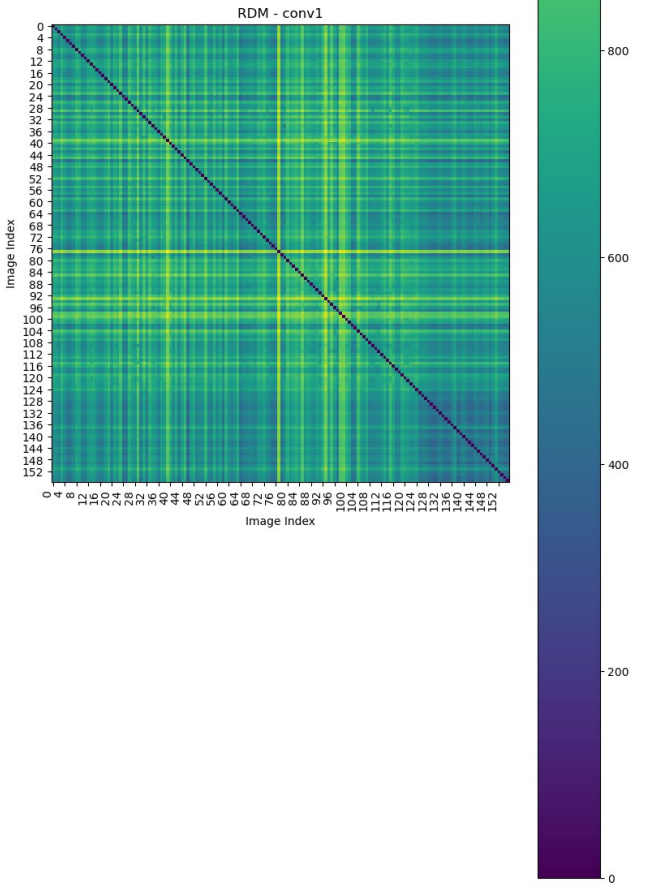
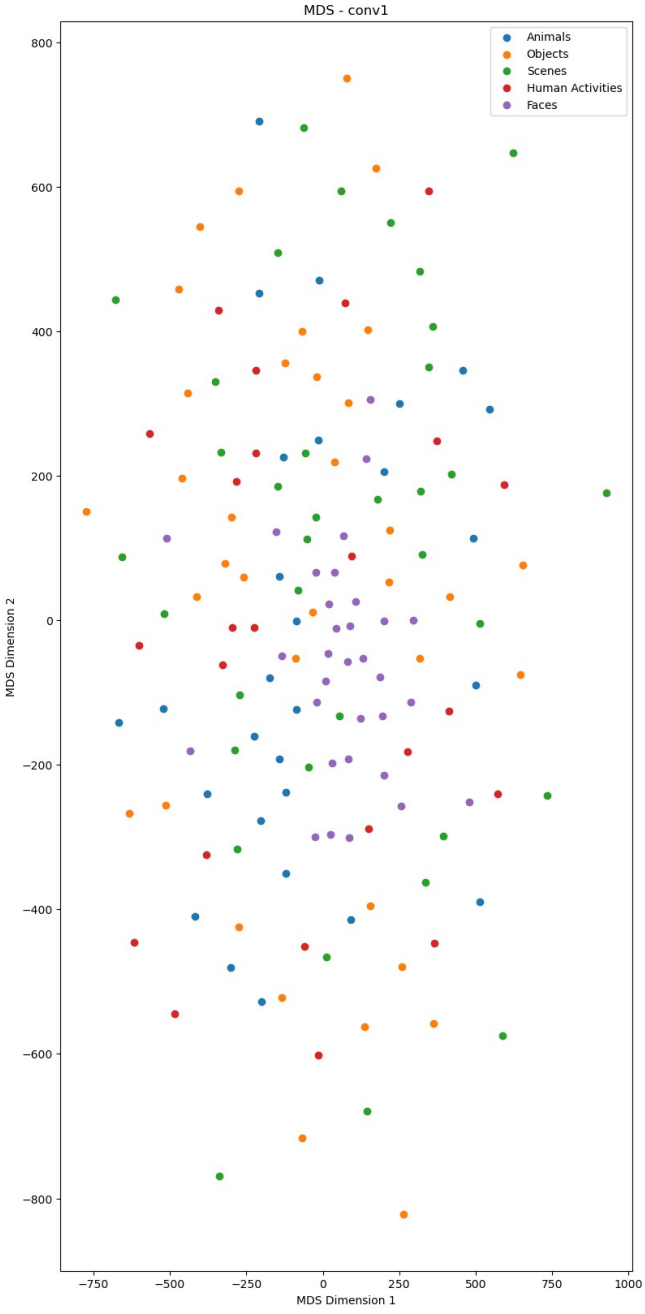
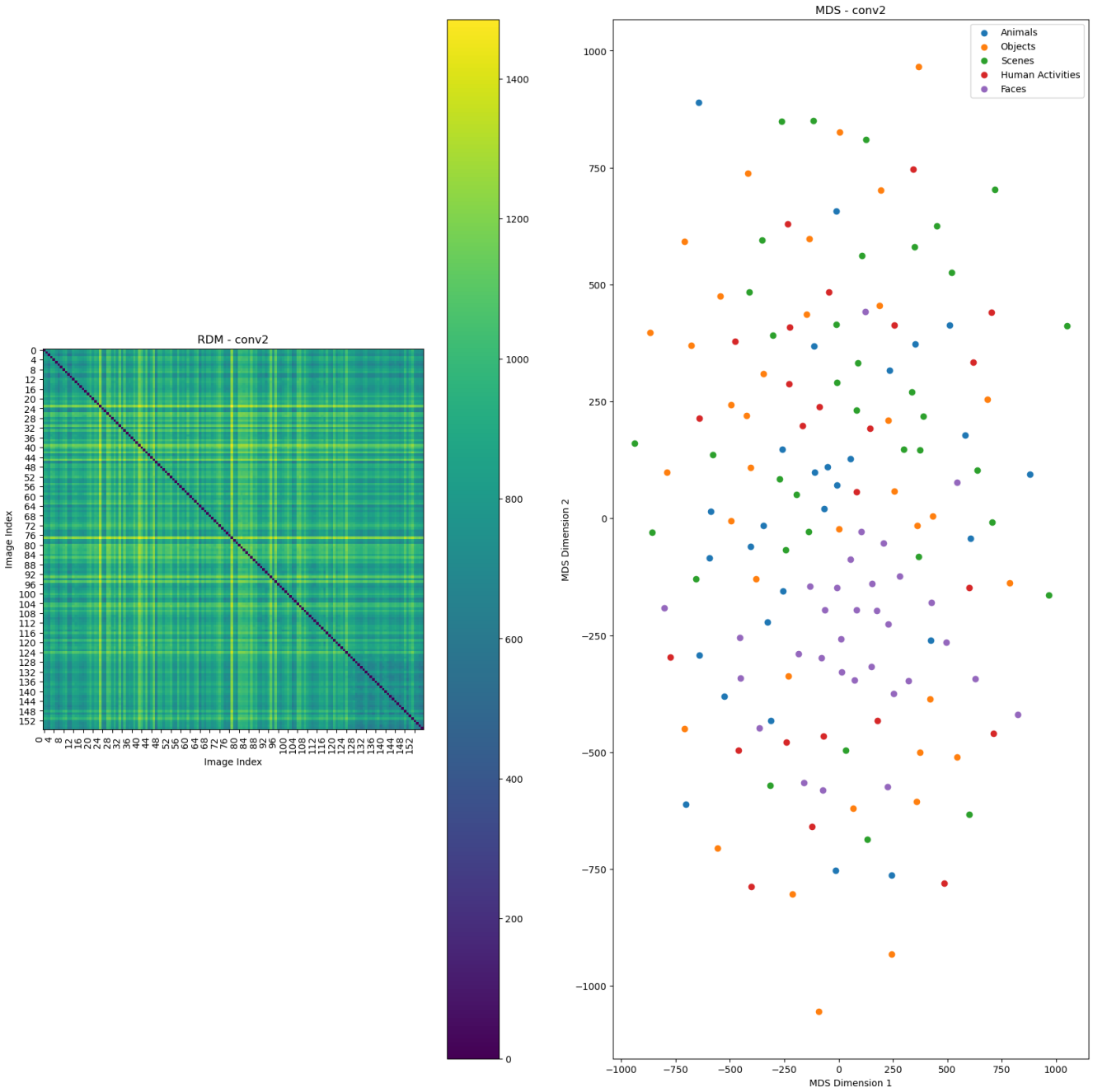
Conv 1

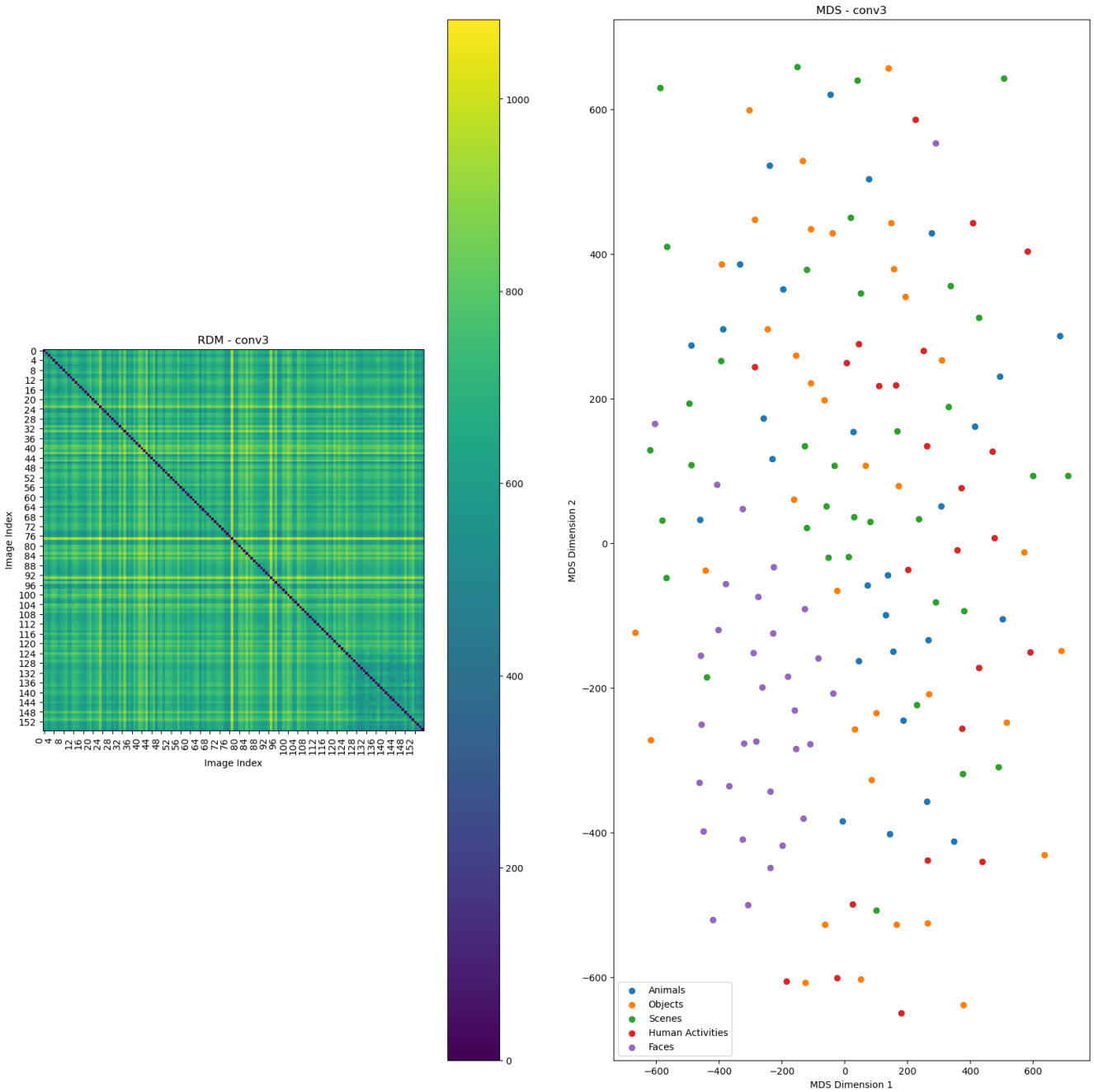

RDM - conv1
Image Index
Image Index


MDS - conv1
Animals
Objects
Scenes
Human Activities
Faces
MDS Dimension 1
MDS Dimension 2

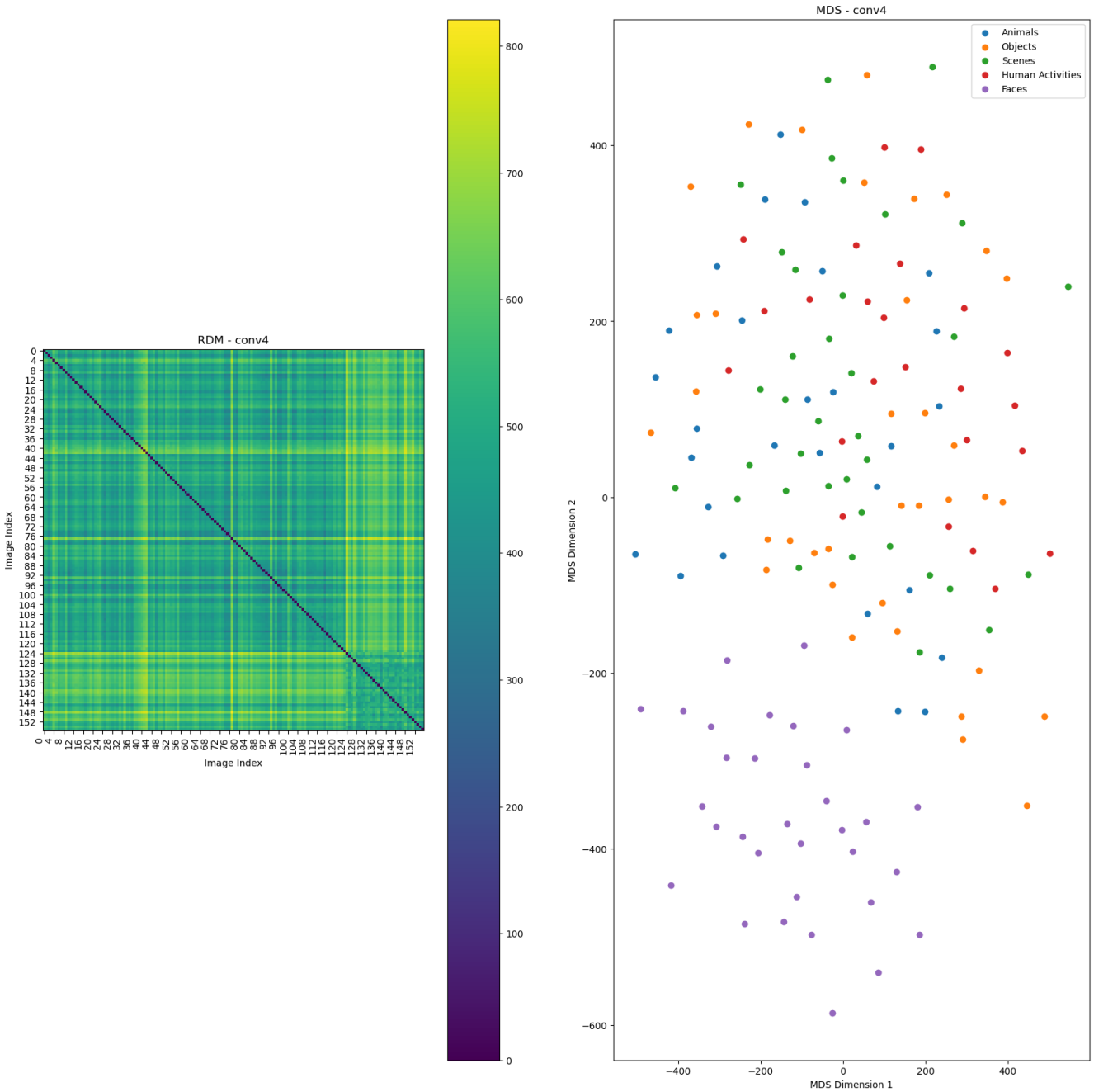
Conv 2


RDM - conv2
Image Index
Image Index
MDS - conv2
Animals
Objects
Scenes
Human Activities
Faces
MDS Dimension 1
MDS Dimension 2

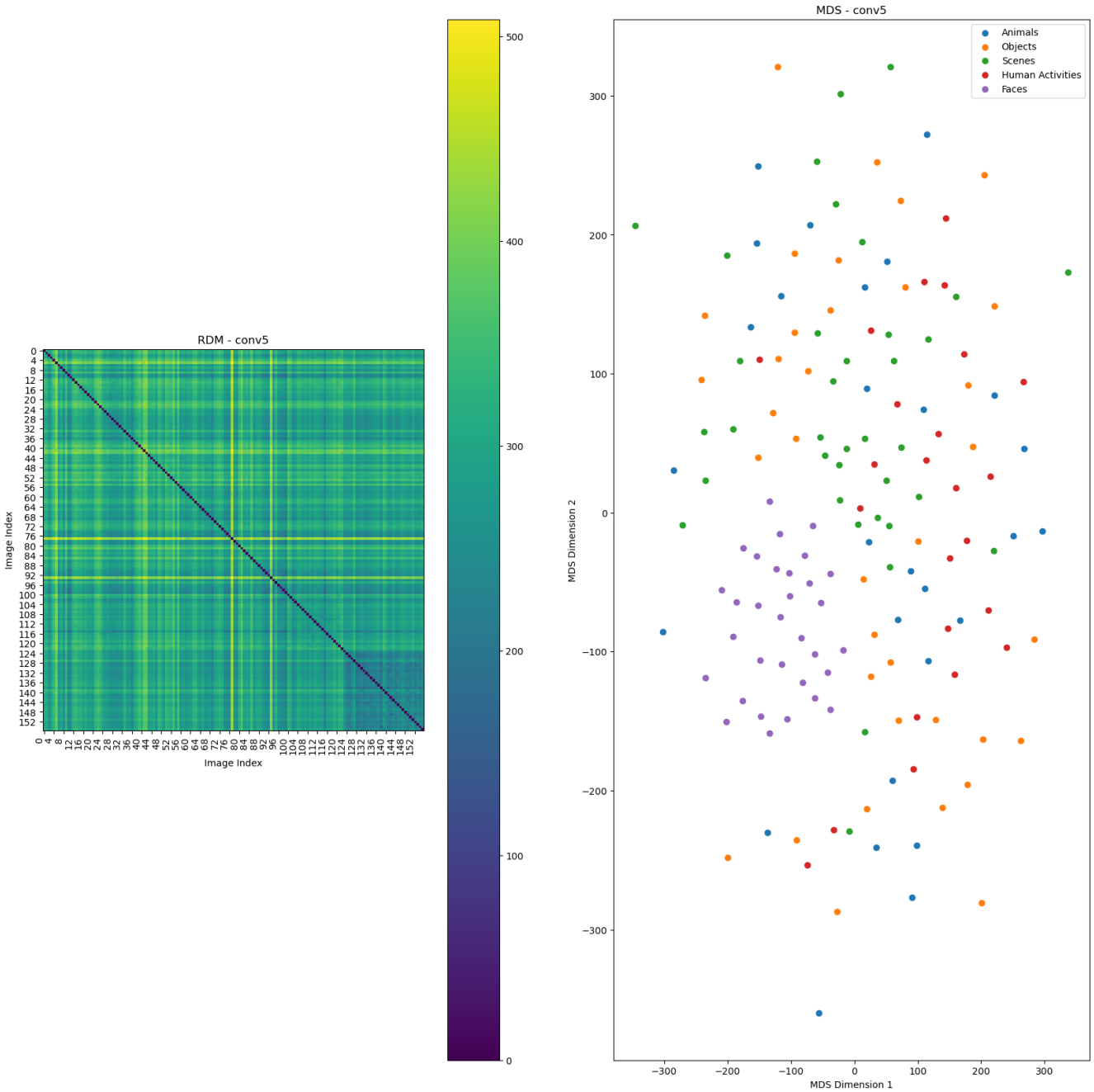
Conv 3


RDM - conv3
Image Index
Image Index
MDS - conv3
MDS Dimension 2
MDS Dimension 1
Animals
Objects
Scenes
Human Activities
Faces

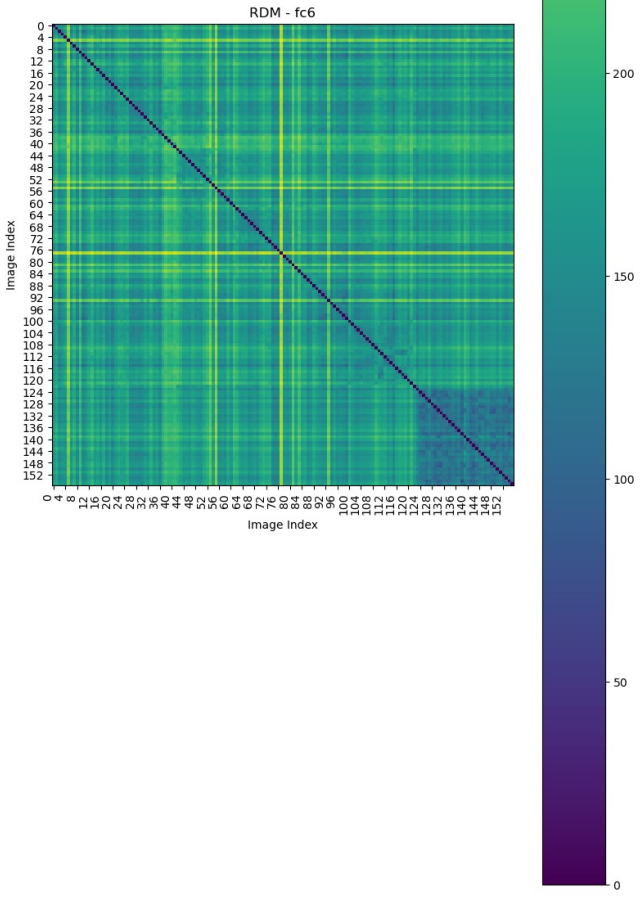
Conv 4

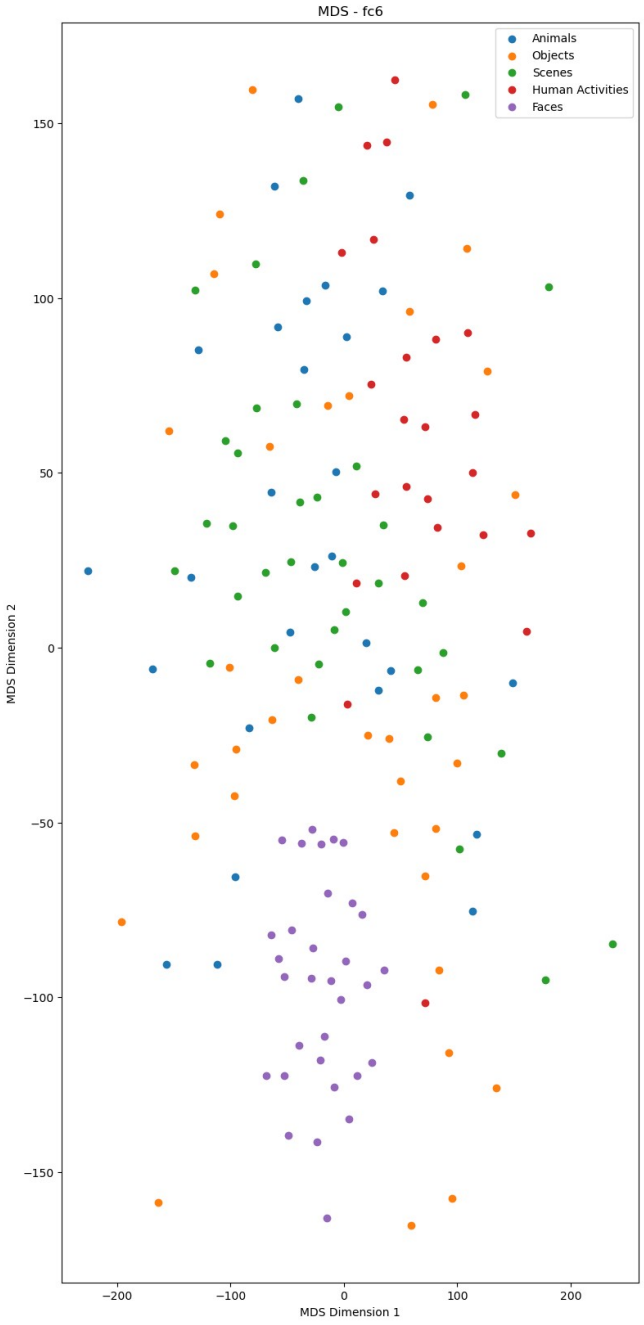

RDM - conv4
Image Index
Image Index
MDS - conv4
Animals
Objects
Scenes
Human Activities
Faces
MDS Dimension 1
MDS Dimension 2

Conv 5

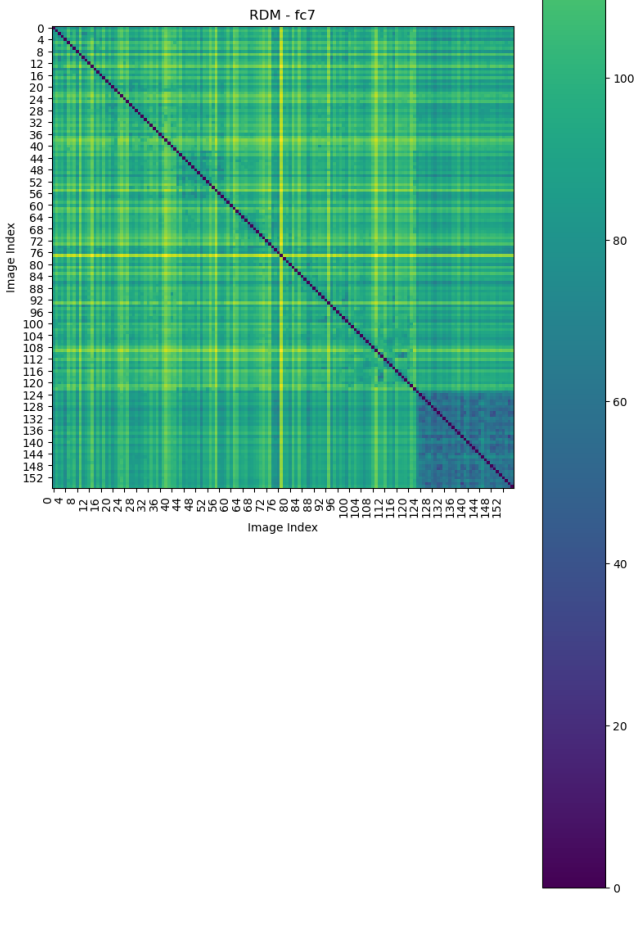

RDM - conv5
Image Index
Image Index
MDS - conv5
Animals
Objects
Scenes
Human Activities
Faces
MDS Dimension 1
MDS Dimension 2

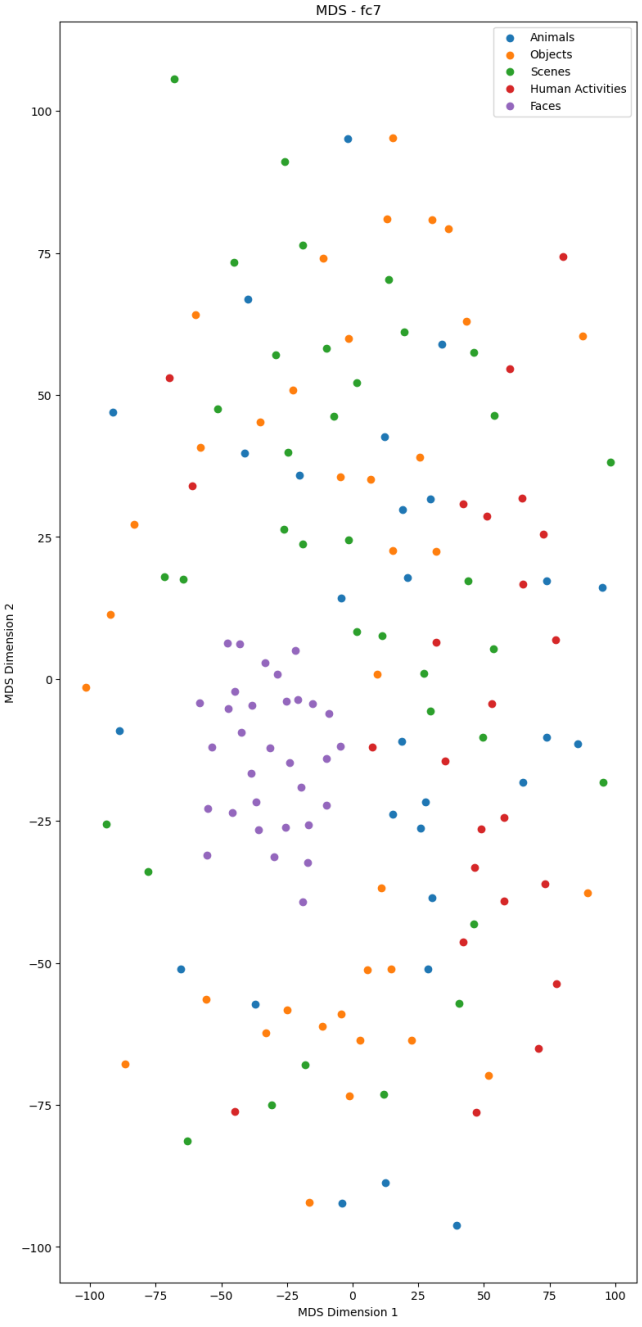
fc 6


RDM - fc6
Image Index
Image Index
300
250
200
150
100
50
0


MDS - fc6
Animals
Objects
Scenes
Human Activities
Faces
MDS Dimension 2
MDS Dimension 1
150
100
50
0
−50
−100
−150
−200
−100
0
100
200

fc7


RDM - fc7
Image Index
Image Index


MDS - fc7
Animals
Objects
Scenes
Human Activities
Faces
MDS Dimension 1
MDS Dimension 2

Report

Short Report

Introduction:

The goal of the assignment was to explore the feature extractions learned by the AlexNet pre-trained model across Conv 1 to 5 and Fc 6,7. There are 5 categories in the image dataset. After getting the activations of each layer, we create a representational dissimilarity matrix (RDM) via Euclidean distance between vectors of activation of corresponding images for each layer. Then we visualize the RDM using multidimensional scaling (MDS) in 2D dimensions.

Method:

In this assignment, I utilized PyTorch for the deep learning framework to load the pre-train AlexNet model. I have 156 images provided for the input of AlexNet to extract features maps from the layers discussed above. All feature maps/activations were vectorized and were used for RDM using Euclidean distance between activation vectors. After, I use MDS to visualize the dissimilarities between activations in a 2D space.

Result:

Convolutional Layers (Conv 1 to 5):

In Conv 1,2,3, the RDMs showed that the layers captured low-level features. That is the reason why they have higher dissimilarities between images.
In Conv 4,5, the dissimilarities decreased, which indicated that the layers capture higher-level features.
In MDS, similar images were starting to grouped closely together especially the face class.

Fully Connected (FC) Layers (fc 6, 7):

The RDMs showed even lower dissimilarity between images than convolutional layers. This states that the FC layers successfully encoded more semantic information.
MDS showed that there is separation between different categories, and similar images are even closer to each other. This revealed the model learns the representation for different similar images and image classes. The model will identify face images.

Conclusion:

The results demonstrate that the feature maps learned by AlexNet advanced from low-level to high-level. The RDMs showed the convolutional layers learned the basic features of images. And the MDS showed that fully connected layers learned the encoded representation of different image categories. The final ouput indicate that the model have learned the feature of each class image because MDSs showes that they are from chaotic, disorder to become order based on image similarity, and different image classes. The model have high accuracy when identify face images.