

Exercise - 3

1 Practical

As you have seen in the lecture, there are several ways to visualize the features extracted by a neural network. Here, we will focus on two dimensionality reduction techniques for visualizing high-dimensional vectors: PCA and t-SNE.

Note - You can use the CNN (not pretrained) from the previous exercise, which you had used for MNIST task for extracting the features. Else you are also allowed to write a basic CNN with 2-3 convolutional layers for extracting the features. The flattened output of the convolutional layers (not the classifier or fully connected layers) are the high dimensional features, which you will use for feature visualization employing PCA and t-SNE.

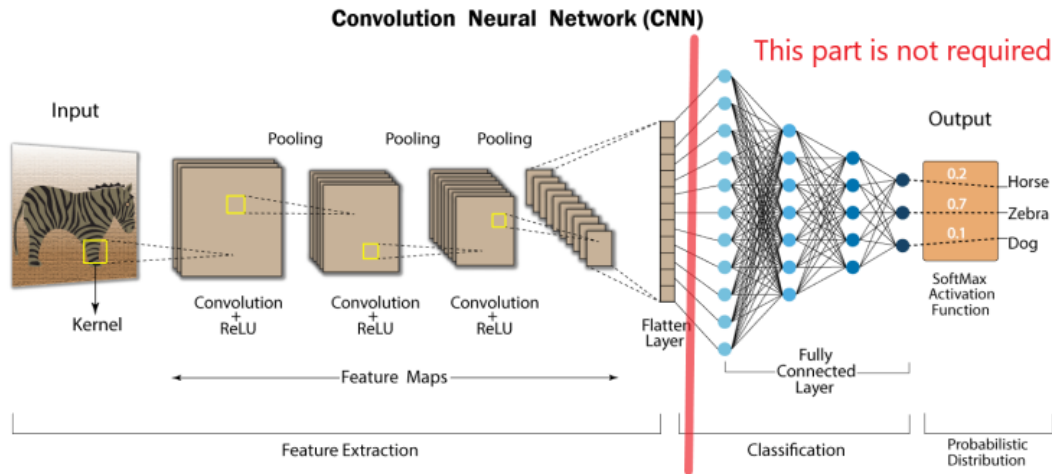


Figure 1: This figure explain how features can be extracted using a convolutional neural network

1.1 Task

1. Train CNN on MNIST with learning rate $1e-6$ (very small) and for one epoch only.
2. Extract the features from that model
3. Visualize the features both with PCA and t-SNE
4. Train CNN on MNIST with learning rate $1e-3$ and for 10-20 epochs
5. Repeat steps 2 and 3 for this model too
6. Report the 4 visualization images and BRIEFLY comment on the differences (a) between PCA and t-SNE and (b) the first model (virtually not trained at all) and the second one.

1.2 Resources

- [t-SNE](#)
- [PCA](#)
- [PCA-sklearn](#)