Akida 2 Models Product Brief

Image Classification						Dev Site
Architecture	Resolution	Dataset	#Params	Quantization	Accuracy	Download
AkidaNet 0.25	160	ImageNet	483K	8 4	48.50% 41.60%	Y
AkidaNet 0.5	160	ImageNet	1.4M	8 4	61.86% 57.93%	Y
AkidaNet	160	ImageNet	4.4M	8 4	69.94% 67.23%	Y
AkidaNet 0.25	224	ImageNet	483K	8 4	52.39% 46.06%	Y
AkidaNet 0.5	224	ImageNet	1.4M	8 4	64.88% 61.47%	Y
AkidaNet	224	ImageNet	4.4M	8 4	72.16% 70.11%	Y
AkidaNet 0.5	224	PlantVillage	1.2M	8 4	99.61% 98.25%	Y
AkidaNet 0.25	96	Visual Wake Words	227K	8 4	87.03% 85.80%	Y
AkidaNet18	160	ImageNet	2.4M	8	64.77%	Y
AkidaNet18	224	ImageNet	2.4M	8	67.32%	Y
MobileNetV1 0.25	160	ImageNet	469K	8 4	45.72% 37.51%	Y
MobileNetV1 0.5	160	ImageNet	1.3M	8 4	60.27% 54.81%	Y
MobileNetV1	160	ImageNet	4.2M	8 4	69.02% 65.28%	Y
MobileNetV1 0.25	224	ImageNet	469K	8 4	49.63% 42.08%	Y
MobileNetV1 0.5	224	ImageNet	1.3M	8 4	63.65% 59.20%	Y
MobileNetV1	224	ImageNet	4.2M	8 4	71.18% 68.52%	Y
GXNOR	28	MNIST	1.6M	4	98.81%	Y

For 2.0 models, both 8-bit PTQ and 4-bit QAT numbers are given. When not explicitly stated 8-bit PTQ accuracy is given as is. The 4-bit QAT is the same as for 1.0. The quantization scheme indicates both weights and activations bit-width. Weights in the first layer are always quantized to 8-bit. When 'edge' means that the model backbone output is quantized to 1-bit to allow Akida edge learning.

Akida 2 Models Product Brief

Object Detection						Dev Site
Architecture	**Resolution**	**Dataset**	**#Params**	**Quantization**	**Accuracy**	**Download**
YOLOv2 (AkidaNet 0.5 backbone)	224	PASCAL-VOC 2007	3.6M	8 4	50.9% / 27.4% / 27.8% 47.2% / 23.3% / 24.7%	Y
CenterNet (AkidaNet18 backbone)	384	PASCAL-VOC 2007	2.4M	8	70.3% /47.3% / 43.9%	Y
YOLOv2 (AkidaNet 0.5 backbone)	224	WIDER FACE	3.6M	8 4	80.19% 78.60%	Y
Regression						
VGG-like	32	UTKFace (age estimation)	458K	8 4	6.0304 5.8858	Y
Face recognition						
AkidaNet 0.5	112×96	CASIA Webface face ID	2.3M	4	69.79%	Y
AkidaNet 0.5	112×96	CASIA Webface face ID	2.3M	8	72.83%	Y
Segmentation						
AkidaUNet 0.5	128	Portrait128	1.1M	8	90.57%	Y
Key Word Spotting						
DS-CNN	**8 bits**	Google speech command	23.8K	8 4 4 + edge	92.87% 92.67% 90.61%	Y
Point Cloud Classification						
PointNet++		ModelNet40 3D Point Cloud	605K	8 4	0.8088 81.56%	Y

For 2.0 models, both 8-bit PTQ and 4-bit QAT numbers are given. When not explicitly stated 8-bit PTQ accuracy is given as is. The 4-bit QAT is the same as for 1.0. The quantization scheme indicates both weights and activations bit-width. Weights in the first layer are always quantized to 8-bit. When 'edge' means that the model backbone output is quantized to 1-bit to allow Akida edge learning.