Net1	Net2
Train error: 0.3347	Train error: 0.2559
Validation error: 0.3910	Validation error: 0.2990
Test set error: 0.3892	Test set error: 0.2918
epochs: 120	epochs: 120

Table 1: Network 2 consisted of 3 convolutional layers with 20,30, and 50 feature maps in each layer. Network 1 only consisted of 1 convolutional layer with 20 feature maps. A maxpooling layer is applied after every convolutional layer

Theoretical: The kernels in the convolutional layers help detecting patterns (i.e edges, hands etc..) in images. More convolutional layers means patterns that are more complex are more easily filtered out. The max pooling layer after each convolutions layer downsizes the features maps (outputted by the convolution layers). This downsizing will essentialy helps maintain important structure, but eliminates things as rotation/shifting within images. Network 2 have more convolutional layer and maxpool layer which should help better differentiate between similar yet complex objects such as cats and dogs present in the CIFAR-10 data set.

Experimental results: The results are shown in Table 1. It can be seen test error for network 1 and network 2 that network 2 is better able to *generalize* than network 1. This seem to agree with theory more convolutional layers are able to differentiate more complex patterns (i.e cats vs dogs).