datamin4

May 1, 2024

Assignment 4

[1]:
```
import os
import warnings
warnings.filterwarnings("ignore")
from skimage import filters
from skimage import io,color
from skimage import exposure
import numpy as np
```

[ ]:
```
directory=r"C:\Users\HP\Downloads\programmingA\Dataset\Cropped"
```

[7]:
```
def angle(dx, dy):
    return np.mod(np.arctan2(dy, dx), np.pi)

hist_images = []
labels = []
for index,breed in enumerate(os.listdir(dog_images)):
    image_folder=os.path.join(dog_images,breed)
    for image in os.listdir(image_folder):
        img    = io.imread(os.path.join(image_folder,image.strip()))
        img=color.rgb2gray(img)
        sobel_img = 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)
```

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

[3]:
```
from sklearn.cluster import KMeans,BisectingKMeans,SpectralClustering
from sklearn.metrics import silhouette_score,fowlkes_mallows_score

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

```
  File "D:\Users\rakesh\anacondainstall\lib\site-
packages\joblib\externals\loky\backend\context.py", line 282, in
_count_physical_cores
    raise ValueError(f"found {cpu_count_physical} physical cores < 1")
```

```
KMeans Random
fowlkes  :0.31210384357136123
silhouette :0.42566998145918294


KMeans KMeans++
fowlkes  :0.3135500174941293
silhouette :0.4219079573583074


BisectingKmeans
fowlkes  :0.3160960680776303
silhouette :0.3246413765119647


SpectralClustering
fowlkes  :0.34462999704928804
silhouette :-0.0320470930142919
```

[8]:
```
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_)
print("fowlkes  :" +str(fowlkes_mallows_score(labels,db.labels_)))
print("silhoutte :"+str(silhouette_score(pca2,db.labels_)))
```

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

```
Estimated number of noise points: 0
fowlkes  :0.49701080553799554
silhoutte :0.4676486289516008
```

[6]:
```
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 :0.297576071811419
 silhouette :0.3733474857164391

complete
folwkes :0.4198789002063884
 silhouette :0.37694991350279805

average
folwkes :0.5018174702339943
 silhouette :0.6157525833578814

single
folwkes :0.5027533800828587
 silhouette :0.6295656109659
```

0.0.1 fowlkes score best to worst : single, average, dbscan,complete,Spectral, bisecting, kmenas++,kmean,ward

0.0.2 silhouette score best to worst :single, average, dbscan, kmeans, kmeans++, complete, ward, bisecting, Spectral

[ ]:
```
```