Neural architecture search automates the design of neural networks. It’s effective because we optimize the network structure as a reinforcement learning problem instead of randomly choosing the structure. For example, we can scale the model along width, depth, channels, resolution, or a combination of them, and try to find the configuration that maximizes the accuracy with a limited number of FLOPs. An optimization to this process is to train the model that allows using different sizes by progressive shrinking. This decouples the training with NAS and thus reduces time and resources needed for NAS.