Gradients distribution for activation function ReLU()

layers.0.weight layers.3.weight

