



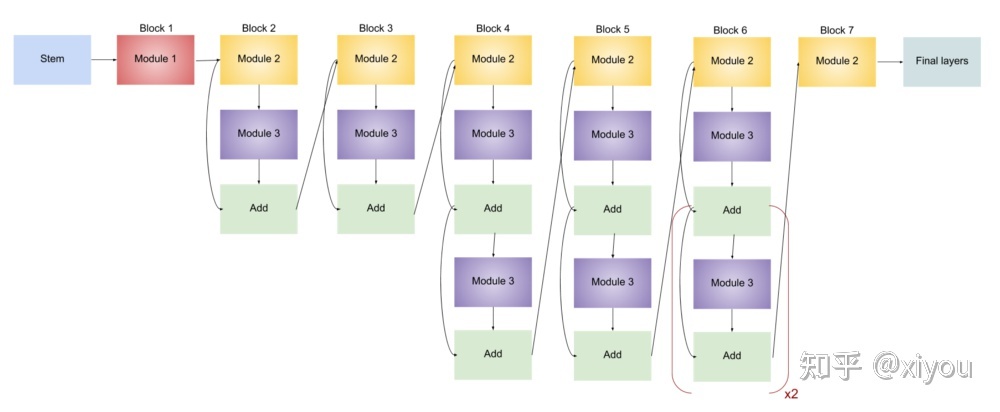
Here I use pre-trained weights for Imagenet.

Image Net is a huge dataset for image recognization, with more than 20 thousand classes.

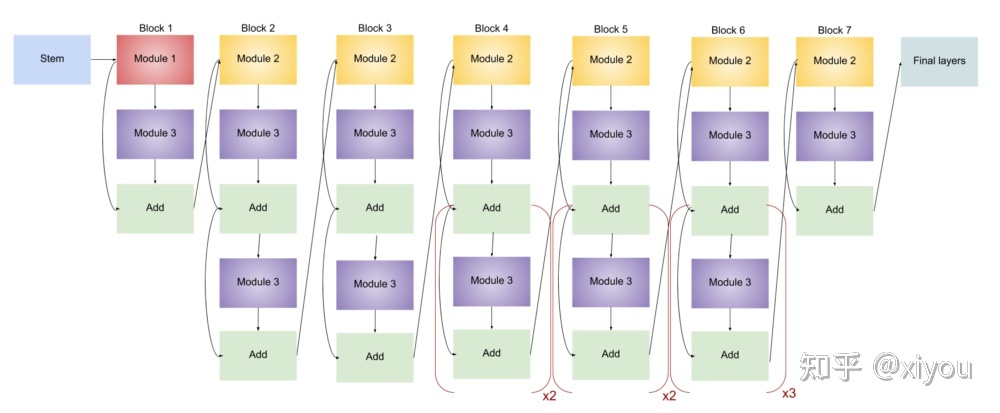
I also compared EfficientNet B0 with pre-trained weights and with no weight parameters, which means this model will train by random initialization.

From B0 to B7, feature maps number increased,

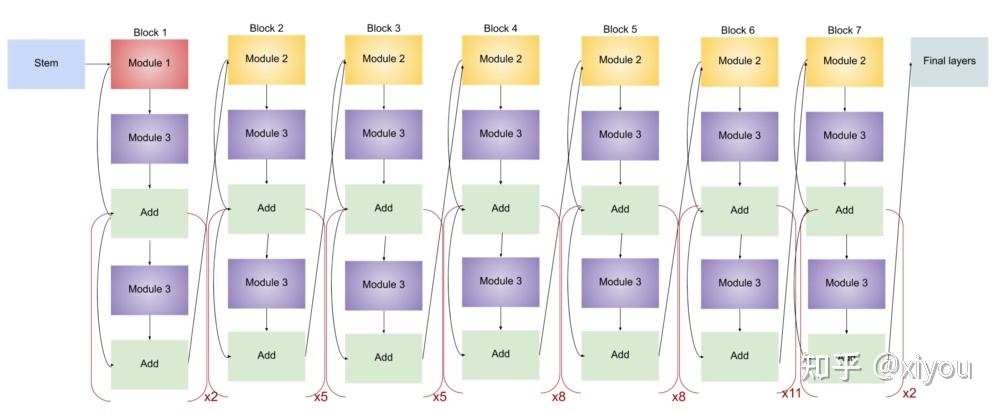
B0



B1

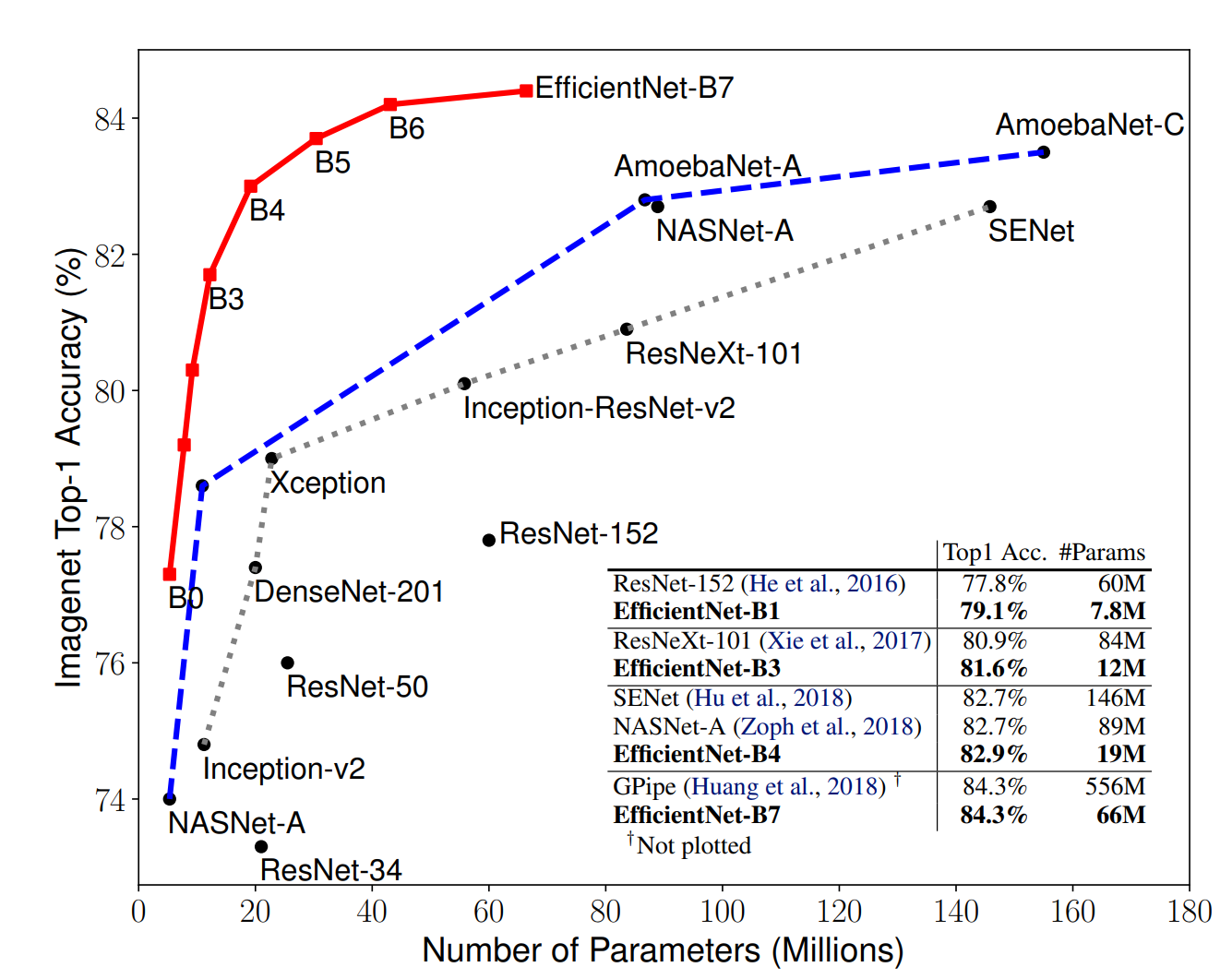


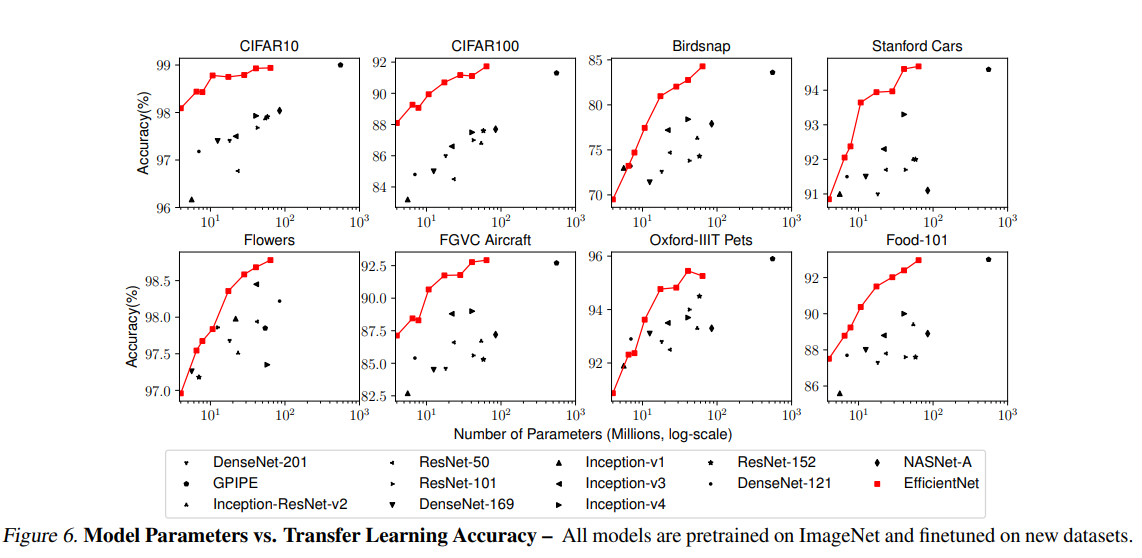
B7



<https://zhuanlan.zhihu.com/p/366738106>

fine adjusting three dimensions, number of parameters are quite lower than other models,like below figure shows.



For CIFAR-10 and some other datasets: 

Krizhevsky, A. (2009). Learning Multiple Layers of Features from Tiny Images. 