

Classification des images

SALORD Florian, MARGUERETTAZ Marion

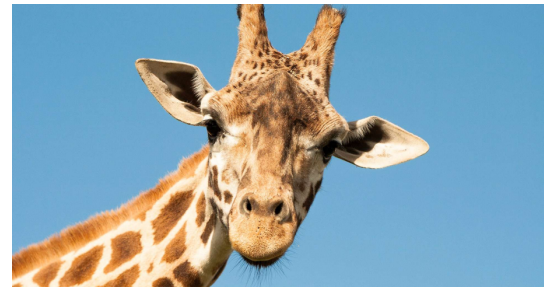
I - Différenciation des classes :

Premier essai :

Le but de notre étude est de différencier les images de girafe et de tardigrade. Pour ce faire nous allons prendre 5 images de girafe et 5 images de tardigrade.



Un tardigrade



Une girafe

La méthode utilisée est celle de la discrétisation de classes grâce à des k-means.

Pour le son nous avons utilisé les MFCC pour différencier les sons. Dans le cas des images nous allons utiliser les descripteurs des SIFT pour les différencier. Pour chaque image nous allons récupérer ces descripteurs et les stocker dans une liste de taille 128 (qui est la taille de leur vecteur).

Nous avons ensuite appliqué la méthode des k-means dessus pour essayer de classer nos images.

La classification étant complexe, les classes n'étaient pas réparties comme nous l'espérions. Nous avons donc essayé différentes techniques :

- Diminuer la résolution des images pour avoir moins de vecteurs, car la méthode va détecter moins de variation. Cela a aidé à avoir plus de différence entre tous les images mais ce n'était toujours pas probant.
- Nous avons décidé de diviser tous les bow par le nombre de bow total d'une image. Ce qui va permettre d'avoir le même nombre de descripteurs pour toutes les images. En faisant cela nous avons eu des résultats un peu plus intéressants, mais ce n'était toujours pas ce qui était attendu.

Nos résultats ne correspondant pas à nos attentes, nous avons choisi d'augmenter notre nombre d'échantillons. Malgré cela, nous n'avons obtenu aucun résultat intéressant.

Pour augmenter nos chances nous avons inclus un troisième type d'image, les voitures. Les comparaisons voiture/tardigrade ou voiture/girafe n'ont toujours pas eu de résultat acceptable.

Les voitures, tardigrades et les girafes ne donnant pas de résultats intéressants nous avons choisi de changer radicalement d'échantillons. Avec des triangles et des carrés, mais encore une fois aucun résultat, cela pouvant être dû au manque de texture des images choisies.

Étant données toutes nos expériences infructueuses nous sommes repartis des sapins et des pères Noël puisque ce sont des échantillons qui sont supposés marcher. La répartition étant celle que nous attendions, nous avons pu confirmer que le problème ne venait pas du script utilisé mais bien des échantillons.

Les résultats de nos collègues ont montré que les images avec un fond blanc marchent plus facilement. Nous avons abandonné les tardigrades car la banque de donnée est trop petite.

Deuxième essai :



Un stylo



Un chien

Le nouveau but de notre étude est de différencier les images de stylo et de chien. Pour ce faire nous allons prendre 6 images de stylo et 6 images de chien.

Résultats du premier test :

```
python3 recoimages2018.py 100 2 True
result of kmeans 1 [65 60 41 ... 91 0 58]
nb of descriptors per file : [434, 359, 543, 243, 522, 521, 231, 209, 293, 71, 93, 82]
BOWs : [[ 7.  3.  4. ...  2.  2.  8.]
 [ 4.  3.  5. ...  2.  2.  2.]
 [ 7.  3. 10. ...  4.  2.  5.]
 ...
 [ 0.  2.  0. ...  0.  1.  1.]
 [ 0.  1.  2. ...  0.  3.  0.]
 [ 1.  3.  1. ...  5.  4.  0.]]
result of kmeans 2 [0 0 0 1 0 0 1 1 1 1 1 1]
```

Les images sont passées dans l'ordre, d'abord tous les chiens puis tous les stylos. On voit que seul un chien est reconnu à tort comme un stylo.

Matrice de confusion :

		Classes estimées	
		Chiens	Stylos
Classes réelles	Chiens	5	1
	Stylos	0	6

Résultat du deuxième test :

Pour le deuxième essai nous avons enlevé le chien qui n'était pas reconnu comme tel. Ce chien était en fait un chiot. Du coup nous avons pris un chien âgé, ressemblant plus aux autres. Avec ce nouveau chien, nous avons de très bons résultats.

```
python3 recoimages2018.py 100 2 True
result of kmeans 1 [31 63 47 ... 40 72 8]
nb of descriptors per file : [434, 359, 543, 484, 522, 521, 231, 209, 293, 71, 93, 82]
BOWs : [[ 1.  3. 14. ...  2.  9.  2.]
 [ 0.  1.  6. ...  2.  1.  7.]
 [ 2. 10.  5. ...  6.  8.  7.]
 ...
 [ 0.  4.  1. ...  0.  0.  0.]
 [ 0.  2.  2. ...  0.  0.  1.]
 [ 0.  0.  0. ...  0.  0.  0.]]
result of kmeans 2 [0 0 0 0 0 0 1 1 1 1 1 1]
```

Encore une fois tous les chiens sont passés en premier puis les stylos. On voit que la classification s'est très bien passée et que chaque classe a été créée correctement.

Matrice de confusion :

		Classes estimées	
		Chiens	Stylos
Classes réelles	Chiens	6	0
	Stylos	0	6

II - Comparaison à des k-means existants :

Le but de la deuxième partie est de lire toutes les images "tests" puis de comparer de nouvelles images à cette répartition pour vérifier si les nouvelles images vont dans l'une ou l'autre des classes existantes, tout cela sans changer les k-means calculés.

En testant sur nos images de test, on obtient bien la même répartition.

En ajoutant d'autres images de chiens ou stylo, on obtient une bonne répartition dans les deux classes obtenues.

Matrice de confusion :

		Classes estimées	
		Chiens	Stylos
Classes réelles	Chiens	4	1
	Stylos	0	5

On observe qu'une seule image de chien n'est pas reconnue comme un chien. Nous avons déjà eu le problème où les chiots n'étaient pas reconnus comme des chiens, ce qui est confirmé par cet exemple. On peut donc en déduire que les bébés chiots sont des stylos.



III - Régression logistique :

Pour faire de la régression logistique nous allons utiliser des données d'apprentissage qui vont nous permettre de donner une répartition des classes fixées par l'utilisateur. Le but de cette méthode est de minimiser les erreurs.

Pour ce faire nous allons représenter toutes les images par un bow, puis nous allons lister les bow et les différents labels. Nous allons ensuite calculer le score sur la base d'apprentissage, ce qui permet de savoir si notre apprentissage est juste ou pas, et enfin nous sauvegardons les paramètres des k-means et de la régression logistique.

Nous avons un dossier d'images de tests depuis lequel nous lisons de nouvelles images, et en faisant les bows de ces images grâce aux k-means trouvés précédemment, nous pouvons prédire nos labels grâce à la régression logistique.

Au premier essai, nous avons obtenu des résultats incohérents [0,0,1,0,0,1,1,1,1,0] (5 chiens et 5 stylos). Nous avons donc chien reconnu comme un stylo et un stylo reconnu comme un chien.

En diminuant notre 1er k-mean de 100 à 50, nous avons eu de meilleurs résultats : [0,0,1,0,0,1,1,1,1,1].

Matrice de confusion avec un k-mean à 50 :

		Classes estimées	
		Chiens	Stylos
Classes réelles	Chiens	4	1
	Stylos	0	5

Au vu de nos résultats concluants, nous avons voulu essayer de nouveau notre idée de départ avec les girafes et les tardigrades par curiosité.

En donnant 5 images de girafes et 4 images de tardigrade en apprentissage, ainsi que 8 images de girafes et 5 de tardigrades en tests, nous obtenons la matrice de confusion suivante :

		Classes estimées	
		Girafe	Tardigrade
Classes réelles	Girafe	8	0
	Tardigrade	0	5

IV – Annexe

Images d'apprentissage :



Images de test :

