input $\mathbf{z}_{l-1}$

neurons $\mathbf{W}_l$

pre-activation $\mathbf{h}_l$

importance switch $\mathbf{s}_l$

output $\mathbf{z}_l = \mathbf{s}_l \circ \sigma(\mathbf{h}_l)$

nonlinearity