A Appendix

Model	Clean %	FGSM	F-FGSM	PGD	HRS
ResNet-18	93.48	52.43	48.33	24.27	38.53
ResNet-50	94.38	50.05	45.78	23.45	37.57
DenseNet-121	94.76	50.94	47.14	24.06	38.38
DenseNet-169	94.74	53.53	49.47	**26.21**	**41.06**
VGG16 BN	94.07	52.42	46.16	20.03	33.03
DARTS	97.03	58.53	45.03	7.09	13.21
PDARTS	**97.12**	58.67	47.62	9.31	16.99
NSGA Net	96.94	**66.08**	56.16	11.1	19.92
Proxyless-NAS	97.92	51.73	**58.38**	3.22	6.23
PC-DARTS	97.05	60.55	48.65	9.84	17.87

Table 2: Comparison of clean accuracy and adversarial robustness on CIFAR-10 dataset (Top-1 Accuracy)

Model	Clean %	FGSM	F-FGSM	PGD	HRS
ResNet-18	63.87	17.08	17.12	6.05	11.05
ResNet-50	73.09	19	18.12	5.63	10.45
DenseNet-121	78.71	22.9	22.22	7.28	13.33
DenseNet-169	82.44	22.73	21.66	**7.37**	**13.53**
VGG16 BN	72.05	17.09	15.15	4.27	8.06
DARTS	82.43	24.91	16.34	2.32	4.51
PDARTS	83.07	27.69	20.23	3.09	5.96
NSGA Net	**85.44**	**34.93**	**24.1**	2.26	4.40
PC-DARTS	81.83	26.22	18.35	2.93	5.66

Table 3: Comparison of clean accuracy and adversarial robustness on CIFAR-100 dataset (Top-1 Accuracy)

Model	Clean %	FGSM	F-FGSM	PGD	HRS
ResNet18	89.08	32.75	18.03	2.41	4.70
ResNet50	92.86	46.28	26.22	4.68	8.90
DenseNet121	91.97	56.20	38.11	6.932	12.89
DenseNet169	92.81	**61.89**	**44.22**	**10.46**	**18.80**
VGG16	91.52	33.34	13.54	1.55	3.05
DARTS	91.26	54.41	31.18	2.94	5.70
P-DARTS	92.61	55.53	33.87	4.11	7.86
PC-DARTS	92.49	58.90	37.86	4.75	9.04
Proxyless-NAS	92.54	59.56	39.69	6.48	12.11
DenseNAS-Large	92.80	47.91	27.25	2.97	5.76
DenseNAS-R3	**93.81**	54.99	32.11	4.32	8.25

Table 4: Comparison of clean accuracy and adversarial robustness on ImageNet dataset (Top-5 Accuracy)

Model	Clean %	FGSM	F-FGSM	PGD	HRS
ResNet-18	95.48	54.33	51.16	11.23	20.10
ResNet-50	97.31	53.97	52.38	11.36	20.34
DenseNet-121	97.19	67.4	58.61	16	27.48
DenseNet-169	**97.44**	69.11	62.76	18.56	31.18
VGG16 BN	95.24	**72.16**	**66.06**	**27.59**	**42.78**
DARTS Liu et al. (2018)	95.97	64.47	59.95	19.29	32.12
PDARTS Chen et al. (2019)	95.12	55.31	51.16	9.52	17.31
NSGA Net Lu et al. (2018)	92.55	40.05	33.58	2.69	5.23
PC-DARTS Xu et al. (2020)	94.02	54.7	45.3	6.84	12.75

Table 5: Comparison of clean accuracy and adversarial robustness on Flowers-102 dataset (Top-1 Accuracy)

Family	Variant	Params (M)	Clean %	PGD	PP-HRS
Efficient-Net	B0	5.29	91.36	8.11	**14.90**
	B1	7.79	88.89	5.47	7.00
	B2	9.11	92.77	11.40	11.79
	B3	12.23	**93.04**	13.37	10.11
	B4	19.34	92.73	**16.99**	7.86
	B5	30.39	90.95	9.37	2.96
	B6	43.04	91.86	11.71	2.55
	B7	**66.35**	91.57	11.20	1.59
DenseNAS	A	4.77	90.94	1.84	3.61
	B	5.58	91.89	2.13	3.56
	C	6.13	92.31	2.29	3.48
	Large	**6.48**	**92.80**	**2.97**	**4.24**
	R1	11.09	91.33	2.01	**3.93**
	R2	19.47	92.47	3.19	3.51
	R3	**24.66**	**93.81**	**4.32**	3.71
ResNet	18	11.69	89.08	2.41	**4.69**
	50	**25.56**	**92.86**	**4.68**	4.08
DenseNet	121	7.98	91.97	6.93	**12.89**
	169	**14.15**	**92.81**	**10.46**	10.60

Table 6: Comparison of parameter count vs Adversarial accuracy for five different family of architectures on ImageNet dataset

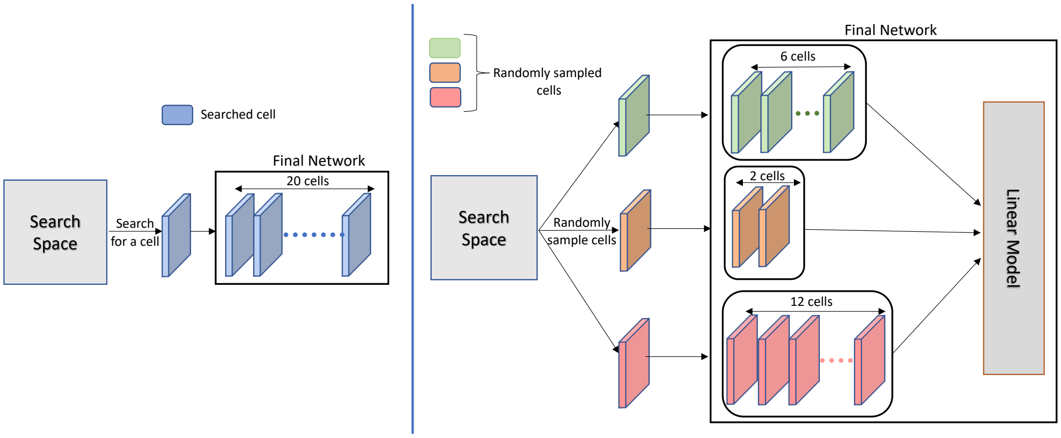

Figure 3: *Left:* Standard procedure for building architectures from DARTS search space; *Right:* Procedure for building ensembles using DARTS search space. 12, 6, 2 can be replaced with any values that sum to 20.

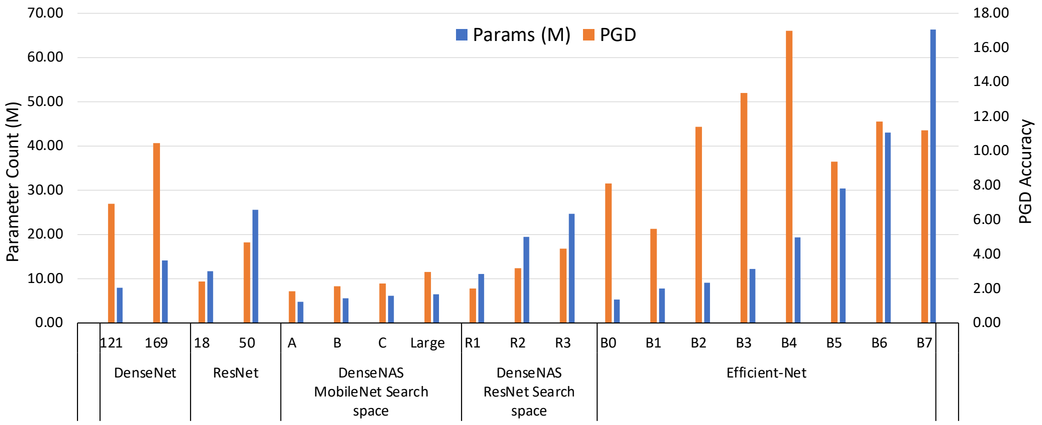

Figure 4: Comparison of PGD accuracy and Parameter count across different family of architectures