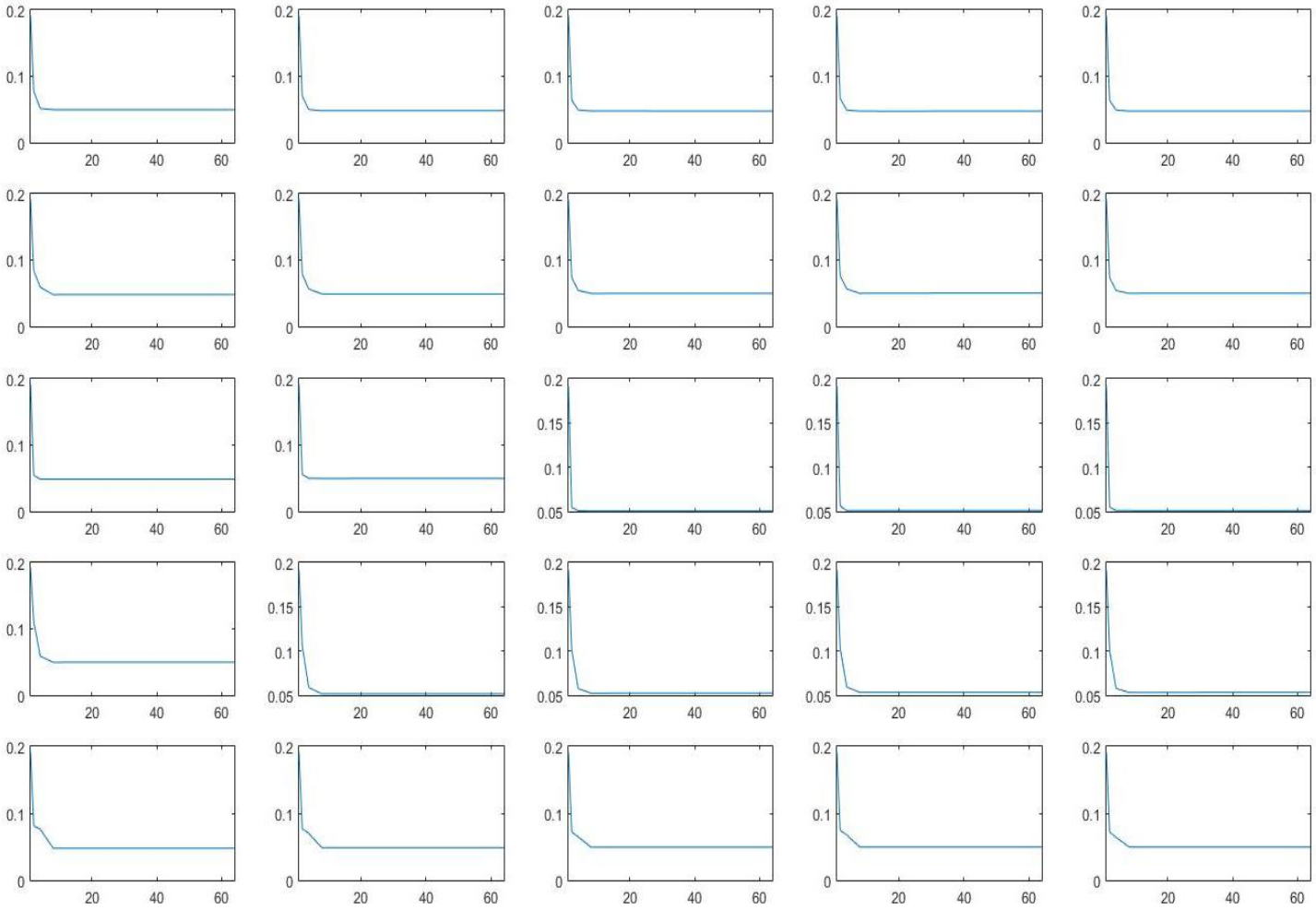
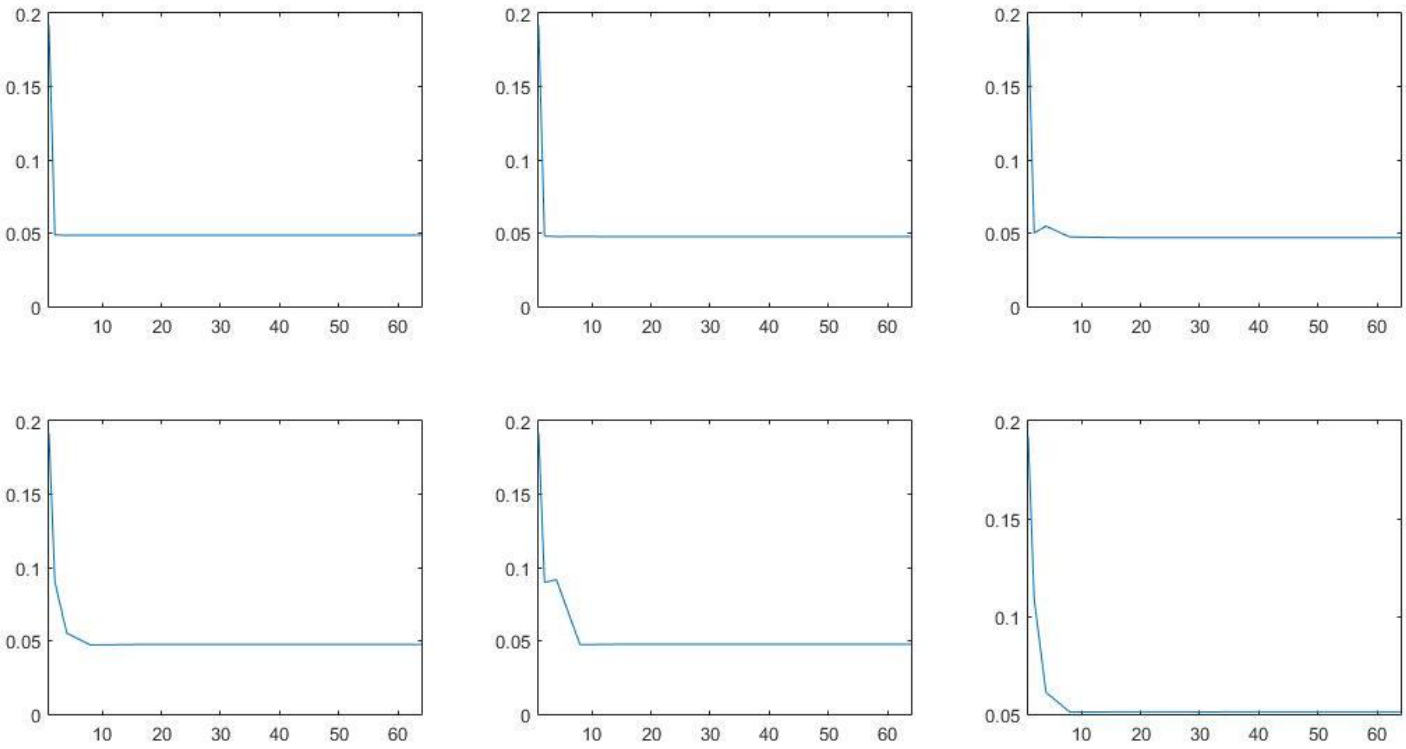
1.



From the 25 classifiers, error probability decrease as dimension increases and reach equilibrium when dimension reaches 16. Different initialization did give slight difference on error probability. Since we are using 64 dimension as initialization and down-sampling into lower dimension, it does not make much difference.

I tried 25 different initializations on 4 dimension itself and error probability could be as high as 18%.

2.



From the graphs, it shows that with more mixture components, error probability reach equilibrium in higher dimension. When C equals 1 ,2 and 4, almost no error probability change after 2D. However, when C becomes larger, the graph stables until it reaches 8D. Although this could be the result of our simple task images. I am not confident of the same result when learning becomes complicated.