**EFFICIENTNET SUMMARY**

**ABSTRACT:-**

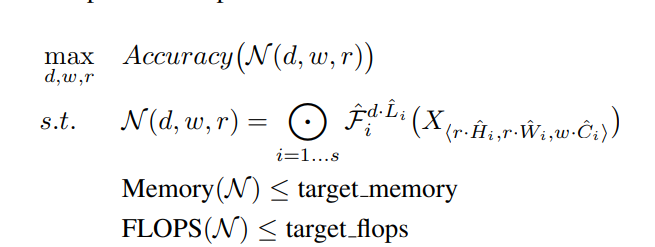
* In this paper, we systematically study model scaling and identify that carefully balancing network depth, width, and resolution can lead to better performance
* A new scaling method that uniformly scales all dimensions of depth, width, and height leads to better performance.
* Used Neural Architecture search to design a new baseline network and scale it to get a family of models **EfficientNets**

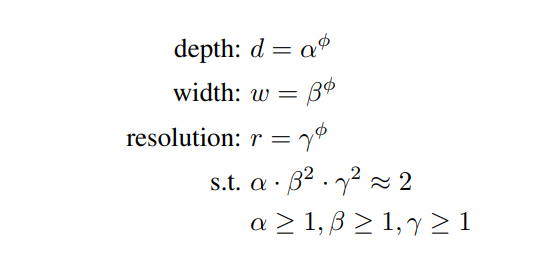
**INTRODUCTION:-**

* In previous work, it is common to scale up only in one of the three dimensions, arbitrarily scaling in two or three dimensions leads to suboptimal performance, so they introduced Compound scaling method.

**Compound scaling method:-**

* + Uniformly scaling the three dimensions with a set of fixed scaling coefficients.
  + Intuitively, the compound scaling method makes sense because if the input image is bigger, then the network needs more layers to increase the receptive field and more channels to capture more fine-grained patterns on the bigger image
  + Scaling is done by optimizing the following relation

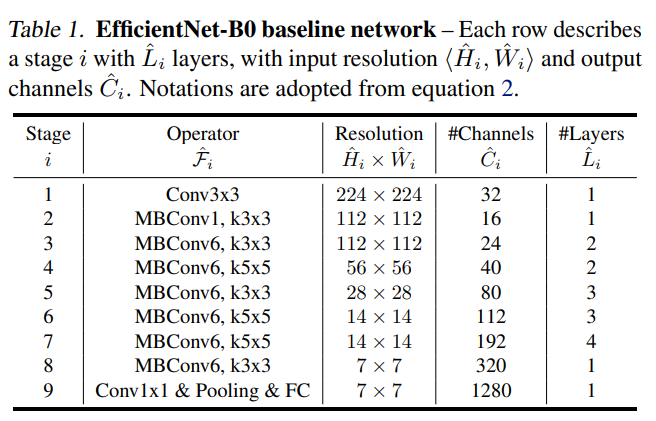




- . In this paper, we constraint α · β2 · γ2 ≈ 2 such that for any new φ, the total FLOPS will approximately increase by 2φ

**EFFICIENT NET ARCHITECTURE:-**

* we develop our baseline network by leveraging a multi-objective neural architecture search that optimizes both accuracy and FLOPS
* use ACC(m)×[FLOPS(m)/T]^w as the optimization goal, where ACC(m) and FLOPS(m) denote the accuracy and FLOPS of model m, T is the target FLOPS and w=-0.07 is a hyperparameter for controlling the trade-off between accuracy and FLOPS.



**NOVELTY:-**

* Compound scaling
* Strong Baseline model

**DRAWBACKS:-**

* For larger models, the search space will be large, computation is difficult
* Several assumptions were made for compound scaling for computational limitations.