

TABLE 1

The performance (%) of gradual AST method on **MNIST** with various intermediate domain numbers (num.). The results in parentheses represent the improvement compared to the vanilla gradual self-training method. We set perturbation bound $\epsilon = 0.1$ and use a 3-layer convolutional neural network.

num.	24	30	42
\mathcal{A}_{cle}	97.09 (+10.27)	97.40 (+6.65)	97.31 (+6.79)
\mathcal{A}_{adv}	90.24 (+84.45)	90.72 (+84.94)	90.30 (+82.22)

TABLE 3

The performance (%) of gradual AST method on **portraits** with various intermediate domain numbers (num.). The results in parentheses represent the improvement compared to the vanilla gradual self-training method. We set perturbation bound $\epsilon = 0.031$ and use a 3-layer convolutional neural network.

num.	8	10	14
\mathcal{A}_{cle}	84.84 (+2.72)	85.99 (+0.26)	84.44 (+1.60)
\mathcal{A}_{adv}	76.43 (+35.45)	75.05 (+32.43)	76.16 (+33.59)

TABLE 5

The performance (%) of gradual AST method on **MNIST** with various perturbation bound (ϵ). The results in parentheses represent the improvement compared to the vanilla gradual self-training method. We use a 3-layer convolutional neural network.

ϵ	0.15	0.2	0.3
\mathcal{A}_{cle}	97.11 (+8.17)	96.65 (+7.71)	96.59 (+8.65)
\mathcal{A}_{adv}	88.07 (+87.21)	83.90 (+83.58)	83.47 (+83.15)

TABLE 2

The performance (%) of gradual AST method on **MNIST** with various backbone networks. The results in parentheses represent the improvement compared to the vanilla gradual self-training method. We set perturbation bound $\epsilon = 0.1$.

backbone	ResNet18	ResNet50
\mathcal{A}_{cle}	98.70 (+0.42)	98.47 (+0.44)
\mathcal{A}_{adv}	96.57 (+89.11)	96.13 (+22.05)

TABLE 4

The performance (%) of gradual AST method on **portraits** with various backbone networks. The results in parentheses represent the improvement compared to the vanilla gradual self-training method. We set perturbation bound $\epsilon = 0.031$.

backbone	ResNet18	ResNet50
\mathcal{A}_{cle}	86.52 (+0.52)	87.52 (+0.43)
\mathcal{A}_{adv}	78.12 (+52.08)	79.01 (+47.24)

TABLE 6

The performance (%) of gradual AST method on **portraits** with various perturbation bound (ϵ). The results in parentheses represent the improvement compared to the vanilla gradual self-training method. We use a 3-layer convolutional neural network.

ϵ	0.01	0.02	0.04	0.05
\mathcal{A}_{cle}	83.89 (+2.50)	85.19 (+3.80)	84.16 (+2.77)	84.92 (+3.53)
\mathcal{A}_{adv}	79.90 (+9.01)	77.94 (+19.80)	75.65 (+52.94)	71.63 (+61.86)

TABLE 7

The performance (%) of gradual AST method on various datasets. The results in parentheses represent the improvement compared to the vanilla gradual self-training method. We use ResNet50 as the backbone network.

dataset	CIFAR10	CIFAR100
\mathcal{A}_{cle}	79.21 (+1.30)	45.02 (+0.80)
\mathcal{A}_{adv}	33.85 (+33.85)	10.29 (+10.28)

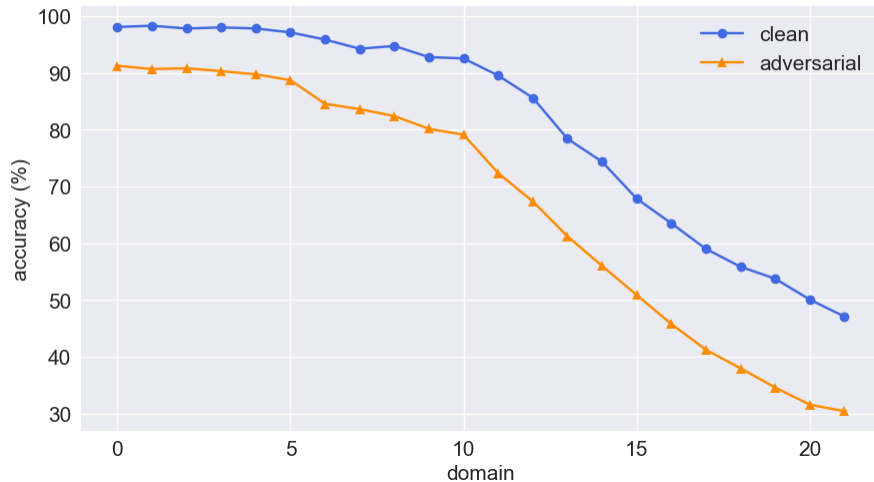


Fig. 1. The source accuracy of the gradual AST method. The abscissa represents the domain on which the model is being trained.