AutoHAS: Efficient Hyperparameter and Architecture Search

Xuanyi Dong[1,2], Mingxing Tan[1], Adams Wei Yu[1]
Daiyi Peng[1], Bogdan Gabrys[2], Quoc Le[1]

[1]Google Research, [2]University of Technology Sydney

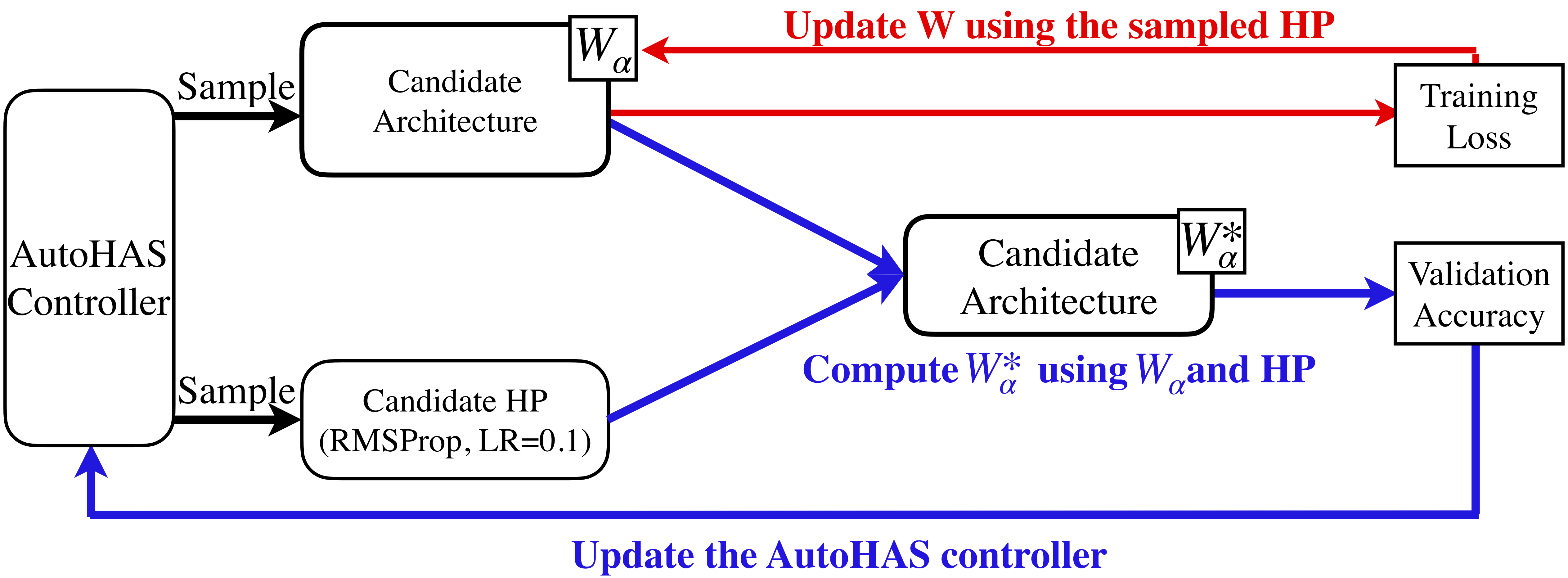
Motivation of AutoHAS

	Model-1	Rank	Model-2
HP-1 (LR=5.5, L2=1.5e-4)	56.9%	>	55.6%
HP-2 (LR=1.1, L2=8.4e-4)	54.7%	<	56.2%

Motivation of AutoHAS

	learning rate	weight decay	augmentation	dropout	architecture	**efficient**
Bayesian	√	√	√	√	√	×
RL or Evolution	√	√	√	√	√	×
PBT	√	√	√	√	×	×
Gradient Descent on LR	√	×	×	×	×	√
Hypergradient	×	√	√	×	√	√
NAS (Weight Sharing)	×	×	×	×	√	√
AutoHAS	√	√	√	√	√	√

AutoHAS Framework



AutoHAS Results

	#Params (MB)	#FLOPs (M)	Accuracy (%)	Search Cost Memory (GB)	Search Cost Time (TPU Hour)
Baseline model	1.5	35.9	50.96	1.0	44.8
AutoHAS (Differentiable)	1.5	36.1	52.17	6.1	92.8
AutoHAS (REINFORCE)	1.5	36.3	53.01	**1.8**	**54.4**

AutoHAS Results

Model	Method	#Params (M)	#FLOPs (M)	Top-1 Accuracy (%)
ResNet-50	Human	25.6	4110	77.20
	AutoHAS	25.6	4110	**77.83** (+0.63)
EfficientNet-B0	NAS	5.3	398	77.15
	AutoHAS	5.2	418	**77.92** (+0.77)