

	Specialized Tasks * Tasks in red are shared between teachers	Backbone Architecture	Parameters		
			Learning rate	Batch size	Epochs
Table 1: Main experiments					
Dataset: PASCAL VOC 2007					
Teacher0	1,2,3,5,9,10,11,12,14,16,17,18	DenseNet	0.01	64	50
Teacher1	0,4,6,7,8,10,13,14,15,16,18,19	ResNet	0.01	64	50
Student	0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19	ResNet	0.01	32	50
Dataset: 3D					
Teacher0	0,2	DenseNet	0.01	64	50
Teacher1	1,2,3	ResNet	0.01	64	50
Student	0,1,2,3	ResNet	0.01	32	50
Dataset: CIFAR-10					
Teacher0	1,2,4,5,7,8	DenseNet	0.01	32	50
Teacher1	0,3,6,7,8,9	ResNet	0.01	32	50
Student	0,1,2,3,4,5,6,7,8,9	ResNet	0.01	16	50
Table 2: Combining teachers with heterogeneous architectures (3D)					
Teacher 1: DenseNet, Teacher 2: VGG					
Teacher0	0,1,2	DenseNet	0.01	64	50
Teacher1	0,2,3	VGG	0.01	64	50
Student	0,1,2,3	ResNet	0.01	32	50
Teacher 1: DenseNet, Teacher 2: AlexNet					
Teacher0	0,1,2	DenseNet	0.01	64	50
Teacher1	0,2,3	AlexNet	0.01	64	50
Student	0,1,2,3	ResNet	0.01	32	50
Teacher 1: VGG, Teacher 2: DenseNet					
Teacher0	0,1,2	DenseNet	0.01	64	50
Teacher1	0,2,3	VGG	0.01	64	50
Student	0,1,2,3	ResNet	0.01	32	50
Table 3: Heterogeneous teachers (PASCAL VOC 2007) in supplementary material					
75%					
Teacher0	0,2,3,4,5,6,7,9,10,11,12,13,14,16,17,18	DenseNet	0.01	64	50
Teacher1	0,1,2,3,5,7,8,9,10,11,14,15,16,17,18,19	ResNet	0.01	64	50
Student	0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19	ResNet	0.01	32	50
50%					
Teacher0	0,2,3,4,6,10,11,12,13,14,16,17,18	DenseNet	0.01	64	50
Teacher1	0,1,5,7,8,9,10,14,15,16,17,18,19	ResNet	0.01	64	50
Student	0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19	ResNet	0.01	32	50
25%					
Teacher0	1,2,3,5,9,10,11,12,14,16,17,18	DenseNet	0.01	64	50
Teacher1	0,4,6,7,8,10,13,14,15,16,18,19	ResNet	0.01	64	50
Student	0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19	ResNet	0.01	32	50