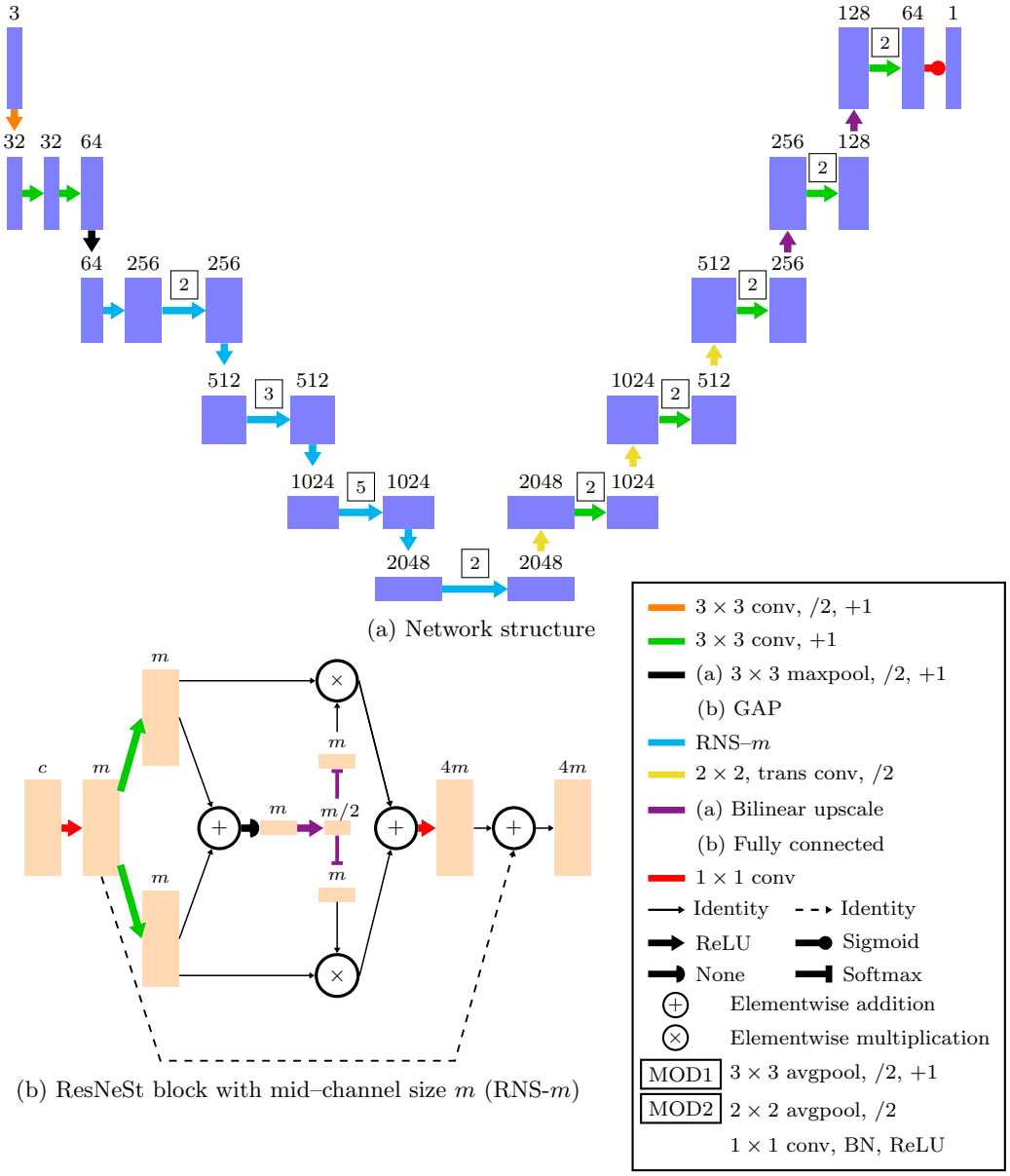


Figure 1. The structure of the neural network: ResNeSt-50 [**?**] + upscale module. Overall, each rectangle and each bold arrow represent an intermediate result (or input) and an operation, respectively. The number at the top of each rectangle indicates the number of channels. If there is a number in the square above the arrow, it means that the arrow repeats that number. The operation types are color-specific and are listed in the lower-right legend. All arrows have batch normalization (BN) after the operation with two exceptions in (b): global average pooling (GAP) and fully connected layers marked with thin lines. The types of the activation functions are indicated by the shape of arrowheads and are also listed in the legend. In the legend, there are notations / and +, which mean stride and padding of the operation with the following number, respectively. All bilinear upscalings are applied with scale factor 2. There are a total of 8 arrows representing RNS-$m$ blocks in (a). For those arrows, $m$ is 64, 64, 128, 128, 256, 256, 512, 512 in that order. Except for the first RNS-$m$, a modification occurs as same as downsampling layers in ResNet [**?**, **?**] whenever $c \neq 4m$ in (b); MOD1 in the legend is added after the first layer, and the dashed identity connection changes to MOD2 in the legend.