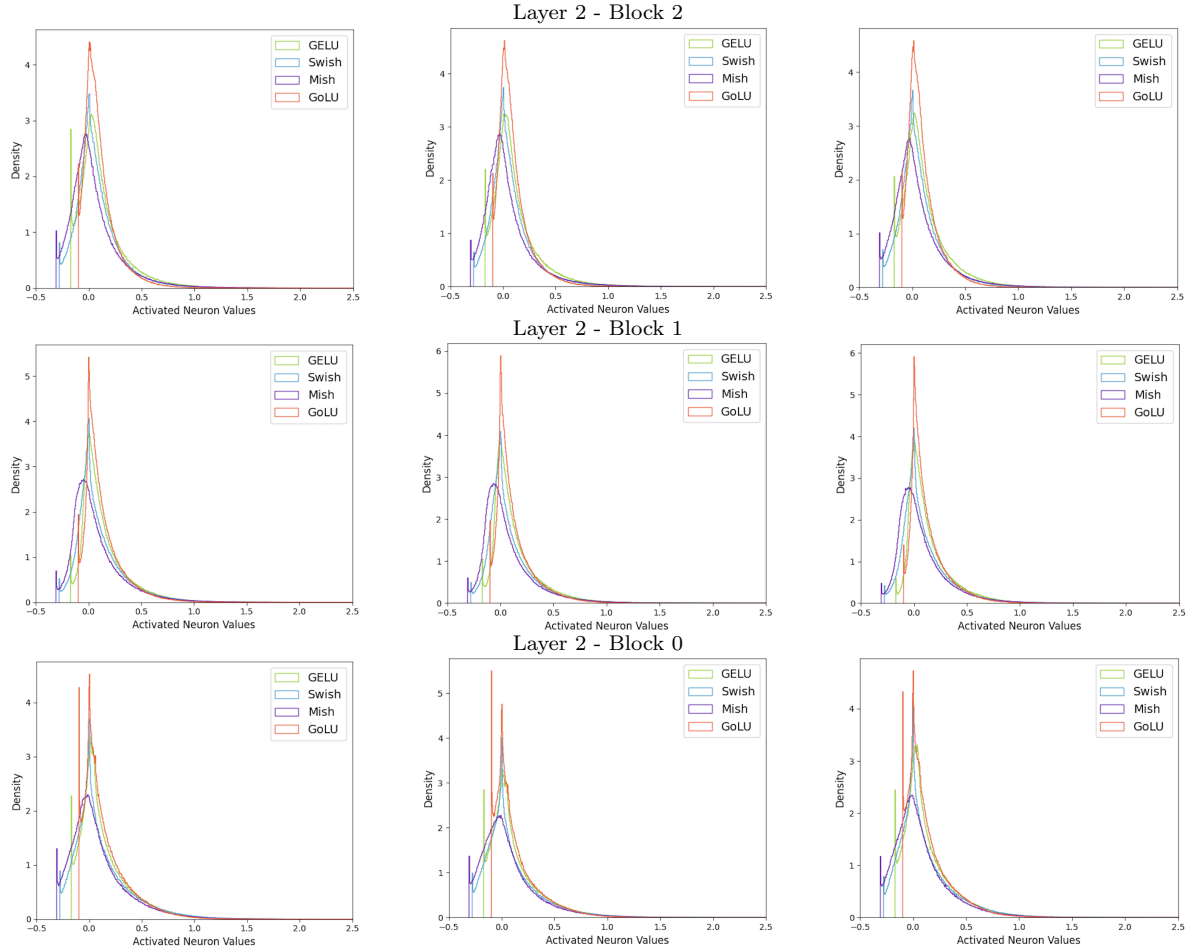
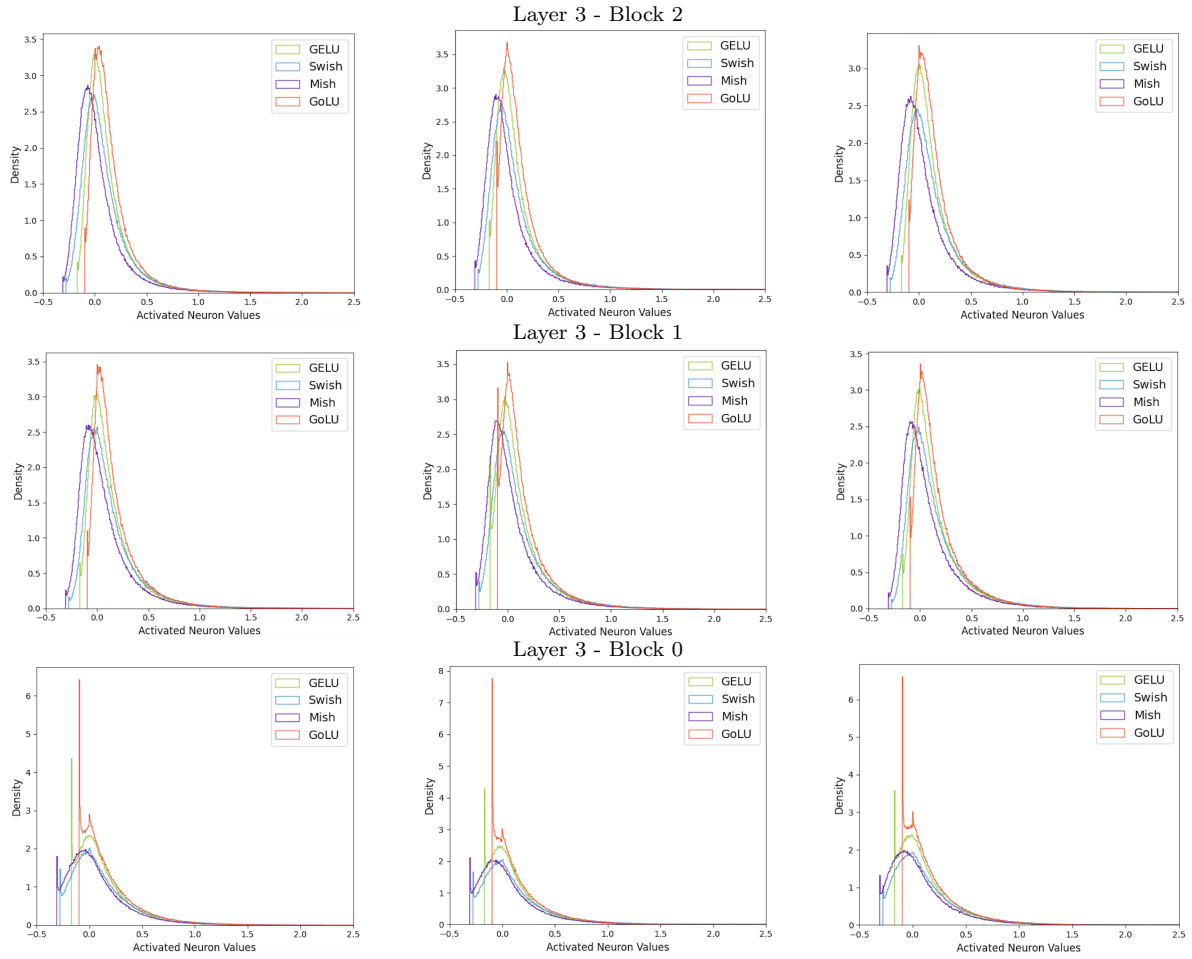


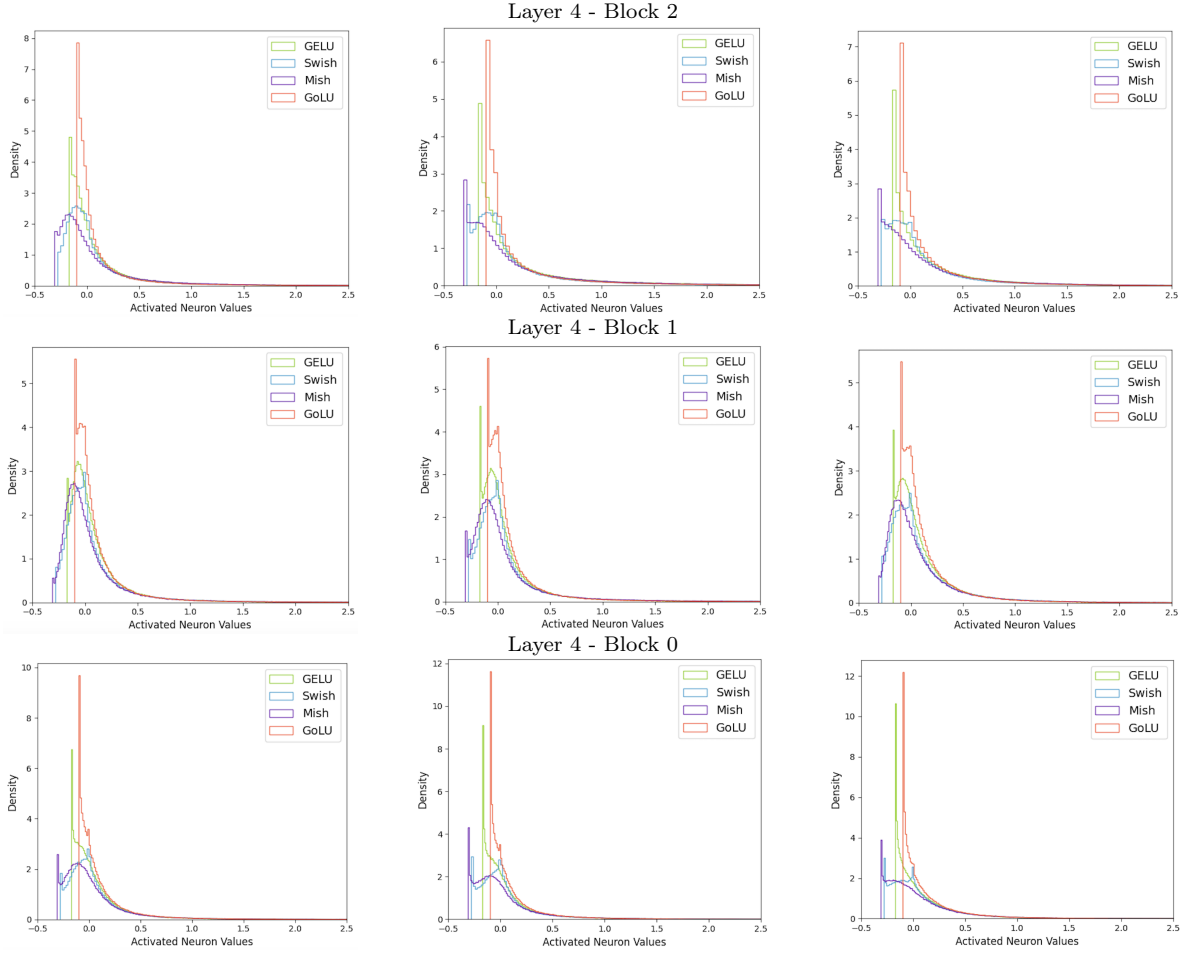
Distributions of (final) activation outputs across different residual blocks of the first layer of ResNet-50 trained on ImageNet-1K, evaluated on three randomly sampled images. GoLU leads to a more peaked distribution.



Distributions of (final) activation outputs across different residual blocks of the second layer of ResNet-50 trained on ImageNet-1K, evaluated on three randomly sampled images. GoLU leads to a more peaked distribution.



Distributions of (final) activation outputs across different residual blocks of the third layer of ResNet-50 trained on ImageNet-1K, evaluated on three randomly sampled images. GoLU leads to a more peaked distribution.



Distributions of (final) activation outputs across different residual blocks of the fourth layer of ResNet-50 trained on ImageNet-1K, evaluated on three randomly sampled images. GoLU leads to a more peaked distribution. The upper row corresponds to the final activation output of the network and reproduces the plots in the manuscript.