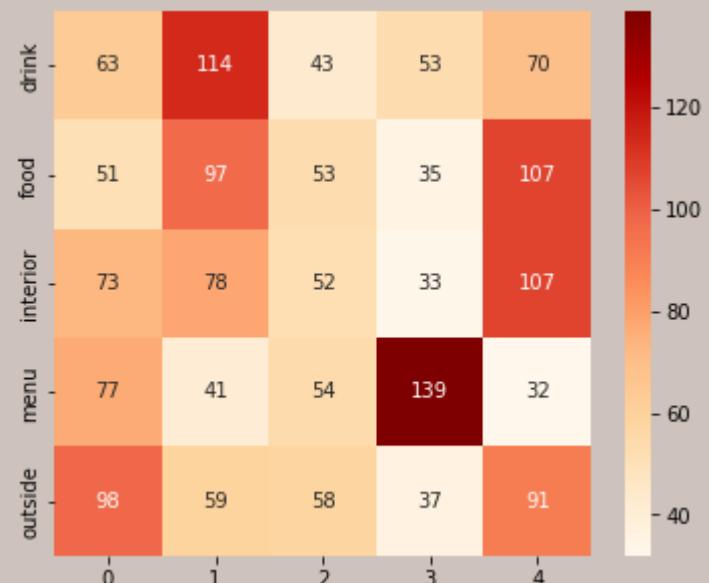


Pré-traitement des photos : Réduction de dimension et visualisation

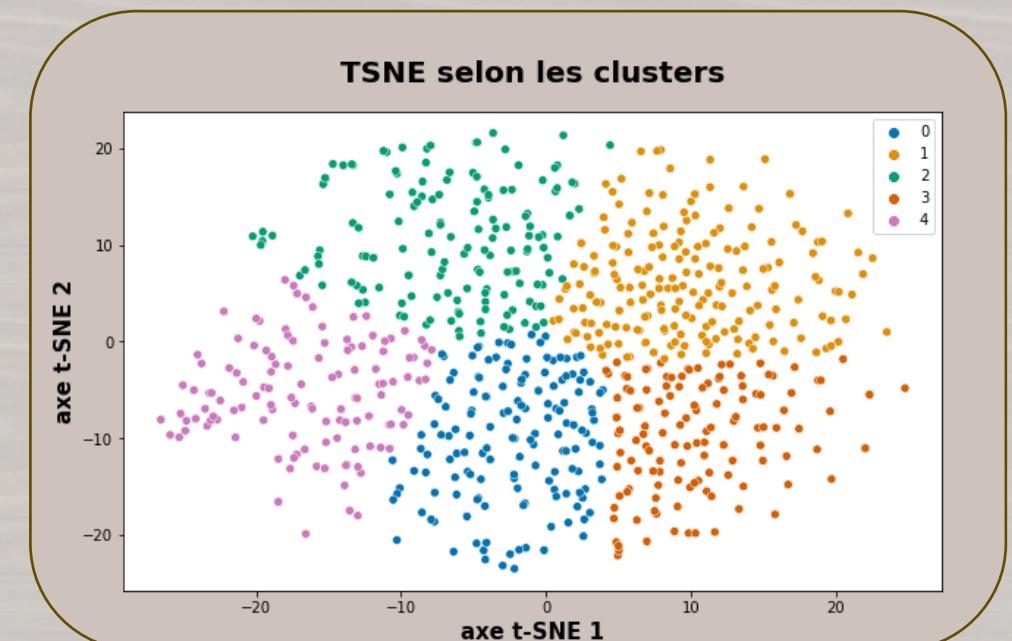
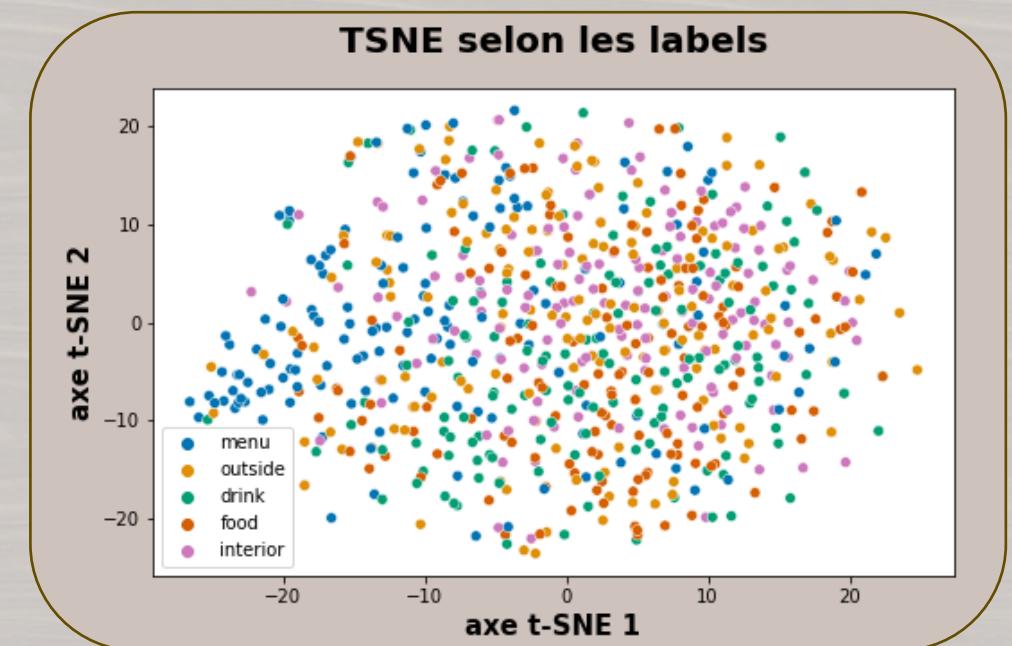
≈ 200 clusters donc 200 variables !!

Dimensions avant l'ACP : (815, 197)

Dimensions après l'ACP : (815, 107)



K-means
(données t-SNE)
ARI : 0.04

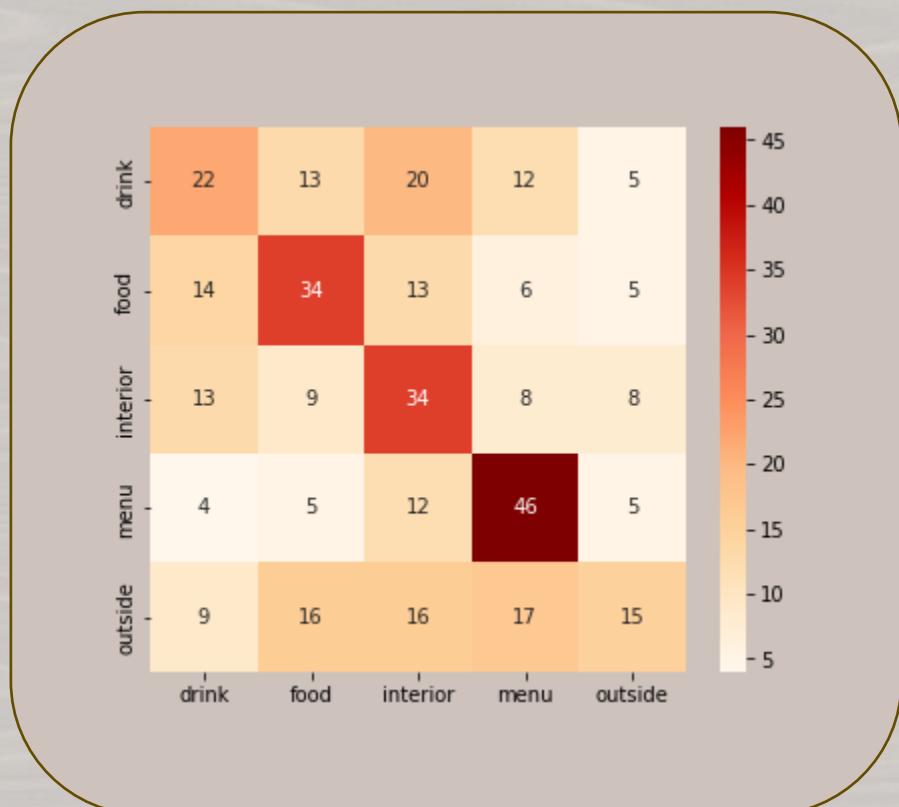


Classification des photos : Classifieur classique

Régression Logistique

Précision = 0,44

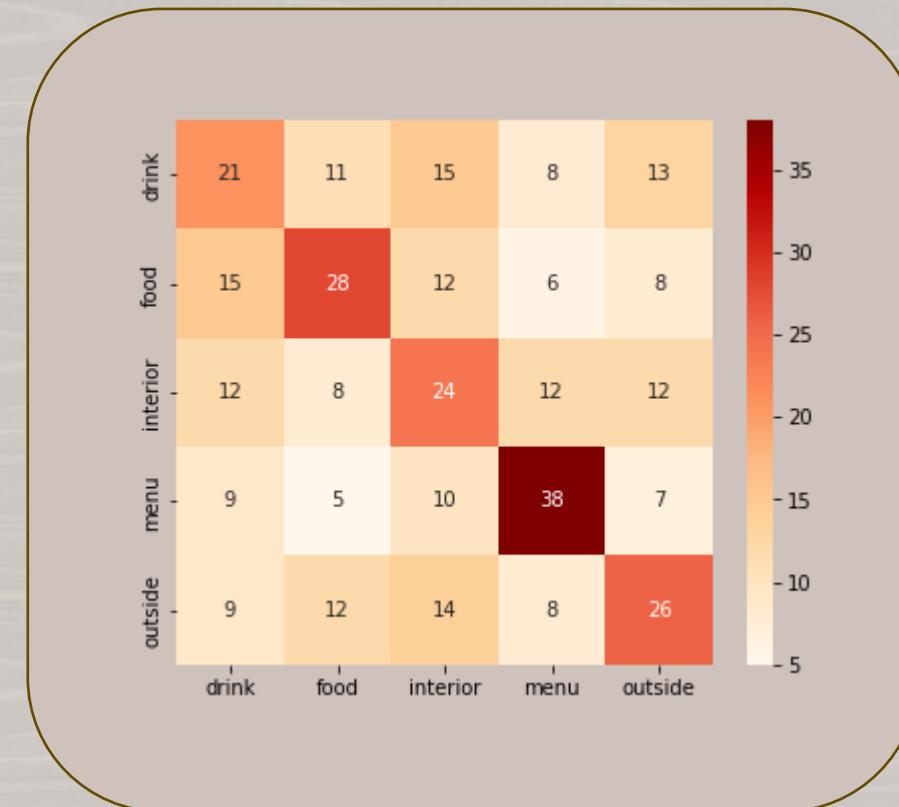
Matrice de confusion



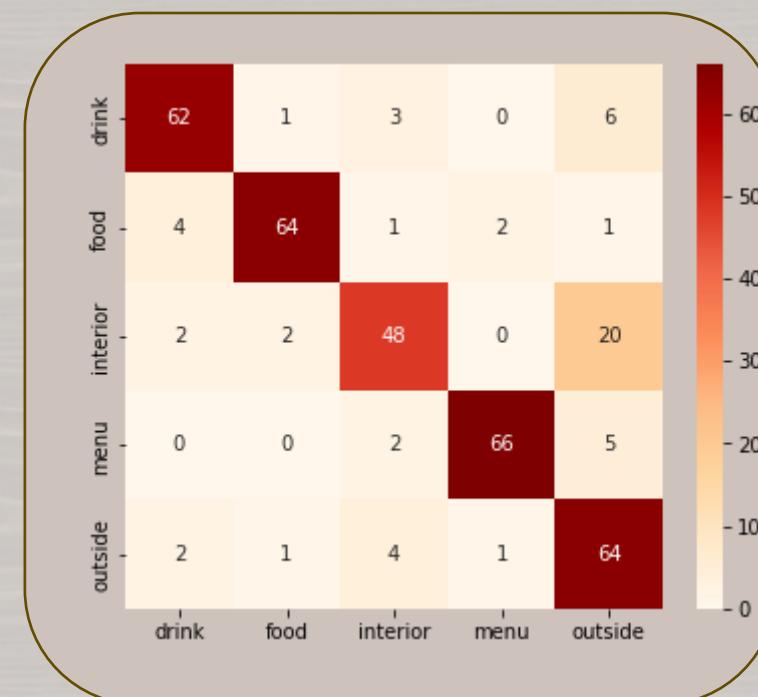
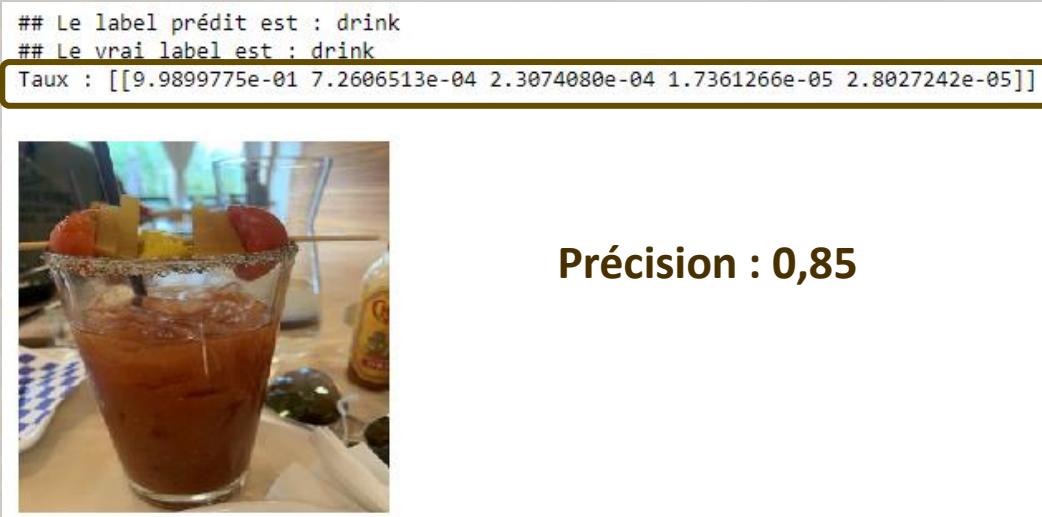
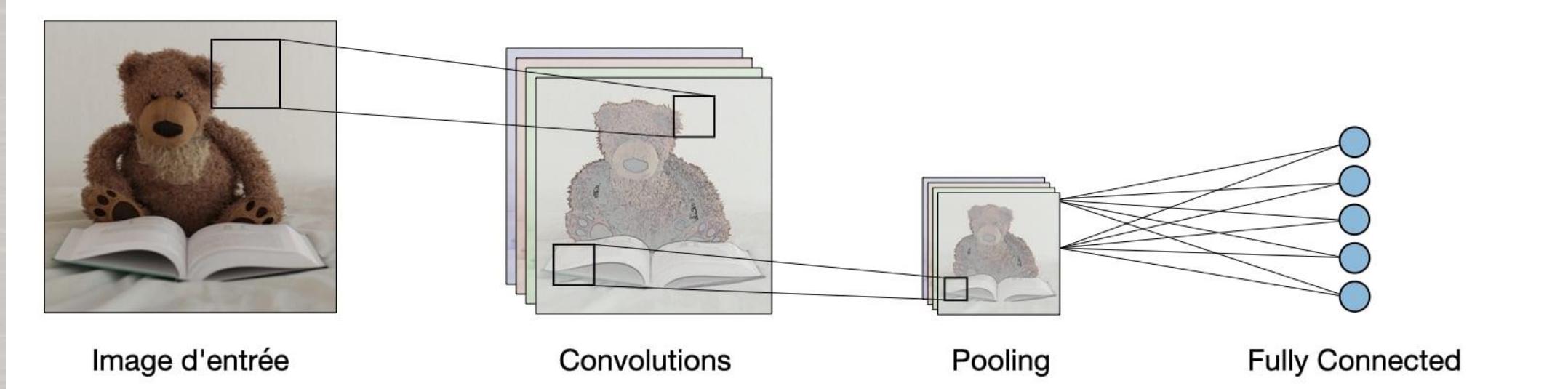
Multi-Layer Perceptron

Précision = 0,33

Matrice de confusion



Classification des photos : Réseau Neuronal Convolutif



Compétences



Pré-traitement expliqués et automatisés
Etablissement des Bags of Visual Words
Détermination des features des images



Réduction de dimensions (PCA)
Visualisation des données de grande dimension (t-SNE)



Test de classification avec Régression Logistique et MLP
Résultats très intéressants avec CNN