

Visualizing last pooling layer activations for each channel, averaged across all images from the poisoned with eyebrows test data set before fine-pruning

