Model Definition

**The Custom Convolutional Neural Network (CNN) features 5 layers with convolution, followed by batch normalization, ReLU activation, and max pooling.**

* Input Channels (1-16-32-64-128-256): The input channels double progressively at each convolutional layer, enabling the network to capture both broad and fine-grained features, which are essential for tasks like image classification, detection, and segmentation.
* Kernel Sizes (3x3): A 3x3 kernel is used for general feature extraction, including edges and textures. Experiments with 5x5 and 7x7 kernels resulted in an error, as the kernel size cannot exceed the given input (torch size 64,1,28,28)
* Padding (1): Padding is selected to complement the kernel size, preserving the spatial dimensions of the input and ensuring the output maintains the same height and width.
* Max Pooling (2x2): Max pooling with a 2x2 stride chosen to efficiently downsamples the spatial dimensions, halving them while retaining important features.