

SUPPLEMENTARY MATERIAL FOR ADVERSARIAL META-LEARNING

TREND OF AVERAGE LOSS & TOP-1 ACCURACY FOR 5-WAY 5-SHOT LEARNING ON MINIIAMNET

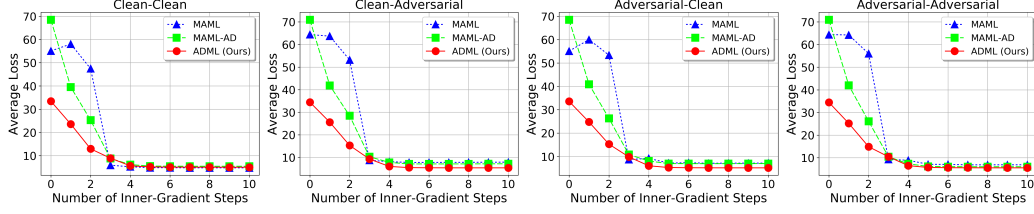


Figure 4: Average loss over the gradient update step for 5-way 5-shot learning on MiniImageNet

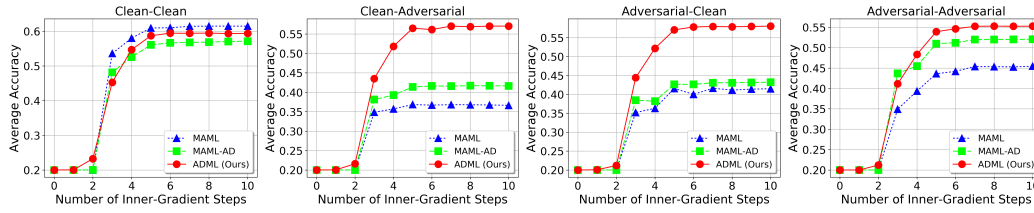


Figure 5: Top-1 accuracy over the gradient update step for 5-way 5-shot learning on MiniImageNet

MODEL SPECIFICATION

For both datasets (i.e, MiniImageNet and CIFAR100), we followed the architecture used by Finn et al. (2017) for image embedding, which contains four 3×3 convolutional blocks with batch normalizations, ReLU activations and 2×2 max-poolings. *Note that this model is the threatened by the aforementioned adversarial attacks.*

CIFAR100 WITH FGSM ATTACK

Note that the maximum perturbations adopted in FGSM Attack are 2 and 0.2.

Table 3: Average classification accuracies on CIFAR100 with FSGM Attack (5-way, 1-shot)

Method	Backbone	Meta-testing	$\epsilon = 2$		$\epsilon = 0.2$	
			Clean	Adversarial	Clean	Adversarial
MAML	32-32-32-32	Clean	57.67 \pm 1.76%	26.40 \pm 1.55%	57.67 \pm 1.76%	43.30 \pm 1.68%
		Adversarial	28.13 \pm 1.56%	28.23 \pm 1.64%	43.03 \pm 1.76%	39.00 \pm 1.70%
MAML-AD	32-32-32-32	Clean	52.70 \pm 1.89%	36.20 \pm 1.65%	52.70 \pm 1.89%	39.17 \pm 1.82%
		Adversarial	37.27 \pm 1.72%	41.67 \pm 1.86%	37.80 \pm 1.70%	37.60 \pm 1.78%
Matching Nets	64-64-64-64	Clean	47.94 \pm 0.56%	25.06 \pm 0.36%	47.68 \pm 0.52%	39.03 \pm 0.51%
		Adversarial	24.82 \pm 0.46%	27.72 \pm 0.43%	40.08 \pm 0.57%	37.79 \pm 0.44%
Relation Nets	64-96-128-256	Clean	58.68 \pm 0.92%	31.11 \pm 0.93%	58.72 \pm 0.90%	45.03 \pm 0.76%
		Adversarial	30.85 \pm 0.92%	30.52 \pm 0.59%	45.85 \pm 1.01%	41.40 \pm 0.80%
R2D2 (64C)	64-64-64-64	Clean	59.76 \pm 2.04%	26.07 \pm 1.00%	59.76 \pm 2.04%	35.53 \pm 1.47%
		Adversarial	27.20 \pm 1.52%	31.51 \pm 1.07%	43.63 \pm 2.17%	37.10 \pm 1.43%
R2D2	96-192-384-512	Clean	60.52\pm2.01%	26.56 \pm 0.93%	60.52\pm2.01%	36.71 \pm 1.45%
		Adversarial	27.90 \pm 1.61%	31.94 \pm 1.26%	43.64 \pm 2.21%	37.73 \pm 1.47%
ADML (Ours)	32-32-32-32	Clean	55.70 \pm 2.00%	50.90\pm1.84%	55.70 \pm 2.00%	49.30\pm1.76%
		Adversarial	54.50\pm1.69%	50.60\pm1.83%	52.90\pm1.92%	45.00\pm1.79%

Table 4: Average classification accuracies on CIFAR100 with FSGM Attack (5-way, 5-shot)

Method	Backbone	Meta-testing	$\epsilon = 2$		$\epsilon = 0.2$	
			Clean	Adversarial	Clean	Adversarial
MAML	32-32-32-32	Clean	74.03 \pm 0.89%	31.29 \pm 0.78%	74.03 \pm 0.89%	54.15 \pm 1.00%
		40%	65.69 \pm 0.92%	36.14 \pm 0.84%	68.99 \pm 0.94%	55.79 \pm 0.98%
		Adversarial	33.34 \pm 0.90%	43.66 \pm 0.86%	59.08 \pm 1.00%	53.93 \pm 0.96%
MAML-AD	32-32-32-32	Clean	67.71 \pm 0.96%	44.61 \pm 0.90%	67.73 \pm 0.96%	56.07 \pm 0.95%
		40%	64.85 \pm 0.90%	53.59 \pm 0.88%	65.93 \pm 0.93%	57.96 \pm 0.93%
		Adversarial	48.37 \pm 0.99%	58.92 \pm 0.97%	59.45 \pm 1.00%	56.33 \pm 0.98%
Matching Nets	64-64-64-64	Clean	62.95 \pm 0.46%	28.14 \pm 0.37%	62.58 \pm 0.49%	47.14 \pm 0.45%
		40%	54.39 \pm 0.48%	28.64 \pm 0.36%	57.86 \pm 0.48%	47.01 \pm 0.48%
		Adversarial	29.40 \pm 0.44%	32.77 \pm 0.42%	53.34 \pm 0.52%	46.50 \pm 0.46%
Relation Nets	64-96-128-256	Clean	75.52 \pm 0.66%	35.37 \pm 0.55%	75.22 \pm 0.70%	55.75 \pm 0.68%
		40%	66.85 \pm 0.79%	36.70 \pm 0.54%	68.67 \pm 0.80%	55.33 \pm 0.69%
		Adversarial	40.46 \pm 0.88%	39.82 \pm 0.57%	60.52 \pm 0.82%	55.50 \pm 0.69%
R2D2 (64C)	64-64-64-64	Clean	76.09 \pm 1.54%	27.83 \pm 1.10%	76.09 \pm 1.54%	38.77 \pm 1.74%
		40%	69.19 \pm 1.53%	38.00 \pm 1.21%	71.39 \pm 1.54%	50.96 \pm 1.55%
		Adversarial	35.14 \pm 1.75%	43.21 \pm 1.19%	58.99 \pm 1.75%	52.20 \pm 1.62%
R2D2	96-192-384-512	Clean	76.29\pm1.44%	29.53 \pm 1.10%	76.29\pm1.44%	40.28 \pm 1.66%
		40%	69.53\pm1.47%	39.32 \pm 1.11%	71.68\pm1.49%	52.04 \pm 1.56%
		Adversarial	35.79 \pm 1.60%	43.28 \pm 1.24%	58.85 \pm 1.89%	52.87 \pm 1.48%
ADML (Ours)	32-32-32-32	Clean	69.90 \pm 0.88%	65.68\pm0.87%	69.90 \pm 0.88%	66.72\pm0.90%
		40%	67.61 \pm 0.93%	62.83\pm0.88%	69.52 \pm 0.88%	63.53\pm0.93%
		Adversarial	65.26\pm0.98%	64.18\pm0.86%	66.81\pm0.95%	66.33\pm0.84%

MINIIMAGENET WITH FFGSM ATTACK

Note that the maximum perturbations adopted in FFGSM Attack are 2, 1 and 0.5, and the step size is set to 10/255.

Table 5: Average classification accuracies on MiniImageNet with FFGSM Attack (5-way, 1-shot)

Method	Backbone	Meta-testing	$\epsilon = 2$		$\epsilon = 1$		$\epsilon = 0.5$	
			Clean	Adversarial	Clean	Adversarial	Clean	Adversarial
MAML	32-32-32-32	Clean	48.47 \pm 1.77%	24.90 \pm 1.39%	48.47 \pm 1.77%	30.73 \pm 1.64%	48.47 \pm 1.77%	39.76 \pm 1.85%
		Adversarial	27.23 \pm 1.61%	23.73 \pm 1.47%	33.97 \pm 1.72%	29.20 \pm 1.67%	42.73 \pm 1.85%	38.63 \pm 1.78%
MAML-AD	32-32-32-32	Clean	40.63 \pm 1.69%	23.03 \pm 0.92%	42.90 \pm 1.88%	33.13 \pm 1.67%	42.43 \pm 1.80%	38.47 \pm 1.75%
		Adversarial	27.87 \pm 1.23%	28.70 \pm 1.65%	33.50 \pm 1.61%	33.47 \pm 1.67%	40.87 \pm 1.82%	39.23 \pm 1.82%
Matching Nets	64-64-64-64	Clean	43.87 \pm 0.41%	25.63 \pm 0.36%	43.87 \pm 0.41%	32.57 \pm 0.50%	43.87 \pm 0.41%	35.54 \pm 0.44%
		Adversarial	26.14 \pm 0.41%	28.96 \pm 0.39%	34.22 \pm 0.48%	33.92 \pm 0.41%	34.99 \pm 0.38%	36.23 \pm 0.43%
Relation Nets	64-96-128-256	Clean	49.67 \pm 0.85%	25.53 \pm 0.46%	49.67 \pm 0.85%	32.64 \pm 0.59%	49.67 \pm 0.85%	42.06 \pm 0.76%
		Adversarial	26.94 \pm 0.82%	24.51 \pm 0.47%	34.07 \pm 0.88%	29.99 \pm 0.61%	42.76 \pm 0.90%	39.37 \pm 0.75%
R2D2 (64C)	64-64-64-64	Clean	49.52 \pm 1.70%	20.51 \pm 0.32%	49.52 \pm 1.70%	22.06 \pm 0.85%	49.52 \pm 1.70%	32.08 \pm 1.60%
		Adversarial	24.46 \pm 1.37%	24.29 \pm 0.85%	29.71 \pm 1.53%	27.15 \pm 0.89%	37.86 \pm 1.76%	34.63 \pm 1.26%
R2D2	96-192-384-512	Clean	51.80\pm1.70%	20.06 \pm 0.26%	51.80\pm1.70%	21.19 \pm 0.54%	51.80\pm1.70%	31.91 \pm 1.42%
		Adversarial	22.68 \pm 1.35%	24.14 \pm 0.92%	26.98 \pm 1.48%	26.94 \pm 0.94%	40.07 \pm 1.79%	34.26 \pm 1.33%
ADML (Ours)	32-32-32-32	Clean	42.20 \pm 1.82%	33.30\pm1.87%	48.60 \pm 1.91%	38.80\pm1.75%	48.80 \pm 1.94%	44.20\pm1.61%
		Adversarial	37.20\pm1.65%	31.00\pm1.71%	40.90\pm1.84%	35.70\pm1.78%	45.10\pm1.78%	39.70\pm1.91%

Table 6: Average classification accuracies on MiniImageNet with FFGSM Attack (5-way, 5-shot)

Method	Backbone	Meta-testing	$\epsilon = 2$		$\epsilon = 1$		$\epsilon = 0.5$	
			Clean	Adversarial	Clean	Adversarial	Clean	Adversarial
MAML	32-32-32-32	Clean	61.45 \pm 0.91%	30.46 \pm 0.70%	61.46 \pm 0.91%	40.91 \pm 0.87%	61.46 \pm 0.91%	53.58 \pm 0.93%
		40%	54.29 \pm 0.94%	30.90 \pm 0.74%	57.77 \pm 0.92%	40.67 \pm 0.87%	59.85 \pm 0.88%	53.38 \pm 0.96%
		Adversarial	32.91 \pm 0.85%	31.90 \pm 0.87%	43.30 \pm 0.89%	41.18 \pm 0.96%	55.06 \pm 0.93%	51.63 \pm 1.01%
MAML-AD	32-32-32-32	Clean	57.55 \pm 0.98%	39.27 \pm 0.93%	58.74 \pm 0.94%	49.53 \pm 0.95%	59.63 \pm 0.95%	56.01 \pm 0.98%
		40%	55.38 \pm 0.93%	40.98 \pm 0.88%	57.35 \pm 0.91%	49.89 \pm 0.92%	59.39 \pm 0.94%	55.83 \pm 0.94%
		Adversarial	41.19 \pm 0.89%	41.17 \pm 0.96%	51.51 \pm 0.89%	49.27 \pm 0.94%	56.97 \pm 0.97%	54.42 \pm 0.99%
Matching Nets	64-64-64-64	Clean	55.99 \pm 0.47%	30.85 \pm 0.45%	55.99 \pm 0.47%	41.28 \pm 0.42%	55.99 \pm 0.47%	48.45 \pm 0.44%
		40%	46.32 \pm 0.41%	32.77 \pm 0.52%	51.66 \pm 0.45%	45.53 \pm 0.46%	53.27 \pm 0.40%	49.61 \pm 0.41%
		Adversarial	32.65 \pm 0.41%	33.56 \pm 0.50%	45.30 \pm 0.45%	46.74 \pm 0.52%	48.98 \pm 0.43%	50.30 \pm 0.45%
Relation Nets	64-96-128-256	Clean	63.85 \pm 0.73%	30.35 \pm 0.50%	63.85 \pm 0.73%	41.74 \pm 0.62%	63.85 \pm 0.73%	54.73 \pm 0.68%
		40%	53.54 \pm 0.83%	28.89 \pm 0.50%	56.15 \pm 0.78%	38.75 \pm 0.65%	60.59 \pm 0.74%	52.59 \pm 0.71%
		Adversarial	37.34 \pm 0.77%	31.31 \pm 0.53%	49.51 \pm 0.75%	42.63 \pm 0.66%	58.90 \pm 0.70%	54.67 \pm 0.71%
R2D2 (64C)	64-64-64-64	Clean	65.48 \pm 1.35%	20.81 \pm 0.45%	65.48 \pm 1.35%	23.44 \pm 0.95%	65.48 \pm 1.35%	38.97 \pm 1.62%
		40%	59.56 \pm 1.58%	28.71 \pm 0.98%	60.27 \pm 1.59%	36.07 \pm 1.40%	61.87 \pm 1.58%	48.33 \pm 1.50%
		Adversarial	30.83 \pm 1.65%	29.80 \pm 0.99%	40.40 \pm 1.68%	38.64 \pm 1.37%	51.03 \pm 1.56%	50.04 \pm 1.46%
R2D2	96-192-384-512	Clean	68.42\pm1.28%	20.61 \pm 0.42%	68.42\pm1.28%	22.97 \pm 0.86%	68.42\pm1.28%	39.53 \pm 1.78%
		40%	60.68\pm1.48%	26.90 \pm 1.05%	61.14\pm1.60%	33.47 \pm 1.14%	63.77\pm1.64%	48.47 \pm 1.51%
		Adversarial	26.50 \pm 1.37%	29.11 \pm 1.02%	35.25 \pm 1.49%	38.00 \pm 1.22%	51.75 \pm 1.56%	50.94 \pm 1.50%
ADML (Ours)	32-32-32-32	Clean	58.68 \pm 0.94%	47.22\pm0.91%	60.22 \pm 0.95%	52.06\pm1.00%	62.07 \pm 0.83%	58.36\pm0.98%
		40%	55.97 \pm 0.84%	43.69\pm0.93%	58.07 \pm 0.87%	51.68\pm0.89%	61.29 \pm 0.88%	57.61\pm0.92%
		Adversarial	50.98\pm0.94%	45.04\pm0.92%	56.56\pm0.91%	52.44\pm0.97%	60.20\pm0.95%	60.25\pm0.98%

MINIIMAGENET WITH RFGSM ATTACK

Note that the maximum perturbations adopted in RFGSM Attack are 0.4, 0.2 and 0.1, the step size is set to 8/255 and the number of steps is 5.

Table 7: Average classification accuracies on MiniImageNet with RFGSM Attack (5-way, 1-shot)

Method	Backbone	Meta-testing	$\epsilon = 0.4$		$\epsilon = 0.2$		$\epsilon = 0.1$	
			Clean	Adversarial	Clean	Adversarial	Clean	Adversarial
MAML	32-32-32-32	Clean	48.47 \pm 1.77%	22.13 \pm 1.27%	48.47 \pm 1.77%	31.73 \pm 1.70%	48.47 \pm 1.77%	45.13 \pm 1.82%
		Adversarial	21.10 \pm 1.51%	22.17 \pm 1.46%	35.03 \pm 1.74%	30.00 \pm 1.73%	42.90 \pm 1.73%	41.23 \pm 1.89%
MAML-AD	32-32-32-32	Clean	41.53 \pm 1.80%	23.13 \pm 1.09%	42.53 \pm 1.87%	34.93 \pm 1.55%	43.07 \pm 1.86%	33.60 \pm 1.48%
		Adversarial	22.67 \pm 1.08%	32.07 \pm 1.69%	35.40 \pm 1.62%	37.27 \pm 1.72%	37.03 \pm 1.65%	38.63 \pm 1.70%
Matching Nets	64-64-64-64	Clean	43.87 \pm 0.41%	23.63 \pm 0.53%	43.87 \pm 0.41%	31.53 \pm 0.46%	43.87 \pm 0.41%	36.32 \pm 0.41%
		Adversarial	25.55 \pm 0.56%	24.61 \pm 0.49%	33.37 \pm 0.42%	34.70 \pm 0.39%	36.59 \pm 0.46%	35.88 \pm 0.51%
Relation Nets	64-96-128-256	Clean	49.67 \pm 0.85%	22.53 \pm 0.39%	49.67 \pm 0.85%	33.07 \pm 0.59%	49.67 \pm 0.85%	44.69 \pm 0.78%
		Adversarial	23.38 \pm 0.75%	22.45 \pm 0.40%	34.30 \pm 0.88%	31.04 \pm 0.63%	45.07 \pm 0.88%	42.12 \pm 0.78%
R2D2 (64C)	64-64-64-64	Clean	49.52 \pm 1.70%	20.34 \pm 0.35%	49.52 \pm 1.70%	22.78 \pm 1.99%	49.52 \pm 1.70%	38.24 \pm 1.66%
		Adversarial	22.08 \pm 1.50%	24.21 \pm 0.78%	30.29 \pm 1.74%	29.68 \pm 1.23%	42.11 \pm 1.73%	39.42 \pm 1.34%
R2D2	96-192-384-512	Clean	51.80\pm1.70%	20.12 \pm 0.31%	51.80\pm1.70%	22.27 \pm 0.81%	51.80\pm1.70%	39.23 \pm 1.56%
		Adversarial	20.94 \pm 1.31%	22.91 \pm 0.72%	28.82 \pm 1.59%	29.23 \pm 1.04%	44.37 \pm 1.88%	38.95 \pm 1.36%
ADML (Ours)	32-32-32-32	Clean	43.20 \pm 1.74%	33.10\pm1.80%	44.50 \pm 1.88%	42.70\pm1.68%	48.90 \pm 1.64%	45.31\pm1.68%
		Adversarial	33.10\pm1.57%	36.90\pm1.80%	41.90\pm1.79%	40.60\pm1.84%	45.70\pm1.68%	42.60\pm1.84%

Table 8: Average classification accuracies on MiniImageNet with RFGSM Attack (5-way, 5-shot)

Method	Backbone	Meta-testing	$\epsilon = 0.4$		$\epsilon = 0.2$		$\epsilon = 0.1$	
			Clean	Adversarial	Clean	Adversarial	Clean	Adversarial
MAML	32-32-32-32	Clean	61.45 \pm 0.91%	25.65 \pm 0.62%	61.46 \pm 0.91%	40.14 \pm 0.88%	61.45 \pm 0.91%	55.62 \pm 0.97%
		40%	52.51 \pm 0.93%	24.80 \pm 0.69%	57.11 \pm 0.91%	42.15 \pm 0.91%	60.63 \pm 0.91%	56.29 \pm 0.96%
		Adversarial	26.92 \pm 0.74%	27.31 \pm 0.71%	44.26 \pm 0.96%	42.27 \pm 0.91%	58.11 \pm 0.99%	54.54 \pm 0.98%
MAML-AD	32-32-32-32	Clean	56.31 \pm 0.99%	27.11 \pm 0.70%	57.52 \pm 0.99%	47.70 \pm 0.97%	58.85 \pm 0.95%	54.21 \pm 0.92%
		40%	50.73 \pm 0.95%	36.37 \pm 0.84%	57.06 \pm 0.90%	51.36 \pm 0.98%	58.89 \pm 0.91%	55.92 \pm 0.93%
		Adversarial	29.69 \pm 0.77%	46.85 \pm 0.93%	50.68 \pm 0.92%	52.72 \pm 0.94%	55.77 \pm 0.97%	54.85 \pm 0.97%
Matching Nets	64-64-64-64	Clean	55.99 \pm 0.47%	24.11 \pm 0.56%	55.99 \pm 0.47%	38.53 \pm 0.41%	55.99 \pm 0.47%	52.13 \pm 0.51%
		40%	47.53 \pm 0.46%	26.81 \pm 0.50%	52.26 \pm 0.44%	36.27 \pm 0.50%	54.83 \pm 0.41%	54.00 \pm 0.46%
		Adversarial	25.33 \pm 0.52%	28.45 \pm 0.55%	40.52 \pm 0.39%	36.61 \pm 0.42%	52.92 \pm 0.46%	53.33 \pm 0.48%
Relation Nets	64-96-128-256	Clean	63.85 \pm 0.73%	24.75 \pm 0.42%	63.85 \pm 0.73%	40.59 \pm 0.64%	63.85 \pm 0.73%	56.94 \pm 0.70%
		40%	52.29 \pm 0.86%	24.24 \pm 0.44%	56.17 \pm 0.81%	39.00 \pm 0.63%	61.36 \pm 0.74%	55.98 \pm 0.73%
		Adversarial	28.69 \pm 0.74%	27.60 \pm 0.47%	48.05 \pm 0.88%	43.20 \pm 0.67%	60.38 \pm 0.77%	56.11 \pm 0.72%
R2D2 (64C)	64-64-64-64	Clean	65.48 \pm 1.35%	20.38 \pm 0.31%	65.48 \pm 1.35%	24.89 \pm 1.10%	65.48 \pm 1.35%	48.47 \pm 1.56%
		40%	59.35 \pm 1.57%	26.89 \pm 0.89%	60.35 \pm 1.50%	38.63 \pm 1.40%	61.83 \pm 1.53%	53.38 \pm 1.51%
		Adversarial	27.11 \pm 1.36%	28.80 \pm 0.98%	38.69 \pm 1.69%	42.28 \pm 1.27%	44.72 \pm 1.75%	54.53 \pm 1.50%
R2D2	96-192-384-512	Clean	68.42\pm1.28%	20.29 \pm 0.26%	68.42\pm1.28%	24.36 \pm 0.92%	68.42\pm1.28%	49.64 \pm 1.68%
		40%	60.04\pm1.53%	24.20 \pm 0.75%	62.15\pm1.54%	35.73 \pm 1.27%	64.54\pm1.61%	54.33 \pm 1.62%
		Adversarial	22.57 \pm 1.33%	27.13 \pm 0.86%	36.22 \pm 1.64%	40.39 \pm 1.14%	56.92 \pm 1.69%	55.49 \pm 1.51%
ADML (Ours)	32-32-32-32	Clean	56.93 \pm 0.89%	43.62\pm0.93%	60.64 \pm 1.03%	53.54\pm0.93%	62.40 \pm 0.91%	58.33\pm0.95%
		40%	55.09 \pm 0.88%	40.54\pm0.91%	58.43 \pm 0.90%	53.83\pm0.93%	62.37 \pm 0.88%	58.08\pm0.94%
		Adversarial	49.36\pm0.87%	47.78\pm0.83%	57.70\pm0.92%	56.77\pm0.89%	61.03\pm0.93%	58.44\pm0.92%

MINIIMAGENET WITH RPGD ATTACK

Note that the maximum perturbations adopted in RPGD Attack are 1.6, 0.8 and 0.4, the step size is set to 2/255 and the number of steps is 40.

Table 9: Average classification accuracies on MiniImageNet with RPGD Attack (5-way, 1-shot)

Method	Backbone	Meta-testing	$\epsilon = 1.6$		$\epsilon = 0.8$		$\epsilon = 0.4$	
			Clean	Adversarial	Clean	Adversarial	Clean	Adversarial
MAML	32-32-32-32	Clean	48.47 \pm 1.77%	25.07 \pm 1.44%	48.47 \pm 1.77%	32.13 \pm 1.67%	48.47 \pm 1.77%	41.37 \pm 1.78%
		Adversarial	27.97 \pm 1.59%	24.77 \pm 1.54%	36.30 \pm 1.77%	30.57 \pm 1.74%	43.20 \pm 1.82%	39.57 \pm 1.76%
MAML-AD	32-32-32-32	Clean	41.03 \pm 1.74%	24.03 \pm 0.87%	42.27 \pm 1.85%	33.87 \pm 1.59%	42.63 \pm 1.87%	36.63 \pm 1.67%
		Adversarial	29.93 \pm 1.49%	29.83 \pm 1.67%	35.37 \pm 1.64%	35.97 \pm 1.76%	37.57 \pm 1.65%	38.07 \pm 1.77%
Matching Nets	64-64-64-64	Clean	43.87 \pm 0.41%	26.36 \pm 0.43%	43.87 \pm 0.41%	31.08 \pm 0.40%	43.87 \pm 0.41%	35.26 \pm 0.56%
		Adversarial	29.36 \pm 0.36%	28.34 \pm 0.39%	33.52 \pm 0.58%	30.02 \pm 0.52%	36.60 \pm 0.51%	33.75 \pm 0.48%
Relation Nets	64-96-128-256	Clean	49.67 \pm 0.85%	26.77 \pm 0.47%	49.67 \pm 0.85%	34.99 \pm 0.64%	49.67 \pm 0.85%	43.67 \pm 0.78%
		Adversarial	28.19 \pm 0.80%	25.32 \pm 0.48%	36.34 \pm 0.88%	32.19 \pm 0.63%	44.18 \pm 0.87%	41.01 \pm 0.77%
R2D2 (64C)	64-64-64-64	Clean	49.52 \pm 1.70%	20.58 \pm 0.39%	49.52 \pm 1.70%	23.46 \pm 0.96%	49.52 \pm 1.70%	34.92 \pm 1.62%
		Adversarial	24.90 \pm 1.43%	25.30 \pm 0.96%	31.70 \pm 1.67%	28.27 \pm 1.21%	40.52 \pm 1.74%	37.49 \pm 1.28%
R2D2	96-192-384-512	Clean	51.80\pm1.70%	20.19 \pm 0.32%	51.80\pm1.70%	22.80 \pm 0.82%	51.80\pm1.70%	34.93 \pm 1.56%
		Adversarial	22.04 \pm 1.26%	25.30 \pm 0.97%	30.63 \pm 1.63%	28.00 \pm 1.04%	42.80 \pm 1.85%	37.18 \pm 1.44%
ADML (Ours)	32-32-32-32	Clean	45.20 \pm 1.85%	40.00\pm1.86%	43.90 \pm 1.82%	40.50\pm1.72%	48.90 \pm 1.82%	44.52\pm1.96%
		Adversarial	37.90\pm1.77%	34.90\pm1.74%	42.60\pm1.64%	39.30\pm1.83%	45.40\pm1.92%	41.50\pm1.83%

Table 10: Average classification accuracies on MiniImageNet with RPGD Attack (5-way, 5-shot)

Method	Backbone	Meta-testing	$\epsilon = 1.6$		$\epsilon = 0.8$		$\epsilon = 0.4$	
			Clean	Adversarial	Clean	Adversarial	Clean	Adversarial
MAML	32-32-32-32	Clean	61.45 \pm 0.91%	32.24 \pm 0.78%	61.45 \pm 0.91%	43.41 \pm 0.92%	61.45 \pm 0.91%	54.96 \pm 0.89%
		40%	55.39 \pm 0.90%	31.55 \pm 0.78%	57.41 \pm 0.95%	43.97 \pm 1.00%	59.99 \pm 0.90%	54.89 \pm 0.94%
		Adversarial	34.48 \pm 0.83%	33.69 \pm 0.83%	46.48 \pm 0.87%	44.67 \pm 0.92%	57.59 \pm 0.91%	54.05 \pm 0.93%
MAML-AD	32-32-32-32	Clean	56.31 \pm 0.98%	42.80 \pm 0.88%	58.56 \pm 0.93%	51.43 \pm 0.97%	60.06 \pm 0.95%	56.18 \pm 0.93%
		40%	53.99 \pm 0.94%	43.58 \pm 0.88%	57.23 \pm 0.97%	51.61 \pm 1.00%	58.35 \pm 0.96%	56.05 \pm 0.94%
		Adversarial	46.41 \pm 0.90%	43.90 \pm 0.90%	53.87 \pm 0.93%	52.68 \pm 0.92%	57.33 \pm 0.91%	55.93 \pm 0.92%
Matching Nets	64-64-64-64	Clean	55.99 \pm 0.47%	33.50 \pm 0.42%	55.99 \pm 0.47%	44.11 \pm 0.36%	55.99 \pm 0.47%	53.31 \pm 0.45%
		40%	49.33 \pm 0.45%	34.58 \pm 0.40%	51.56 \pm 0.42%	44.89 \pm 0.39%	54.02 \pm 0.47%	53.78 \pm 0.53%
		Adversarial	35.02 \pm 0.40%	34.67 \pm 0.39%	43.08 \pm 0.40%	45.84 \pm 0.42%	52.15 \pm 0.49%	51.36 \pm 0.44%
Relation Nets	64-96-128-256	Clean	63.85 \pm 0.73%	32.15 \pm 0.54%	63.85 \pm 0.73%	44.52 \pm 0.64%	63.85 \pm 0.73%	56.53 \pm 0.68%
		40%	54.06 \pm 0.82%	30.69 \pm 0.54%	57.20 \pm 0.80%	41.96 \pm 0.70%	58.53 \pm 0.74%	54.91 \pm 0.75%
		Adversarial	39.64 \pm 0.79%	34.20 \pm 0.57%	51.91 \pm 0.74%	46.04 \pm 0.66%	59.12 \pm 0.70%	56.13 \pm 0.70%
R2D2 (64C)	64-64-64-64	Clean	65.48 \pm 1.35%	21.06 \pm 0.55%	65.48 \pm 1.35%	26.08 \pm 1.14%	65.48 \pm 1.35%	43.64 \pm 1.54%
		40%	59.79 \pm 1.53%	30.65 \pm 1.07%	60.54 \pm 1.59%	39.01 \pm 1.48%	62.36 \pm 1.61%	37.49 \pm 1.28%
		Adversarial	33.70 \pm 1.60%	32.25 \pm 1.25%	42.28 \pm 1.64%	42.32 \pm 1.40%	54.35 \pm 1.70%	53.65 \pm 1.35%
R2D2	96-192-384-512	Clean	68.42\pm1.28%	20.85 \pm 0.53%	68.42\pm1.28%	25.35 \pm 1.13%	68.42\pm1.28%	44.63 \pm 1.70%
		40%	60.66\pm1.56%	28.77 \pm 1.01%	61.48\pm1.60%	36.32 \pm 1.31%	63.15\pm1.59%	52.01 \pm 1.52%
		Adversarial	27.59 \pm 1.39%	31.16 \pm 1.11%	40.29 \pm 1.51%	41.16 \pm 1.34%	55.50 \pm 1.63%	54.00 \pm 1.50%
ADML (Ours)	32-32-32-32	Clean	58.40 \pm 0.94%	48.62\pm0.97%	59.94 \pm 0.86%	54.62\pm1.01%	61.54 \pm 0.97%	57.20\pm0.94%
		40%	56.28 \pm 0.94%	45.83\pm0.90%	58.27 \pm 0.92%	53.60\pm0.99%	60.68 \pm 0.93%	56.33\pm0.98%
		Adversarial	52.28\pm0.88%	45.82\pm0.94%	57.36\pm0.96%	54.39\pm0.96%	59.68\pm0.92%	57.16\pm1.00%