

```
In [1]: import pandas as pd
import numpy as np
from collections import Counter
from matplotlib import pyplot as plt
import os
import cv2
%matplotlib inline
```

```
In [2]: def load_filepaths():
    imdir_ideology = 'ideology_image_dataset/'
    ideology_files=os.listdir('ideology_image_dataset/')
    ideology_files_path=[os.path.join(imdir_ideology,file) for file in ideology_files]

    return ideology_files_path

ideology_files_path=load_filepaths()
len(ideology_files_path)
```

Out[2]: 2942

```
In [4]: def showClustering(predicted_labels,label):
    label_indexs= np.where(predicted_labels==label)[0]
    print("CLUSTER--> ",label,"TOTAL IMAGES--> ",len(label_indexs))
    if(len(label_indexs)>=500):
        fig=plt.figure(figsize=(10, 400))

    elif(len(label_indexs)>100 and len(label_indexs)<500):
        fig=plt.figure(figsize=(10, 70))
    elif(len(label_indexs)>=50 and len(label_indexs)<100):
        fig=plt.figure(figsize=(10, 30))

    elif(len(label_indexs)>=20 and len(label_indexs)<50):
        fig=plt.figure(figsize=(10, 20))

    elif(len(label_indexs)>=0 and len(label_indexs)<20):
        fig=plt.figure(figsize=(10, 8))

    for i,index in enumerate(label_indexs):

        image = cv2.imread(ideology_files_path[index])
        image= cv2.cvtColor(image, cv2.COLOR_BGR2RGB)

        columns = 4
        rows = np.ceil(len(label_indexs)/float(columns))

        fig.add_subplot(rows,columns, i+1)
        plt.imshow(image)

    plt.show()
```

```
In [5]: results_df=pd.read_csv('D://Himani-work/gsoc2020/code/image-results-updateGrid/ideology_model_vgg_16.npy-pca.csv')  
results_df
```

Out[5]:

	Unnamed: 0	n_clusters	silhouette	calinski	davies
0	HDBSCAN	665	0.084728	3.688289	1.385730
1	DBSCAN	4	0.265952	9.874167	1.795206
2	Mean-shift	7	0.316918	9.279240	0.804949
3	Optics	215	-0.120561	4.536935	1.420074

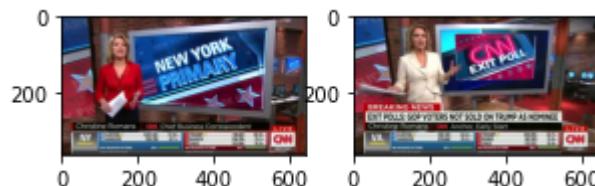
Visualizing the clusters of HDBSCAN and optics

```
In [6]: predicted_Labels=np.load('D://Himani-work/gsoc2020/code/image-results-updateGr  
id/ideology_model_vgg_16.npy-pca_hdbscan_labels.npy')  
unique_labels=set(predicted_Labels)  
for label in unique_labels:  
    if(label!=-1):  
        showClustering(predicted_Labels,label)
```

CLUSTER--> 0 TOTAL IMAGES--> 3



CLUSTER--> 1 TOTAL IMAGES--> 2



CLUSTER--> 2 TOTAL IMAGES--> 2



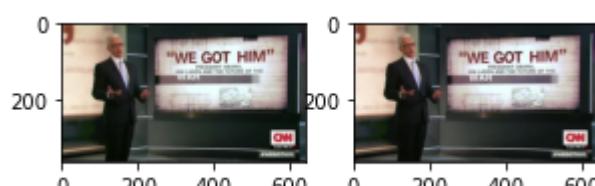
CLUSTER--> 3 TOTAL IMAGES--> 2



CLUSTER--> 4 TOTAL IMAGES--> 2



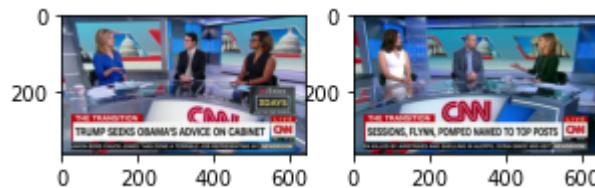
CLUSTER--> 5 TOTAL IMAGES--> 2



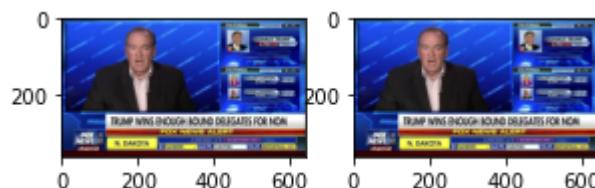
CLUSTER--> 6 TOTAL IMAGES--> 2



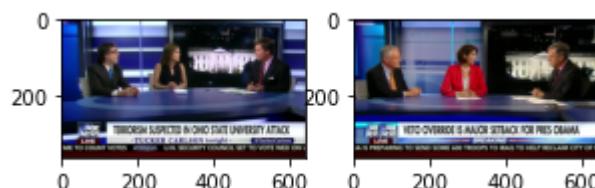
CLUSTER--> 7 TOTAL IMAGES--> 2



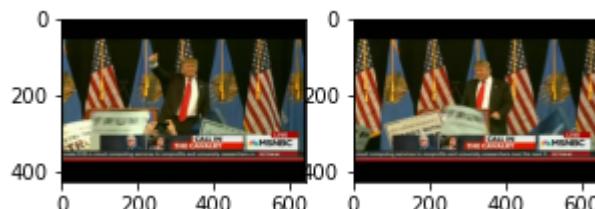
CLUSTER--> 8 TOTAL IMAGES--> 2



CLUSTER--> 9 TOTAL IMAGES--> 2



CLUSTER--> 10 TOTAL IMAGES--> 2



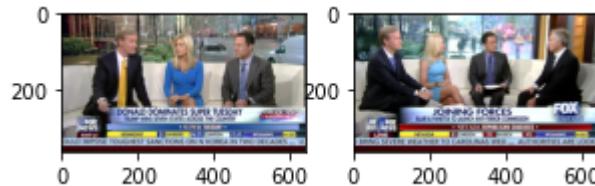
CLUSTER--> 11 TOTAL IMAGES--> 2



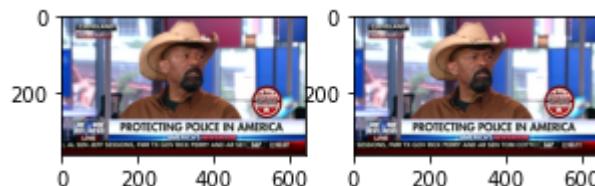
CLUSTER--> 12 TOTAL IMAGES--> 3



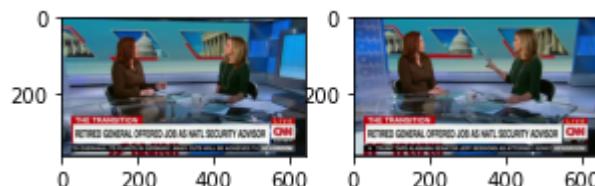
CLUSTER--> 13 TOTAL IMAGES--> 2



CLUSTER--> 14 TOTAL IMAGES--> 2



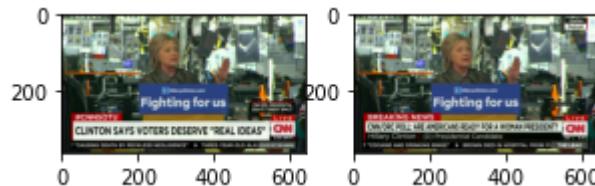
CLUSTER--> 15 TOTAL IMAGES--> 2



CLUSTER--> 16 TOTAL IMAGES--> 3



CLUSTER--> 17 TOTAL IMAGES--> 2



CLUSTER--> 18 TOTAL IMAGES--> 3



CLUSTER--> 19 TOTAL IMAGES--> 3



CLUSTER--> 20 TOTAL IMAGES--> 2



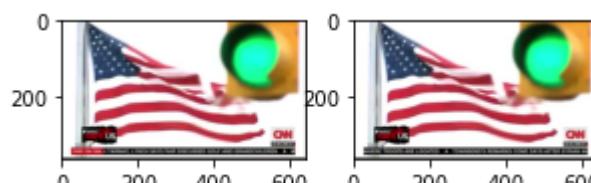
CLUSTER--> 21 TOTAL IMAGES--> 2



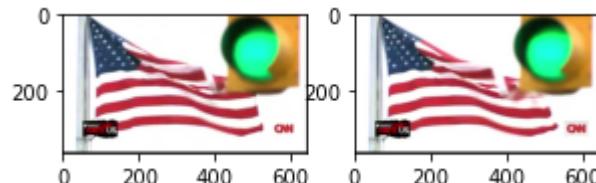
CLUSTER--> 22 TOTAL IMAGES--> 2



CLUSTER--> 23 TOTAL IMAGES--> 2



CLUSTER--> 24 TOTAL IMAGES--> 2



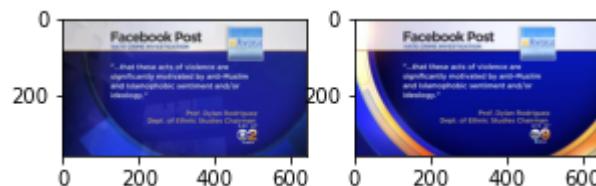
CLUSTER--> 25 TOTAL IMAGES--> 3



CLUSTER--> 26 TOTAL IMAGES--> 2



CLUSTER--> 27 TOTAL IMAGES--> 2



CLUSTER--> 28 TOTAL IMAGES--> 2



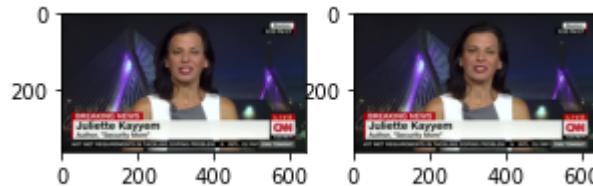
CLUSTER--> 29 TOTAL IMAGES--> 3



CLUSTER--> 30 TOTAL IMAGES--> 3



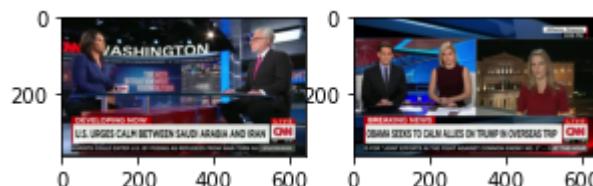
CLUSTER--> 31 TOTAL IMAGES--> 2



CLUSTER--> 32 TOTAL IMAGES--> 2



CLUSTER--> 33 TOTAL IMAGES--> 2



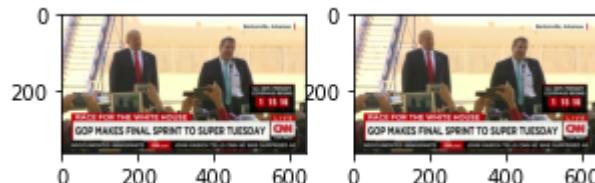
CLUSTER--> 34 TOTAL IMAGES--> 2



CLUSTER--> 35 TOTAL IMAGES--> 2



CLUSTER--> 36 TOTAL IMAGES--> 2



CLUSTER--> 37 TOTAL IMAGES--> 2



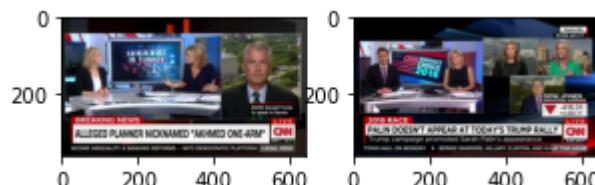
CLUSTER--> 38 TOTAL IMAGES--> 2



CLUSTER--> 39 TOTAL IMAGES--> 2



CLUSTER--> 40 TOTAL IMAGES--> 2



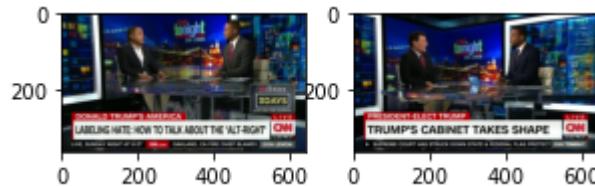
CLUSTER--> 41 TOTAL IMAGES--> 3



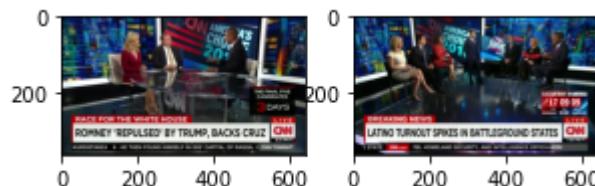
CLUSTER--> 42 TOTAL IMAGES--> 3



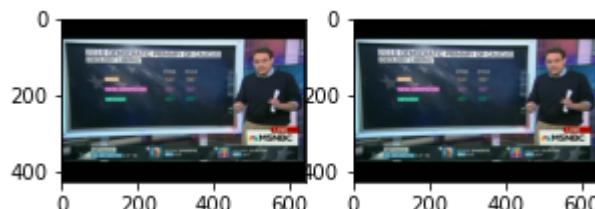
CLUSTER--> 43 TOTAL IMAGES--> 2



CLUSTER--> 44 TOTAL IMAGES--> 2



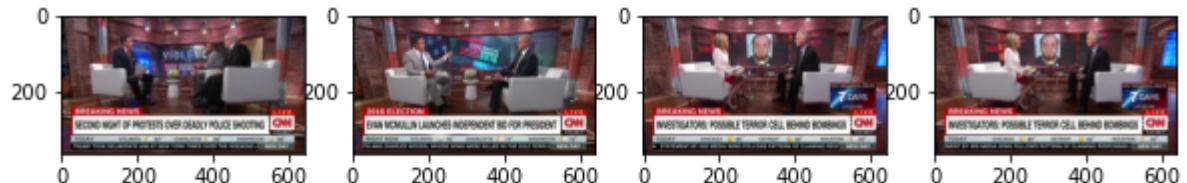
CLUSTER--> 45 TOTAL IMAGES--> 2



CLUSTER--> 46 TOTAL IMAGES--> 3



CLUSTER--> 47 TOTAL IMAGES--> 5



CLUSTER--> 48 TOTAL IMAGES--> 2



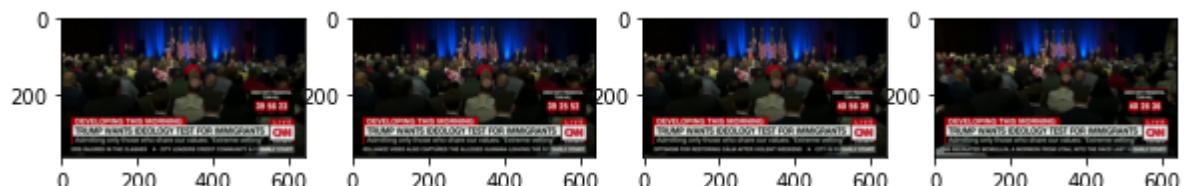
CLUSTER--> 49 TOTAL IMAGES--> 2



CLUSTER--> 50 TOTAL IMAGES--> 7



CLUSTER--> 51 TOTAL IMAGES--> 4



CLUSTER--> 52 TOTAL IMAGES--> 2



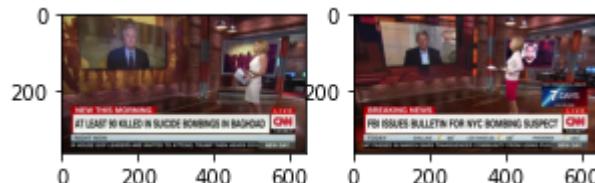
CLUSTER--> 53 TOTAL IMAGES--> 2



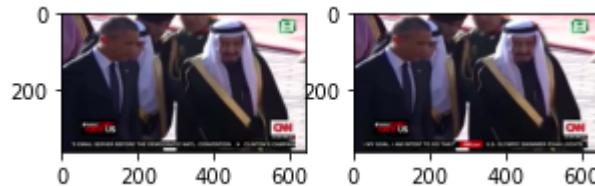
CLUSTER--> 54 TOTAL IMAGES--> 3



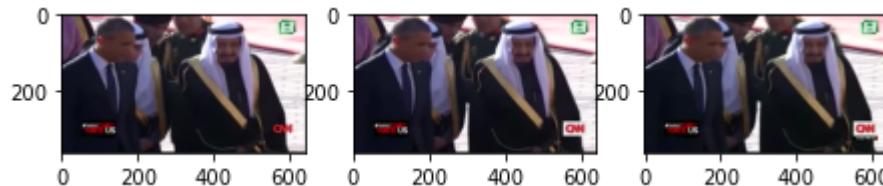
CLUSTER--> 55 TOTAL IMAGES--> 2



CLUSTER--> 56 TOTAL IMAGES--> 2



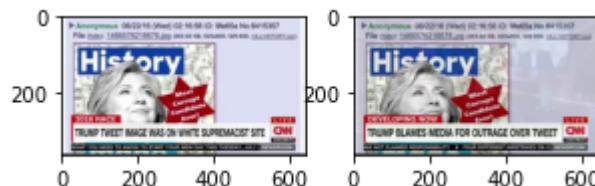
CLUSTER--> 57 TOTAL IMAGES--> 3



CLUSTER--> 58 TOTAL IMAGES--> 2



CLUSTER--> 59 TOTAL IMAGES--> 2



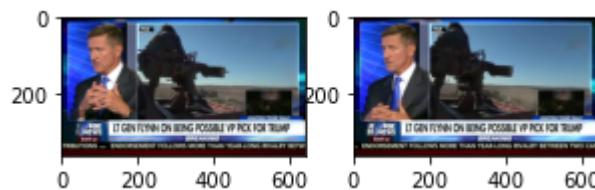
CLUSTER--> 60 TOTAL IMAGES--> 3



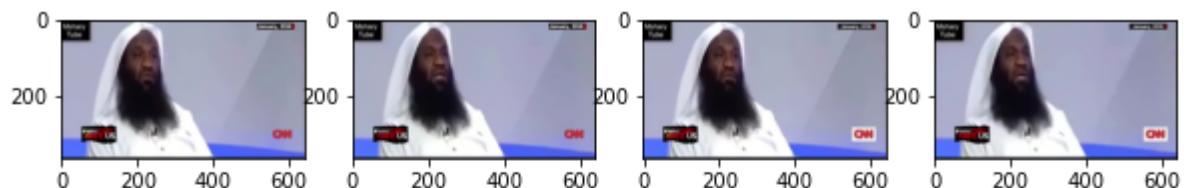
CLUSTER--> 61 TOTAL IMAGES--> 2



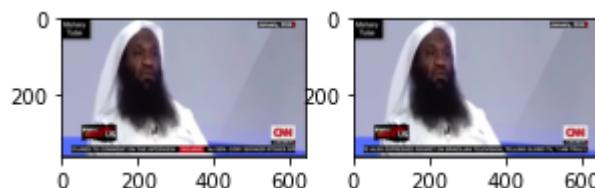
CLUSTER--> 62 TOTAL IMAGES--> 2



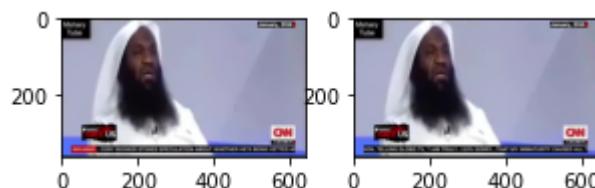
CLUSTER--> 63 TOTAL IMAGES--> 7



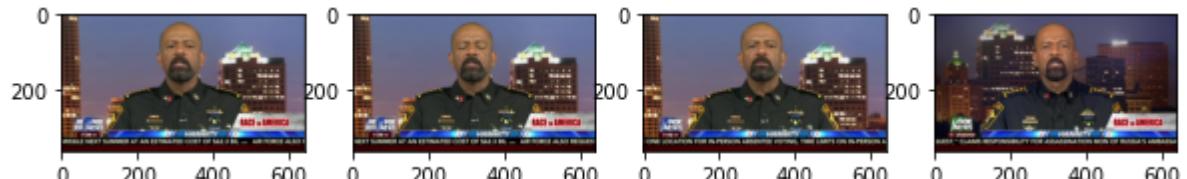
CLUSTER--> 64 TOTAL IMAGES--> 2



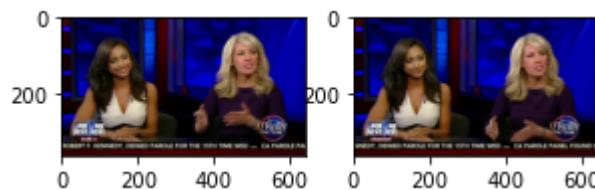
CLUSTER--> 65 TOTAL IMAGES--> 2



CLUSTER--> 66 TOTAL IMAGES--> 4



CLUSTER--> 67 TOTAL IMAGES--> 2



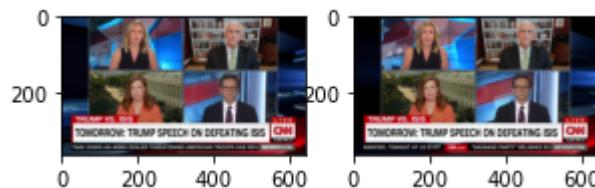
CLUSTER--> 68 TOTAL IMAGES--> 4



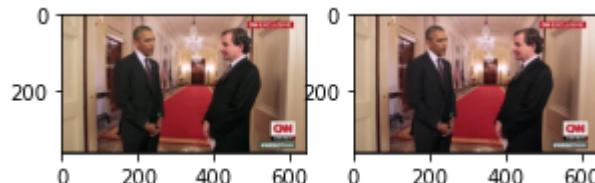
CLUSTER--> 69 TOTAL IMAGES--> 5



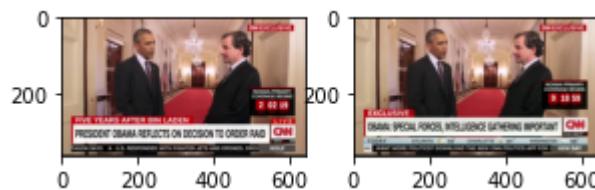
CLUSTER--> 70 TOTAL IMAGES--> 2



CLUSTER--> 71 TOTAL IMAGES--> 2



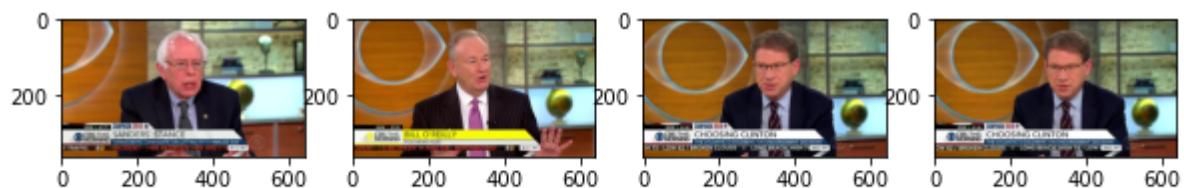
CLUSTER--> 72 TOTAL IMAGES--> 2



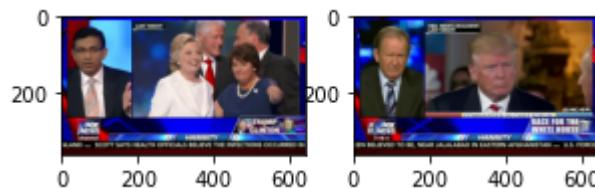
CLUSTER--> 73 TOTAL IMAGES--> 2



CLUSTER--> 74 TOTAL IMAGES--> 5



CLUSTER--> 75 TOTAL IMAGES--> 2



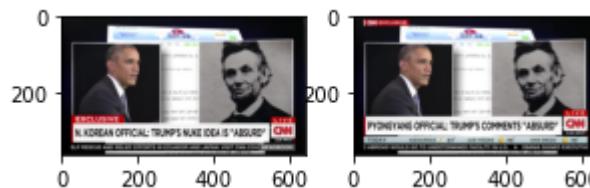
CLUSTER--> 76 TOTAL IMAGES--> 3



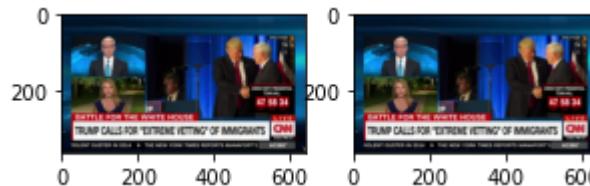
CLUSTER--> 77 TOTAL IMAGES--> 2



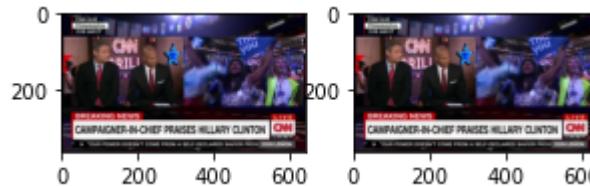
CLUSTER--> 78 TOTAL IMAGES--> 2



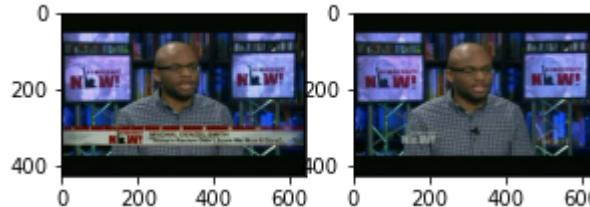
CLUSTER--> 79 TOTAL IMAGES--> 2



CLUSTER--> 80 TOTAL IMAGES--> 2



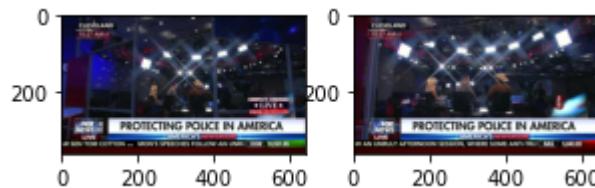
CLUSTER--> 81 TOTAL IMAGES--> 2



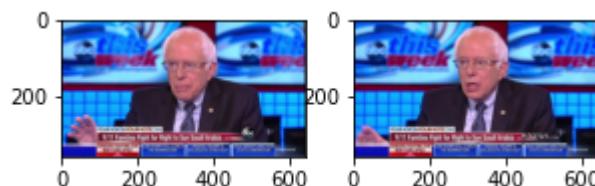
CLUSTER--> 82 TOTAL IMAGES--> 2



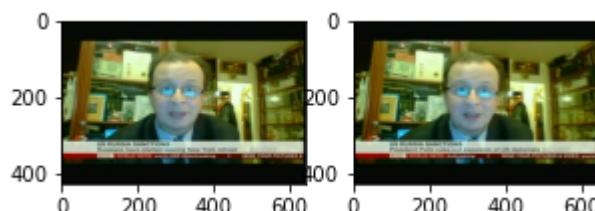
CLUSTER--> 83 TOTAL IMAGES--> 2



CLUSTER--> 84 TOTAL IMAGES--> 2



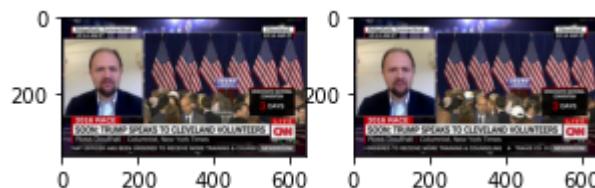
CLUSTER--> 85 TOTAL IMAGES--> 2



CLUSTER--> 86 TOTAL IMAGES--> 3



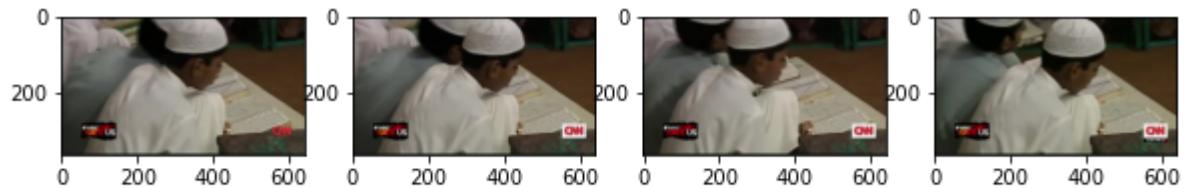
CLUSTER--> 87 TOTAL IMAGES--> 2



CLUSTER--> 88 TOTAL IMAGES--> 3



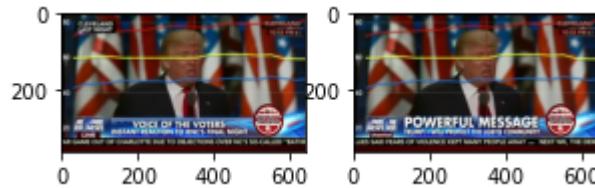
CLUSTER--> 89 TOTAL IMAGES--> 5



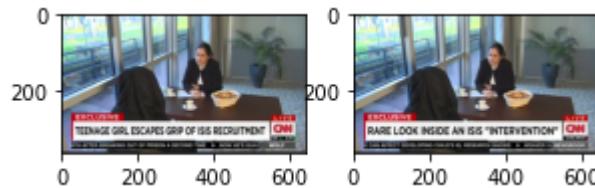
CLUSTER--> 90 TOTAL IMAGES--> 3



CLUSTER--> 91 TOTAL IMAGES--> 2



CLUSTER--> 92 TOTAL IMAGES--> 2



CLUSTER--> 93 TOTAL IMAGES--> 2



CLUSTER--> 94 TOTAL IMAGES--> 2



CLUSTER--> 95 TOTAL IMAGES--> 3



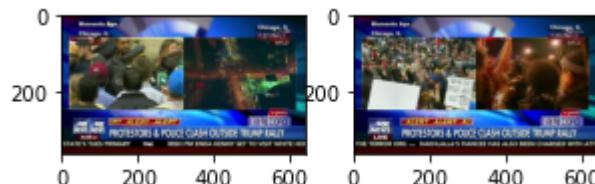
CLUSTER--> 96 TOTAL IMAGES--> 3



CLUSTER--> 97 TOTAL IMAGES--> 4



CLUSTER--> 98 TOTAL IMAGES--> 2



CLUSTER--> 99 TOTAL IMAGES--> 6



CLUSTER--> 100 TOTAL IMAGES--> 5



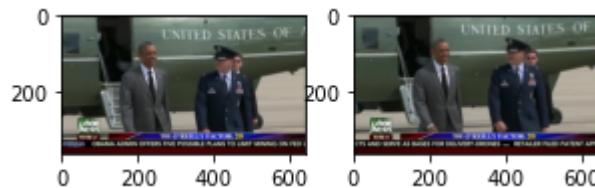
CLUSTER--> 101 TOTAL IMAGES--> 3



CLUSTER--> 102 TOTAL IMAGES--> 3



CLUSTER--> 103 TOTAL IMAGES--> 2



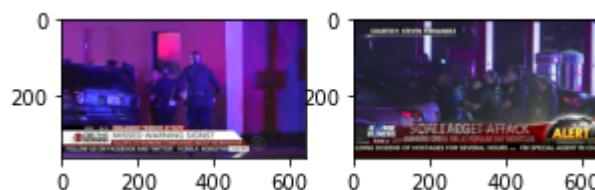
CLUSTER--> 104 TOTAL IMAGES--> 5



CLUSTER--> 105 TOTAL IMAGES--> 2



CLUSTER--> 106 TOTAL IMAGES--> 2



CLUSTER--> 107 TOTAL IMAGES--> 2



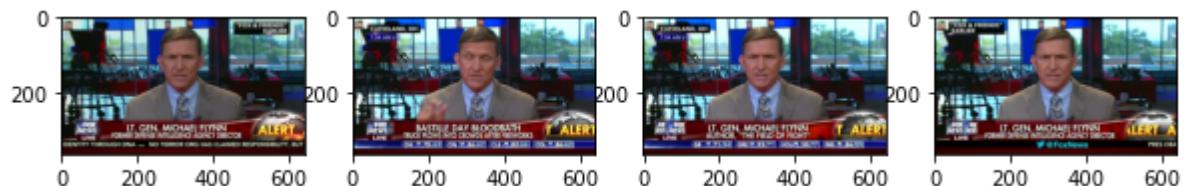
CLUSTER--> 108 TOTAL IMAGES--> 2



CLUSTER--> 109 TOTAL IMAGES--> 4



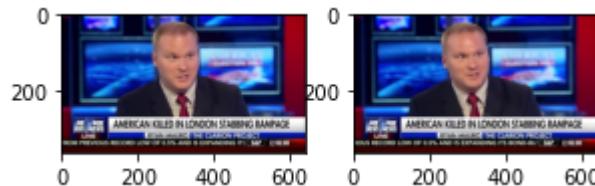
CLUSTER--> 110 TOTAL IMAGES--> 4



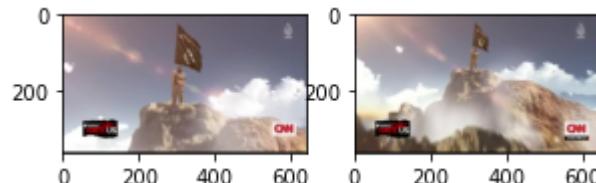
CLUSTER--> 111 TOTAL IMAGES--> 2



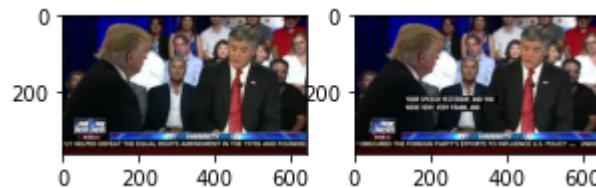
CLUSTER--> 112 TOTAL IMAGES--> 2



CLUSTER--> 113 TOTAL IMAGES--> 2



CLUSTER--> 114 TOTAL IMAGES--> 2



CLUSTER--> 115 TOTAL IMAGES--> 2



CLUSTER--> 116 TOTAL IMAGES--> 2



CLUSTER--> 117 TOTAL IMAGES--> 2



CLUSTER--> 118 TOTAL IMAGES--> 2



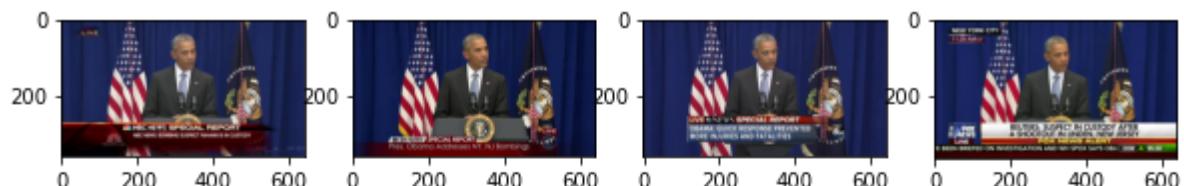
CLUSTER--> 119 TOTAL IMAGES--> 3



CLUSTER--> 120 TOTAL IMAGES--> 3



CLUSTER--> 121 TOTAL IMAGES--> 5



CLUSTER--> 122 TOTAL IMAGES--> 2



CLUSTER--> 123 TOTAL IMAGES--> 3



CLUSTER--> 124 TOTAL IMAGES--> 2



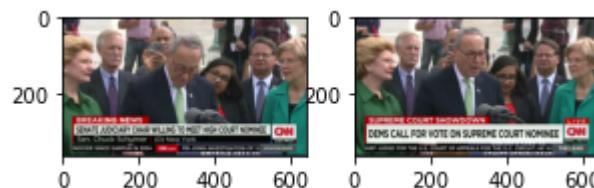
CLUSTER--> 125 TOTAL IMAGES--> 4



CLUSTER--> 126 TOTAL IMAGES--> 2



CLUSTER--> 127 TOTAL IMAGES--> 2



CLUSTER--> 128 TOTAL IMAGES--> 2



CLUSTER--> 129 TOTAL IMAGES--> 3



CLUSTER--> 130 TOTAL IMAGES--> 2



CLUSTER--&gt; 131 TOTAL IMAGES--&gt; 4



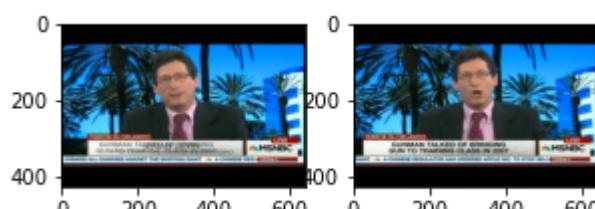
CLUSTER--&gt; 132 TOTAL IMAGES--&gt; 2



CLUSTER--&gt; 133 TOTAL IMAGES--&gt; 2



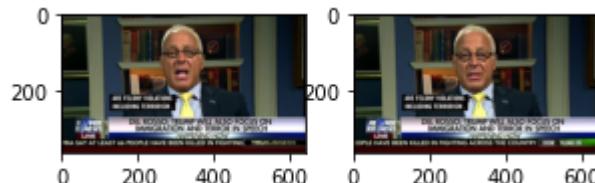
CLUSTER--&gt; 134 TOTAL IMAGES--&gt; 2



CLUSTER--&gt; 135 TOTAL IMAGES--&gt; 2



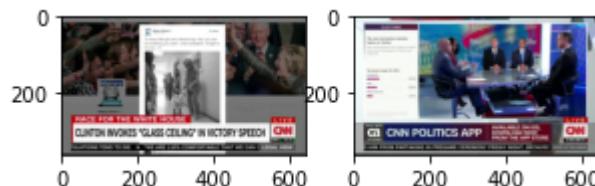
CLUSTER--&gt; 136 TOTAL IMAGES--&gt; 2



CLUSTER--> 137 TOTAL IMAGES--> 3



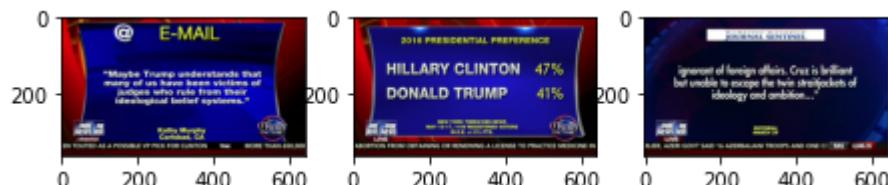
CLUSTER--> 138 TOTAL IMAGES--> 2



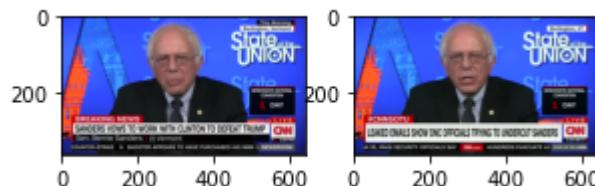
CLUSTER--> 139 TOTAL IMAGES--> 2



CLUSTER--> 140 TOTAL IMAGES--> 3



CLUSTER--> 141 TOTAL IMAGES--> 2



CLUSTER--> 142 TOTAL IMAGES--> 3



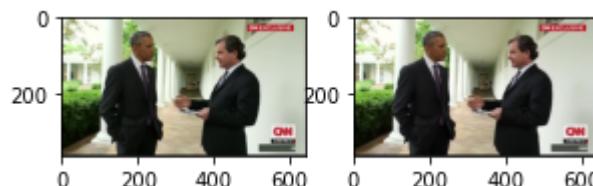
CLUSTER--> 143 TOTAL IMAGES--> 3



CLUSTER--> 144 TOTAL IMAGES--> 2



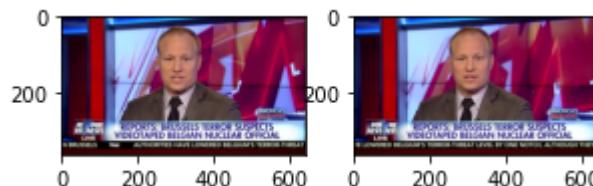
CLUSTER--> 145 TOTAL IMAGES--> 2



CLUSTER--> 146 TOTAL IMAGES--> 3



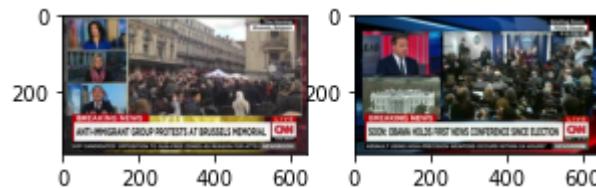
CLUSTER--> 147 TOTAL IMAGES--> 2



CLUSTER--> 148 TOTAL IMAGES--> 2



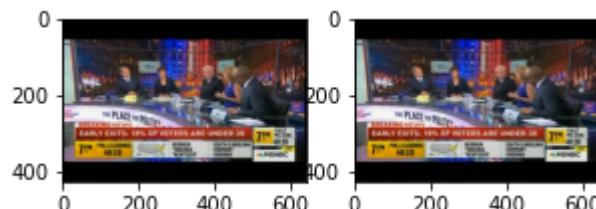
CLUSTER--> 149 TOTAL IMAGES--> 2



CLUSTER--> 150 TOTAL IMAGES--> 2



CLUSTER--> 151 TOTAL IMAGES--> 2



CLUSTER--> 152 TOTAL IMAGES--> 3



CLUSTER--> 153 TOTAL IMAGES--> 2



CLUSTER--> 154 TOTAL IMAGES--> 2



CLUSTER--> 155 TOTAL IMAGES--> 3



CLUSTER--> 156 TOTAL IMAGES--> 2



CLUSTER--> 157 TOTAL IMAGES--> 3



CLUSTER--> 158 TOTAL IMAGES--> 2



CLUSTER--> 159 TOTAL IMAGES--> 3



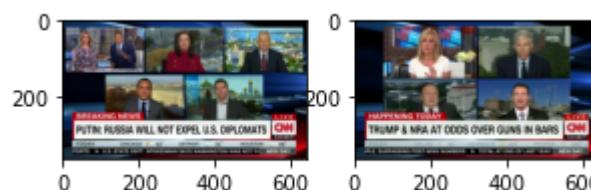
CLUSTER--> 160 TOTAL IMAGES--> 3



CLUSTER--> 161 TOTAL IMAGES--> 3



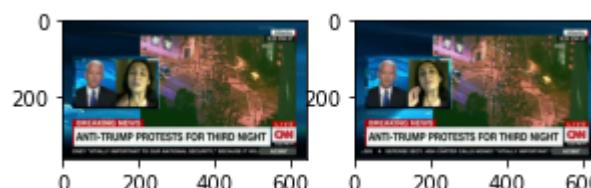
CLUSTER--> 162 TOTAL IMAGES--> 2



CLUSTER--> 163 TOTAL IMAGES--> 2



CLUSTER--> 164 TOTAL IMAGES--> 2



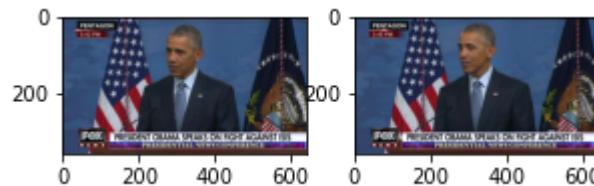
CLUSTER--> 165 TOTAL IMAGES--> 2



CLUSTER--> 166 TOTAL IMAGES--> 3



CLUSTER--> 167 TOTAL IMAGES--> 2



CLUSTER--> 168 TOTAL IMAGES--> 2



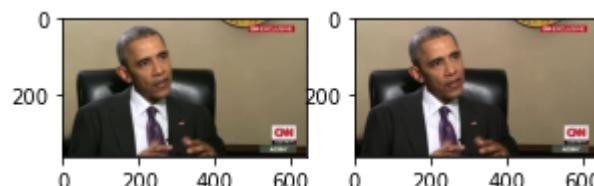
CLUSTER--> 169 TOTAL IMAGES--> 3



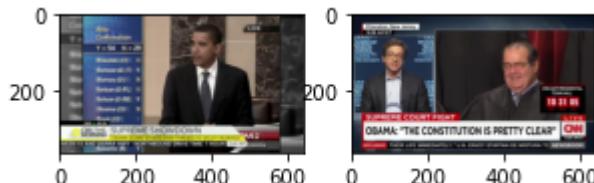
CLUSTER--> 170 TOTAL IMAGES--> 3



CLUSTER--> 171 TOTAL IMAGES--> 2



CLUSTER--> 172 TOTAL IMAGES--> 2



CLUSTER--> 173 TOTAL IMAGES--> 2



CLUSTER--> 174 TOTAL IMAGES--> 2



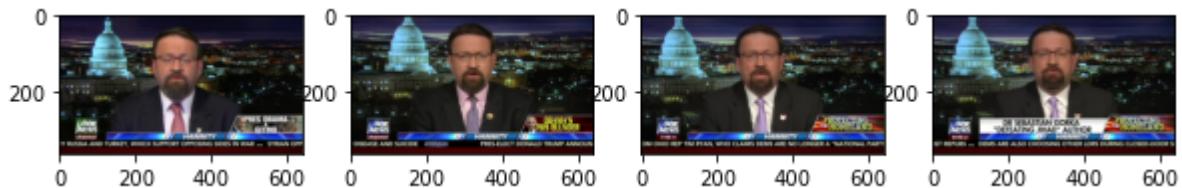
CLUSTER--> 175 TOTAL IMAGES--> 3



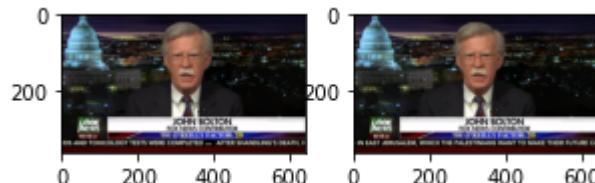
CLUSTER--> 176 TOTAL IMAGES--> 3



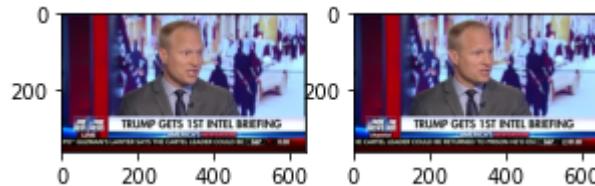
CLUSTER--> 177 TOTAL IMAGES--> 4



CLUSTER--> 178 TOTAL IMAGES--> 2



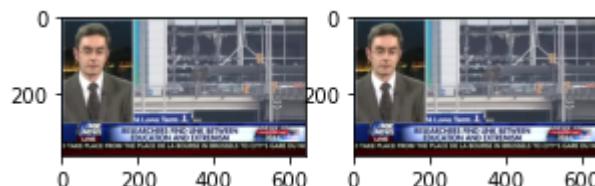
CLUSTER--> 179 TOTAL IMAGES--> 2



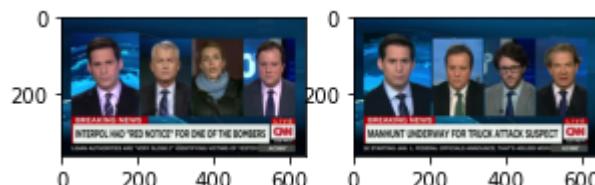
CLUSTER--> 180 TOTAL IMAGES--> 2



CLUSTER--> 181 TOTAL IMAGES--> 2



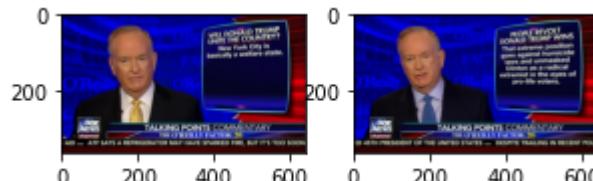
CLUSTER--> 182 TOTAL IMAGES--> 2



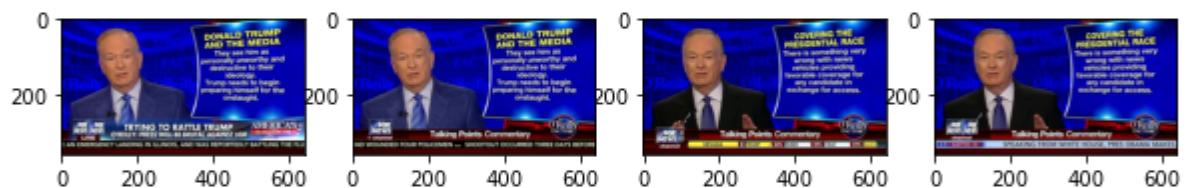
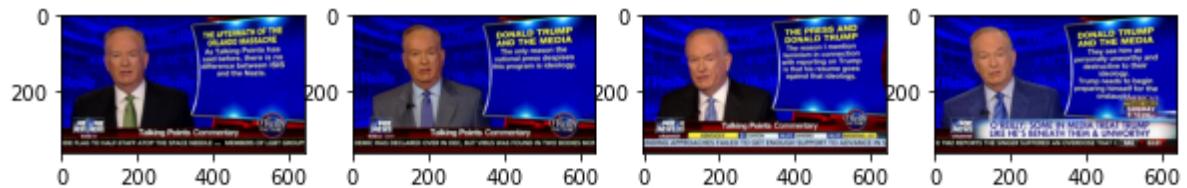
CLUSTER--> 183 TOTAL IMAGES--> 3



CLUSTER--> 184 TOTAL IMAGES--> 2



CLUSTER--> 185 TOTAL IMAGES--> 11



CLUSTER--> 186 TOTAL IMAGES--> 3



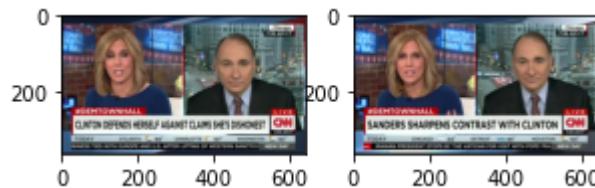
CLUSTER--> 187 TOTAL IMAGES--> 3



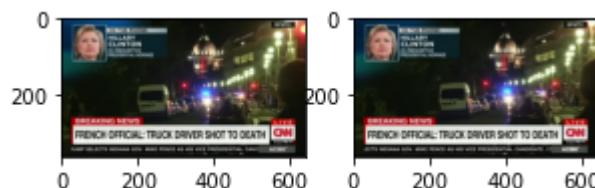
CLUSTER--> 188 TOTAL IMAGES--> 2



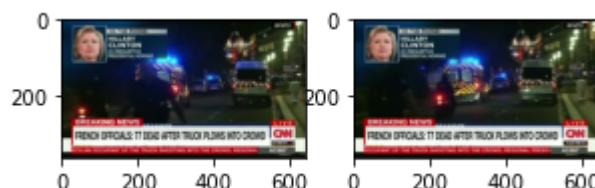
CLUSTER--> 189 TOTAL IMAGES--> 2



CLUSTER--> 190 TOTAL IMAGES--> 2



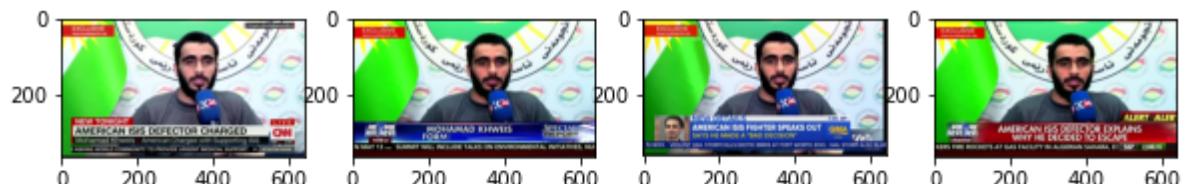
CLUSTER--> 191 TOTAL IMAGES--> 2



CLUSTER--> 192 TOTAL IMAGES--> 3



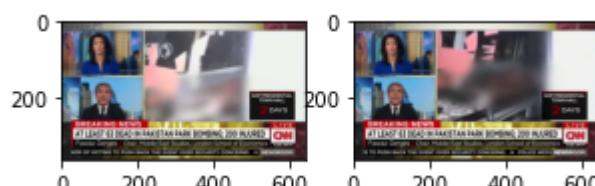
CLUSTER--> 193 TOTAL IMAGES--> 8



CLUSTER--> 194 TOTAL IMAGES--> 2



CLUSTER--> 195 TOTAL IMAGES--> 2



CLUSTER--> 196 TOTAL IMAGES--> 2



CLUSTER--> 197 TOTAL IMAGES--> 7



CLUSTER--> 198 TOTAL IMAGES--> 4



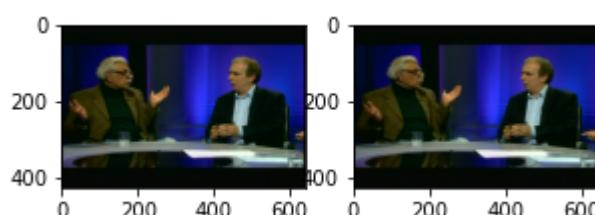
CLUSTER--> 199 TOTAL IMAGES--> 2



CLUSTER--> 200 TOTAL IMAGES--> 2



CLUSTER--> 201 TOTAL IMAGES--> 2



CLUSTER--> 202 TOTAL IMAGES--> 3



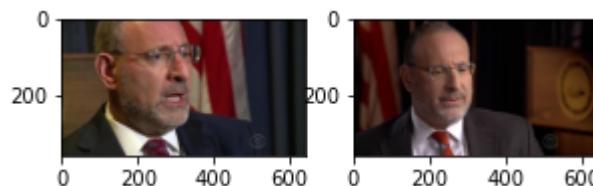
CLUSTER--> 203 TOTAL IMAGES--> 2



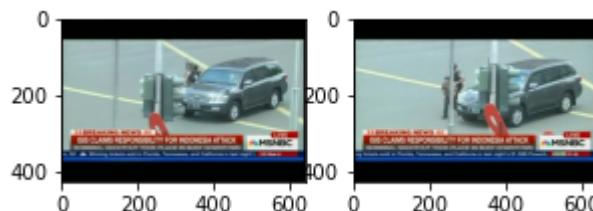
CLUSTER--> 204 TOTAL IMAGES--> 3



CLUSTER--> 205 TOTAL IMAGES--> 2



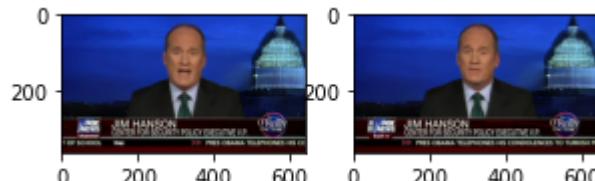
CLUSTER--> 206 TOTAL IMAGES--> 2



CLUSTER--> 207 TOTAL IMAGES--> 3



CLUSTER--> 208 TOTAL IMAGES--> 2



CLUSTER--> 209 TOTAL IMAGES--> 2



CLUSTER--> 210 TOTAL IMAGES--> 2



CLUSTER--> 211 TOTAL IMAGES--> 3



CLUSTER--> 212 TOTAL IMAGES--> 4



CLUSTER--> 213 TOTAL IMAGES--> 2



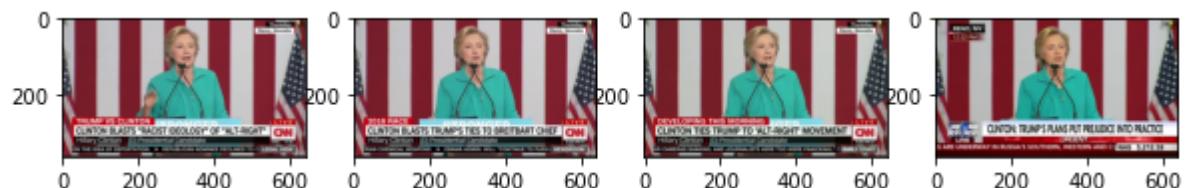
CLUSTER--> 214 TOTAL IMAGES--> 4



CLUSTER--> 215 TOTAL IMAGES--> 2



CLUSTER--> 216 TOTAL IMAGES--> 4



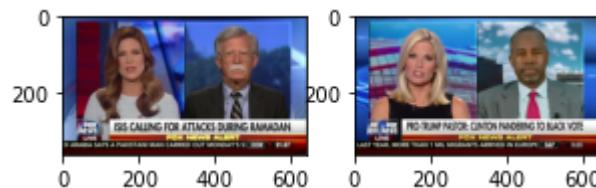
CLUSTER--> 217 TOTAL IMAGES--> 3



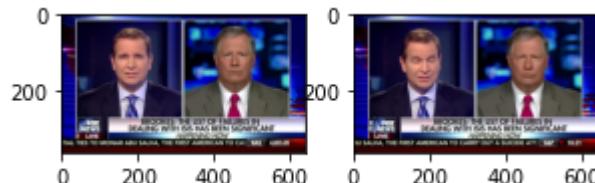
CLUSTER--> 218 TOTAL IMAGES--> 2



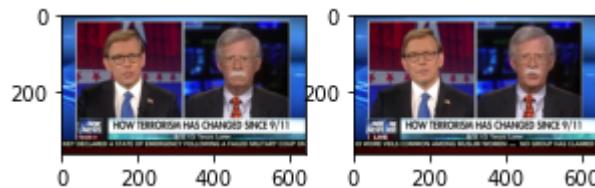
CLUSTER--> 219 TOTAL IMAGES--> 2



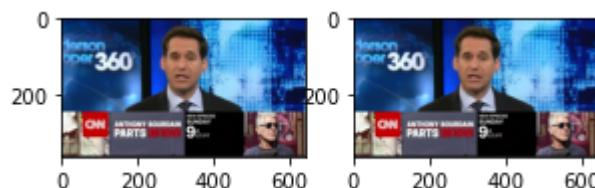
CLUSTER--> 220 TOTAL IMAGES--> 2



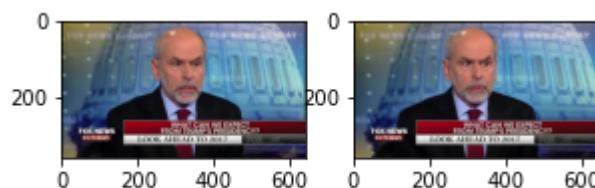
CLUSTER--> 221 TOTAL IMAGES--> 2



CLUSTER--> 222 TOTAL IMAGES--> 2



CLUSTER--> 223 TOTAL IMAGES--> 2



CLUSTER--> 224 TOTAL IMAGES--> 2



CLUSTER--> 225 TOTAL IMAGES--> 2



CLUSTER--> 226 TOTAL IMAGES--> 4



CLUSTER--> 227 TOTAL IMAGES--> 3



CLUSTER--> 228 TOTAL IMAGES--> 2



CLUSTER--> 229 TOTAL IMAGES--> 3



CLUSTER--> 230 TOTAL IMAGES--> 3



CLUSTER--> 231 TOTAL IMAGES--> 3



CLUSTER--> 232 TOTAL IMAGES--> 4



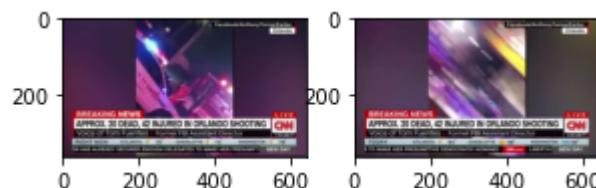
CLUSTER--> 233 TOTAL IMAGES--> 2



CLUSTER--> 234 TOTAL IMAGES--> 2



CLUSTER--> 235 TOTAL IMAGES--> 2



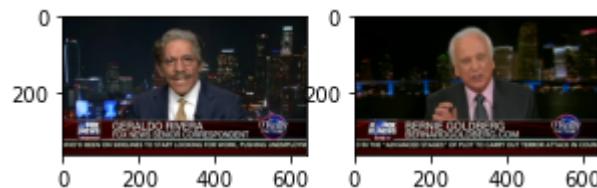
CLUSTER--> 236 TOTAL IMAGES--> 6



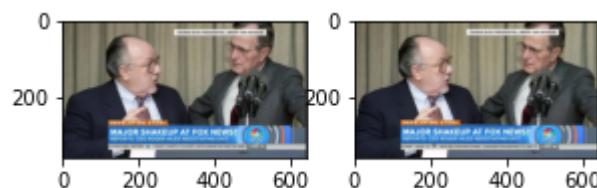
CLUSTER--> 237 TOTAL IMAGES--> 2



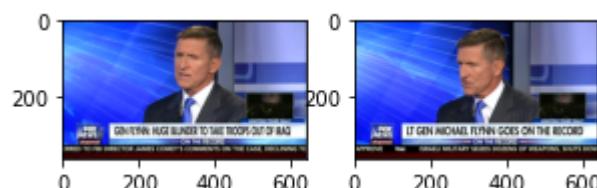
CLUSTER--> 238 TOTAL IMAGES--> 2



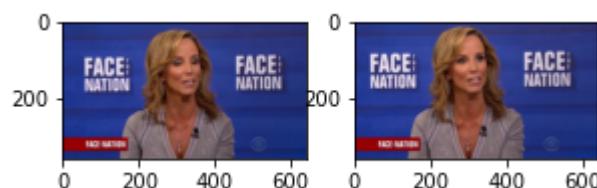
CLUSTER--> 239 TOTAL IMAGES--> 2



CLUSTER--> 240 TOTAL IMAGES--> 2



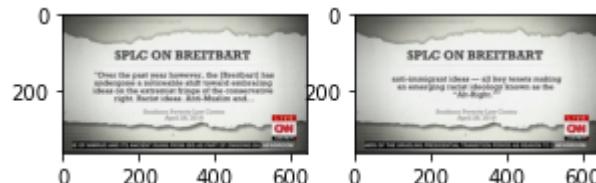
CLUSTER--> 241 TOTAL IMAGES--> 2



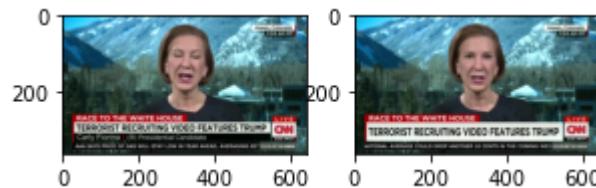
CLUSTER--> 242 TOTAL IMAGES--> 3



CLUSTER--> 243 TOTAL IMAGES--> 2



CLUSTER--> 244 TOTAL IMAGES--> 2



CLUSTER--> 245 TOTAL IMAGES--> 2



CLUSTER--> 246 TOTAL IMAGES--> 5



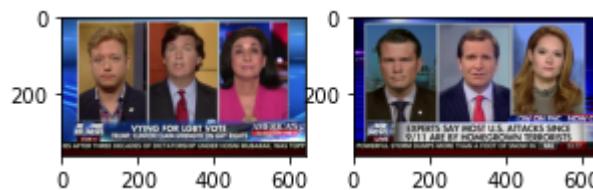
CLUSTER--> 247 TOTAL IMAGES--> 3



CLUSTER--> 248 TOTAL IMAGES--> 3



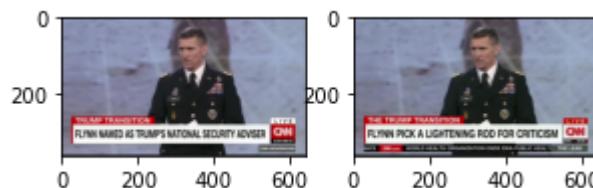
CLUSTER--> 249 TOTAL IMAGES--> 2



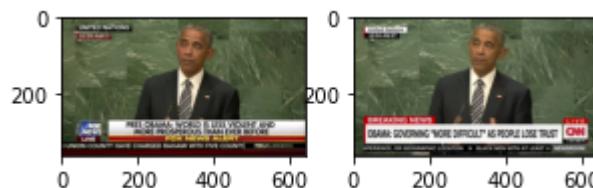
CLUSTER--> 250 TOTAL IMAGES--> 2



CLUSTER--> 251 TOTAL IMAGES--> 2



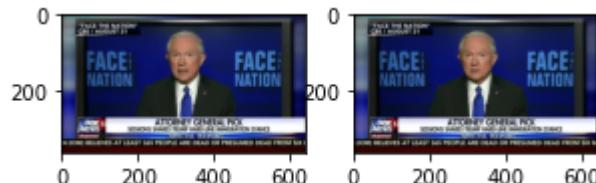
CLUSTER--> 252 TOTAL IMAGES--> 2



CLUSTER--> 253 TOTAL IMAGES--> 2



CLUSTER--> 254 TOTAL IMAGES--> 2



CLUSTER--> 255 TOTAL IMAGES--> 2