programming-4

May 1, 2024

Assignment 4

[19]:
```
import os
import warnings
warnings.filterwarnings("ignore")

dog_image_directory = r'C:\Users\injam\Desktop\DM_Assignment_1\Cropped'

from skimage import filters
from skimage import io, color
from skimage import exposure
import numpy as np

def calculate_angle(dx, dy):
    return np.mod(np.arctan2(dy, dx), np.pi)

hist_images = []
labels = []

for index, breed in enumerate(os.listdir(dog_image_directory)):
    image_folder = os.path.join(dog_image_directory, breed)
    for image in os.listdir(image_folder):
        img = io.imread(os.path.join(image_folder, image.strip()))
        img = color.rgb2gray(img)
        sobel_img = calculate_angle(filters.sobel_h(img), filters.sobel_v(img))
        hist, _ = exposure.histogram(sobel_img, nbins=36)
        hist_images.append(hist / np.sum(hist))
        labels.append(index)

hist_images = np.array(hist_images)
labels = np.array(labels)
```

[20]:
```
from sklearn.decomposition import PCA
model=PCA(2)
pca2 = model.fit_transform(hist_images)
```

[25]:
```
from sklearn.cluster import KMeans, BisectingKMeans, SpectralClustering
from sklearn.metrics import silhouette_score, fowlkes_mallows_score
```

```
clustering_models = {
    "KMeans Random": KMeans(init="random", n_clusters=4),
    "KMeans KMeans++": KMeans(init="k-means++", n_clusters=4),
    "BisectingKmeans": BisectingKMeans(init="random", n_clusters=4),
    "SpectralClustering": SpectralClustering(n_clusters=4)
}

for method, model in clustering_models.items():
    fitted_model = model.fit(pca2)
    print(method)
    print("fowlkes  :" + str(fowlkes_mallows_score(labels, fitted_model.
 ↪labels_)))
    print("silhouette :" + str(silhouette_score(pca2, fitted_model.labels_)))
    print("\n")
```

```
KMeans Random
fowlkes  :0.30353266800098666
silhouette :0.4185683795750631


KMeans KMeans++
fowlkes  :0.31520188962993373
silhouette :0.4135128261083564


BisectingKmeans
fowlkes  :0.333017685831437
silhouette :0.4332733726639791


SpectralClustering
fowlkes  :0.34834824513062457
silhouette :-0.029977113490016673
```

[22]:
```
from sklearn.cluster import DBSCAN
db = DBSCAN(eps=0.015, min_samples=2).fit(pca2)
# Number of clusters in labels, ignoring noise if present.
n_clusters_ = len(set(labels)) - (1 if -1 in labels else 0)
n_noise_ = list(labels).count(-1)

print("Estimated number of clusters: %d" % n_clusters_)
print("Estimated number of noise points: %d" % n_noise_)
```

```
Estimated number of clusters: 4
```

```
Estimated number of noise points: 0
```

0.1 eps=0.015, min_samples=2

[23]:
```
print("fowlkes  :" +str(fowlkes_mallows_score(labels,db.labels_)))
print("silhoutte :"+str(silhouette_score(pca2,db.labels_)))
```

```
fowlkes  :0.48807103785693184
silhoutte :0.4802708103721345
```

[24]:
```
from sklearn.cluster import AgglomerativeClustering
linkage=["ward", "complete", "average", "single" ]
for link in linkage:
    clustering=AgglomerativeClustering(n_clusters=4,linkage=link)
    pred=clustering.fit(pca2)
    print(link)
    print("folwkes :"+str(fowlkes_mallows_score(labels,pred.labels_)))
    print(" silhouette :"+str(silhouette_score(pca2,pred.labels_))+"\n")
```

```
ward
folwkes :0.3146235000818094
 silhouette :0.40256571902377786

complete
folwkes :0.40659482650984735
 silhouette :0.401907724039627

average
folwkes :0.4965405634587954
 silhouette :0.6257171216011216

single
folwkes :0.4977891732358187
 silhouette :0.6869276296961174
```

0.1.1 fowlkes score best to worst : agglomerative single,Average,dbscan,complete,spectral,bisecting,k-means++,ward,random

0.1.2 silhouette score best to worst : agglomerative single,average,dbscan,bisecting,random,k-means++,ward,complete,spectral

[ ]:
```
```