

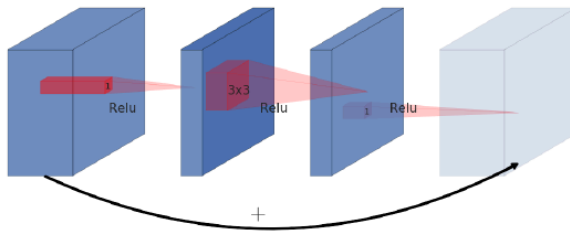
# Mobile-Net V2

The model is based on an inverted residual structure where the short-cut connections are between the thin bottleneck layers. The intermediate expansion layer uses lightweight depthwise convolution to filter features as a source of non-linearity.

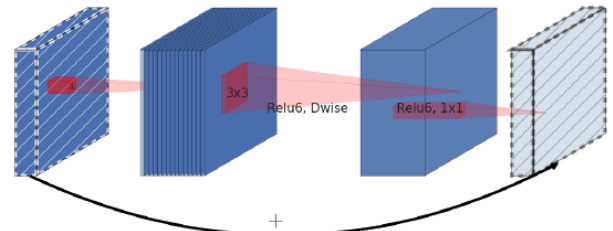
- **Inverted bottleneck residual layer:-**

This module takes as input a low dimensional compressed representation which is first expanded to high dimension and filtered with a lightweight depthwise convolution. Then the features are subsequently projected back to a low-dimensional representation with linear convolution. The module is different from the traditional residual module as instead of connecting the expanded layers (higher no. of channels) it connects the linear bottleneck layers.

(a) Residual block

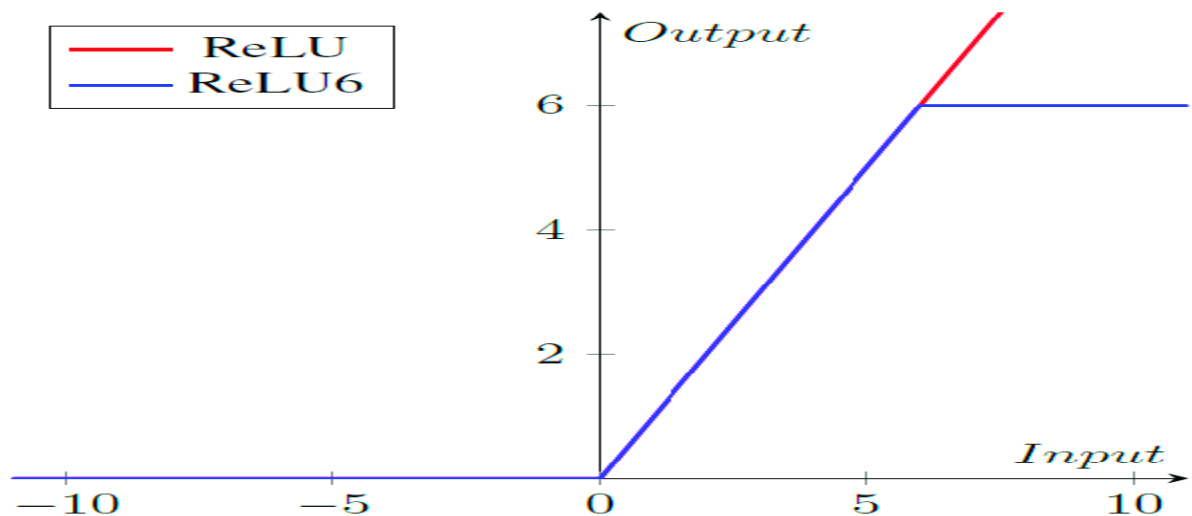


(b) Inverted residual block



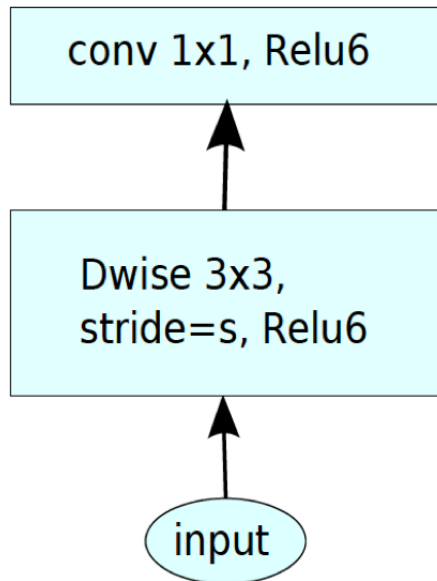
- **Relu-6**

We also see that the paper makes use of the activation function 'Relu6' instead of Relu. Relu6 is a variant of relu where the value is cut-off after the value 6, restricting the output of the activation from 0 to 6. The reason of using Relu6 is so that we can handle floating point numbers and prevent underflow/overflow.

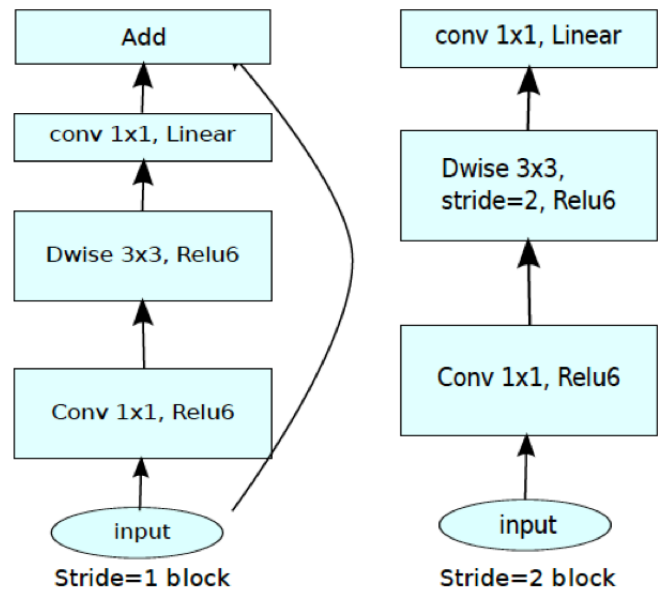


- **Comparisons between Mobilenet V1 and Mobilenet V2**

The main differences between v1 and v2 are that v1 just makes use of the depthwise and pointwise convolution module, whereas v2 is a bit more nuanced as it introduces residual connections between linear bottlenecks. These connections are introduced to improve the gradient flow among the narrow layers as they have densely packed information and represent the capacity of the network. Another difference we observe that in the v2 module we avoid using activation after the linear bottleneck layer. This is done because the authors believe that applying relu in an low-dimensional input space causes inevitable loss of information, but relu is capable of preserving information when the input manifold lies in a low-dimensional subspace of the input space. Hence the activation relu6 is used before and after the depthwise convolution (having expanded input space).



(b) MobileNet[26]



(d) Mobilenet V2