Mehmet Berk Türkçapar
21902570

CS-464: Introduction to Machine Learning
Homework 2

Question 1.1

```
Red channel PVE sum for the first 10 principal components:
Principal component 1: 0.2350696993650973
Principal component 2: 0.3915808514223919
Principal component 3: 0.48163338999534244
Principal component 4: 0.5499329368273457
Principal component 5: 0.5874602763381437
Principal component 6: 0.6114078154737781
Principal component 7: 0.6341724742528841
Principal component 8: 0.6553006837195982
Principal component 9: 0.6732366043040807
Principal component 10: 0.6867302133532457
```

```
Green channel PVE for the first 10 principal components:
Principal component 1: 0.228590359053683 52
Principal component 2: 0.385082938326110 74
Principal component 3: 0.4729888940831668
Principal component 4: 0.5350243758269491
Principal component 5: 0.5724257178602572
Principal component 6: 0.5965915917203778
Principal component 7: 0.6206389256874243
Principal component 8: 0.6412350602788931
Principal component 9: 0.6596940546392752
Principal component 10: 0.6739797747077685
```

```
Blue channel PVE for the first 10 principal components:
Principal component 1: 0.2087371485475201
Principal component 2: 0.36758280817282707
Principal component 3: 0.46017137679920705
Principal component 4: 0.5282824942617058
Principal component 5: 0.5662675470195494
Principal component 6: 0.5907348644687901
Principal component 7: 0.615014027883697
Principal component 8: 0.6365045561464204
Principal component 9: 0.6553745590486495
Principal component 10: 0.6695858942544113
```

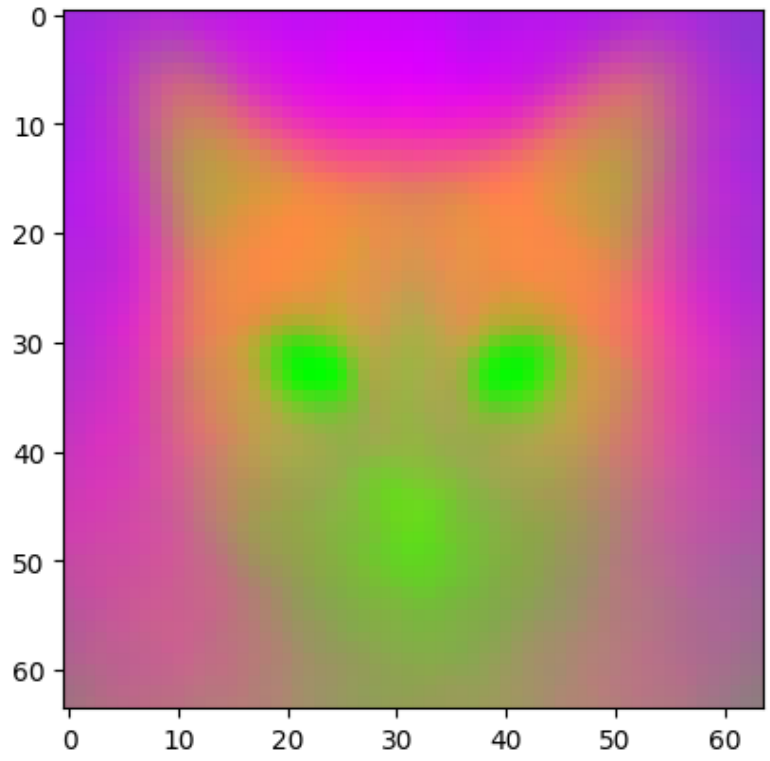
As seen from the data below, the minimum number of principal components required for at least 70% PVE for all channels is 13. As the number of principal components increases, the PVE also increases which means we get a better representation of the original data.

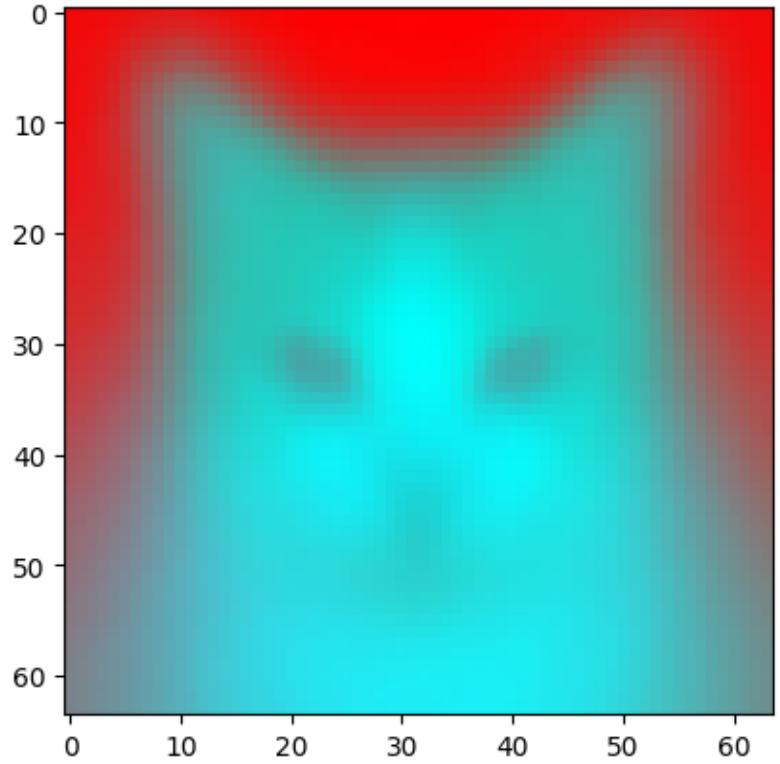
```
Red channel PVE for the principal components:
Principal component 11: 0.6981816476502414
Principal component 12: 0.7081842760520995
Principal component 13: 0.71788997208389
Principal component 14: 0.7266976985389427
Green channel PVE for the principal components:
Principal component 11: 0.6856969305285905
Principal component 12: 0.6959232719754039
Principal component 13: 0.7057291119296866
Principal component 14: 0.7150642607684055
Blue channel PVE for the principal components:
Principal component 11: 0.6811286368261021
Principal component 12: 0.6913055152453241
Principal component 13: 0.7012453873716817
Principal component 14: 0.7105624353635005
```

Mehmet Berk Türkçapar
21902570

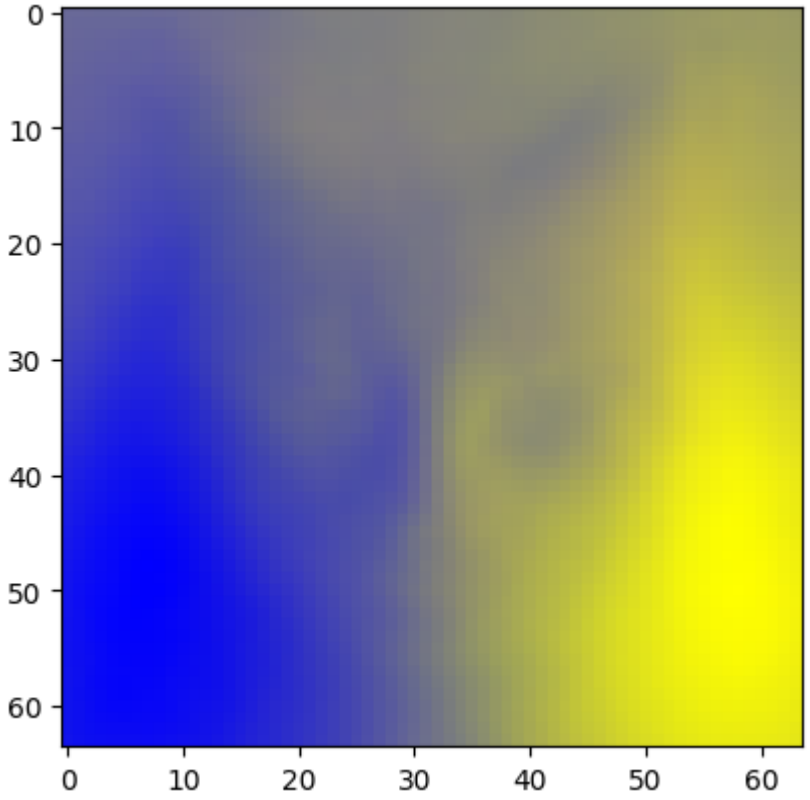
Question 1.2

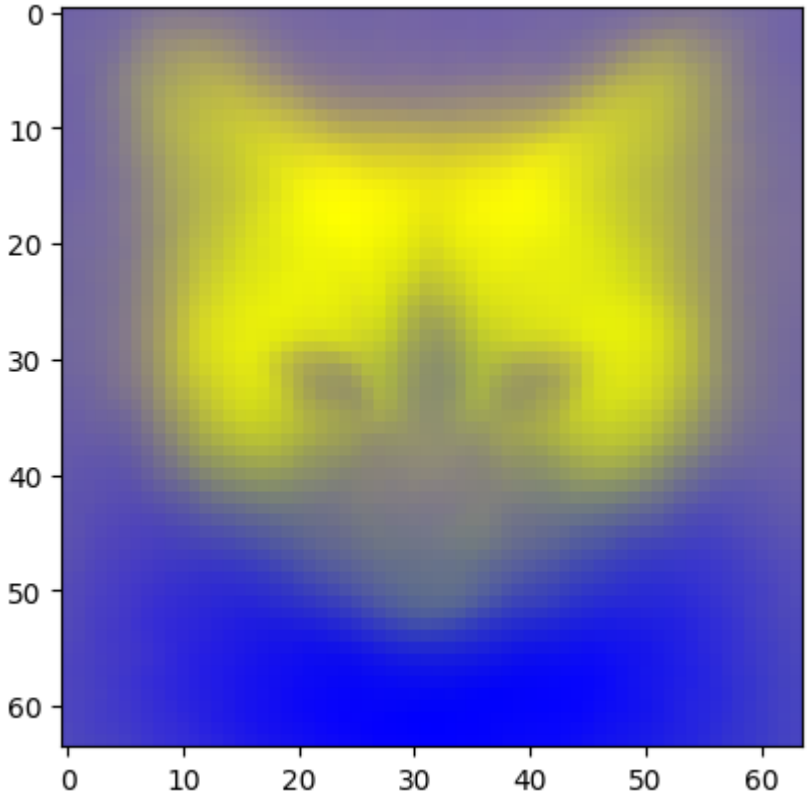
Below shows the 10 RGB images corresponding to the first 10 principal components found. As it can be seen, some of them focus on the overall edges while some focus on textures and variations. These images are important to understand the general features and patterns of the dataset.



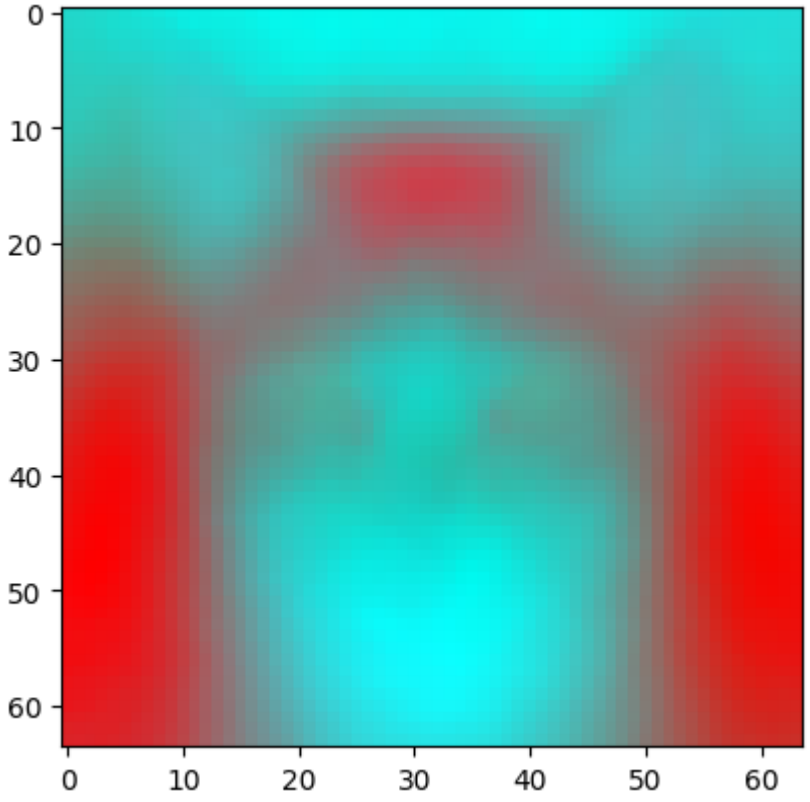


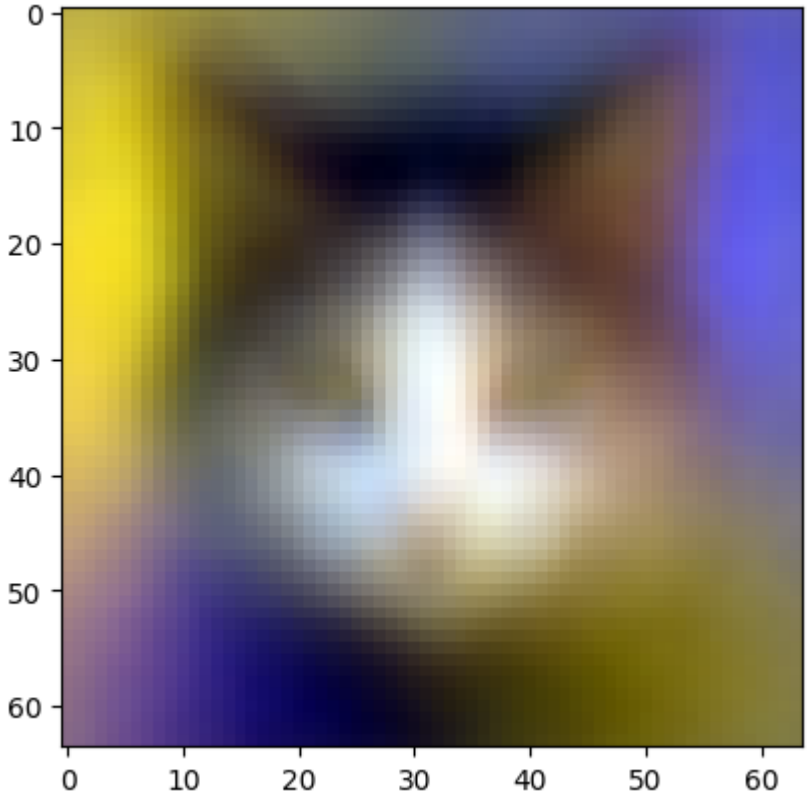
Mehmet Berk Türkçapar
21902570



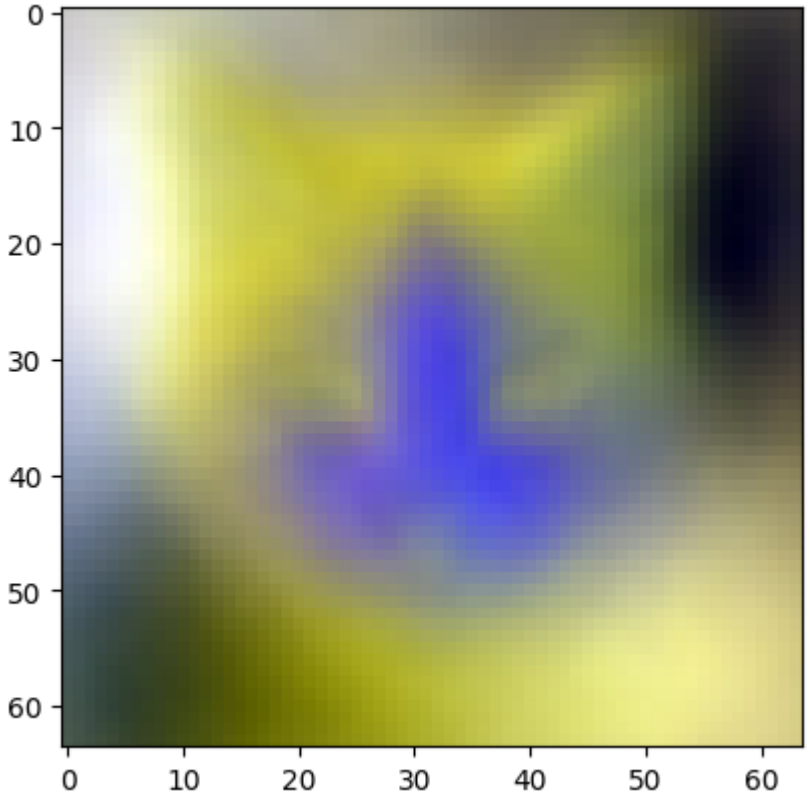


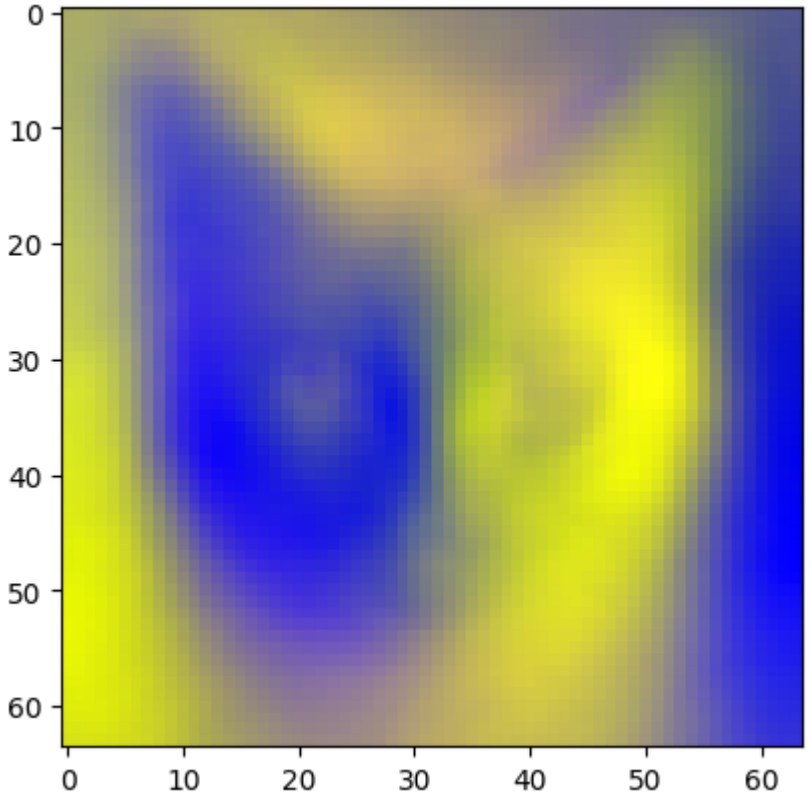
Mehmet Berk Türkçapar
21902570



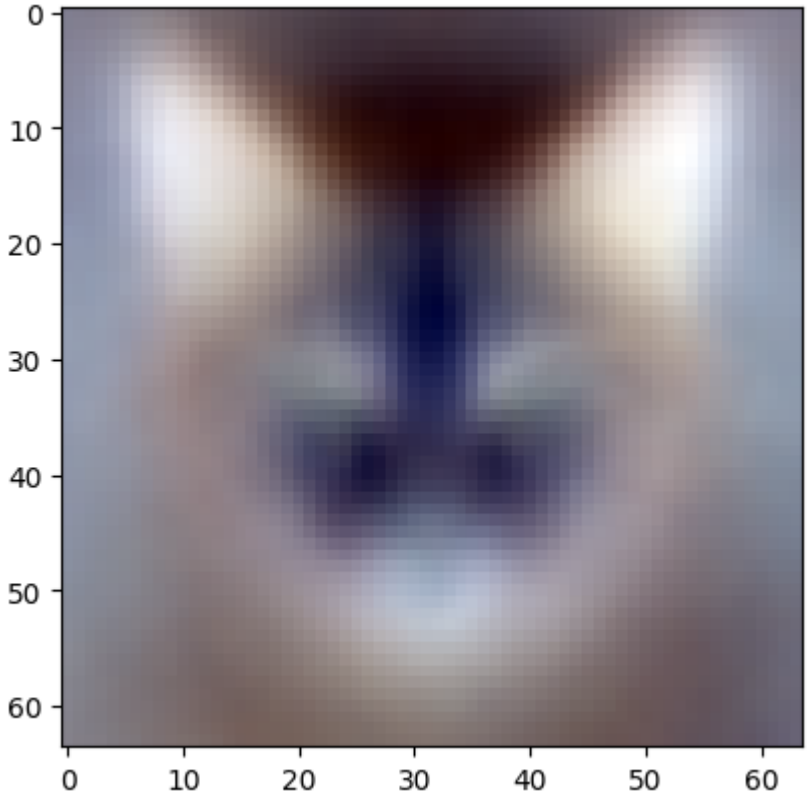


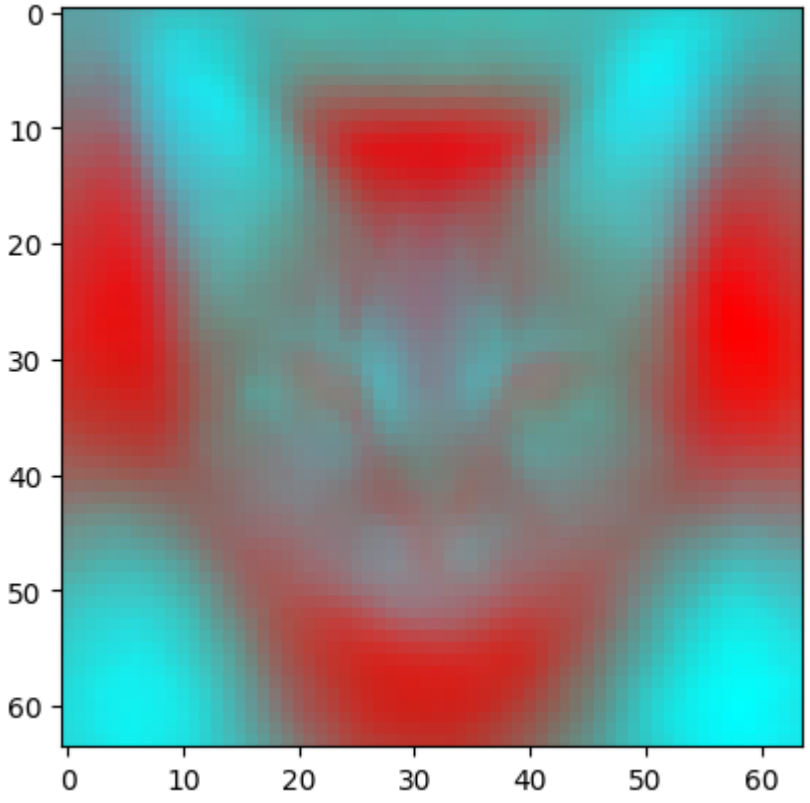
Mehmet Berk Türkçapar
21902570





Mehmet Berk Türkçapar
21902570



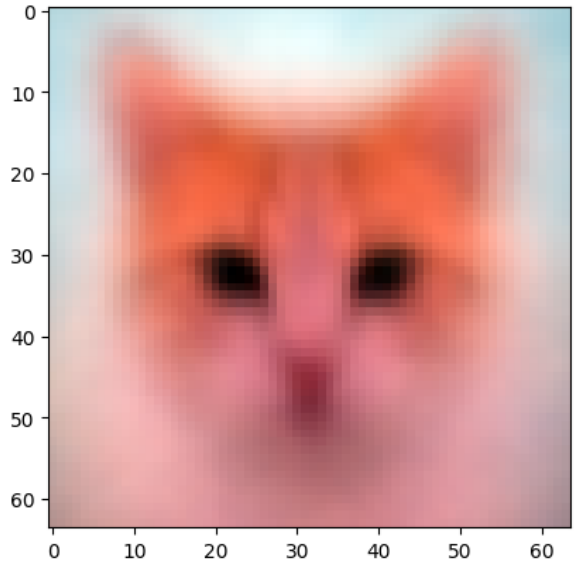


Mehmet Berk Türkçapar
21902570

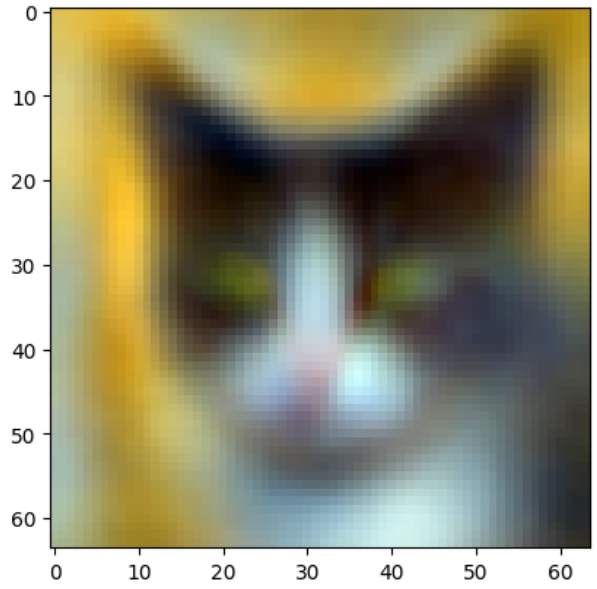
Question 1.3

In order to reconstruct the second image (flickr_cat_000003.jpg), I took the dot product of the image's color channels with the principal components and projected the data back to the original space using k principal components. As it can be seen from the images below, as the principal components, k, increased the image became much more similar to the original image. With more principal components used, we are able to capture information about the data, and with more information, we are able to reconstruct a more similar image to the original one.

For k = 1:

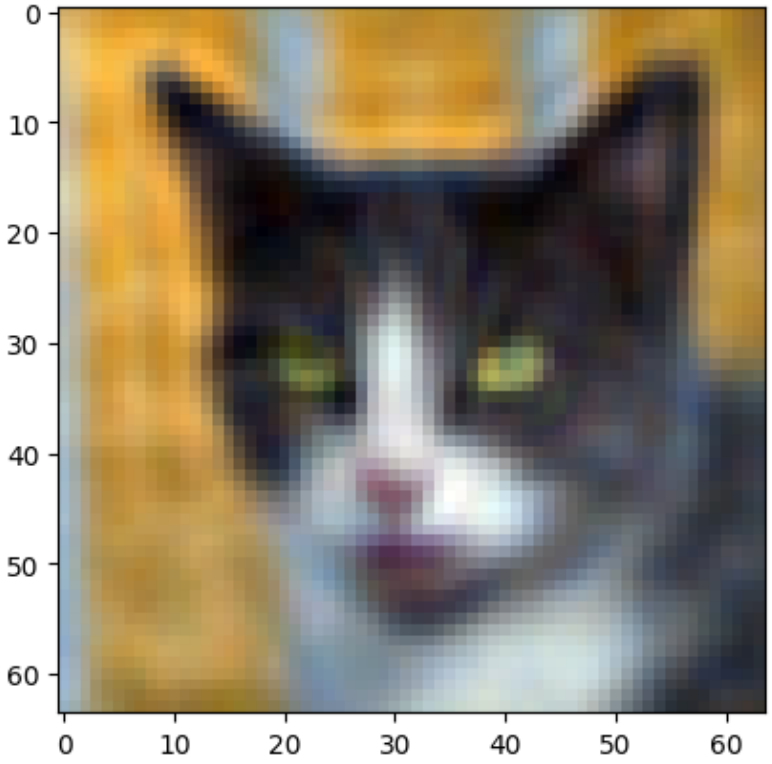


For k = 50:

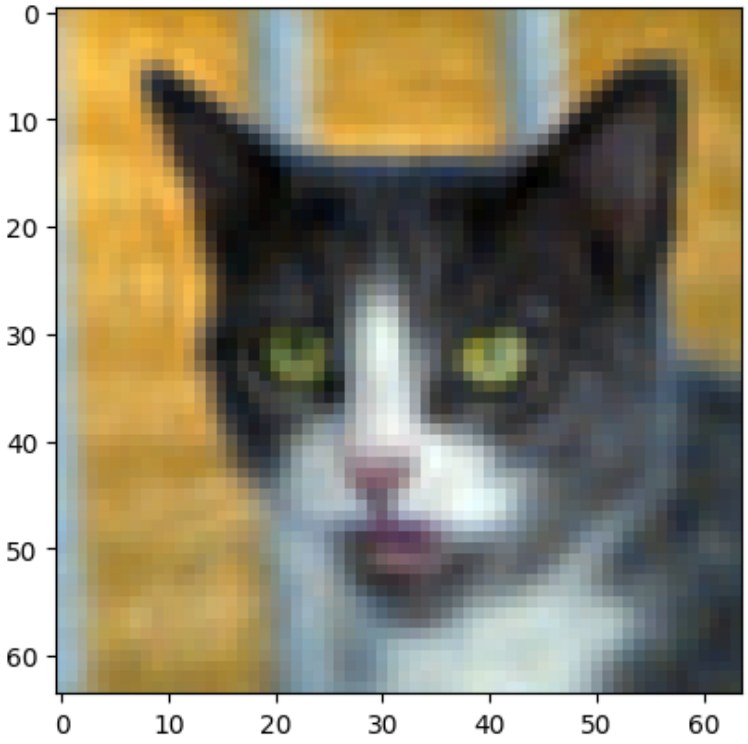


Mehmet Berk Türkçapar
21902570

For k = 250:

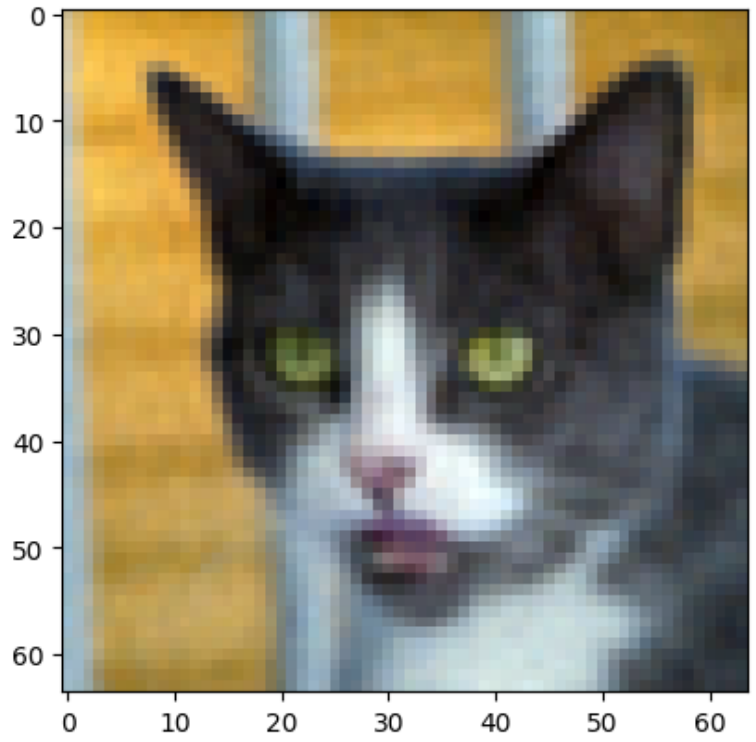


For k = 500:



Mehmet Berk Türkçapar
21902570

For k = 1000:



For k = 4096: