





Figure 3: Performance comparison of existing data subset selection methods and AL baselines on CIFAR100 and ImageNet30 with ResNet-18.

Table 2: Performance comparison of data subset selection and AL on CIFAR10 and CIFAR100 with VGG-19, and on ImageNet30 with VGG-16. The best results are in bold.

Datasets	Select Ratios	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
CIFAR10	Random	76.7±0.8	85.6±0.6	89.8±0.4	91.5±0.3	92.5±0.3	93.1±0.2	93.3±0.2	93.4±0.3	93.7±0.2	94.1±0.2
	Margin	48.3±1.5	81.6±0.7	89.0±0.4	91.4±0.4	92.9±0.3	93.3±0.3	93.5±0.2	<b>93.8±0.2</b>	94.0±0.1	94.1±0.2
	Forgetting	66.7±1.1	85.6±0.5	89.8±0.5	91.5±0.3	92.5±0.3	93.1±0.2	93.3±0.2	93.4±0.2	93.8±0.2	94.1±0.2
	GraNd	52.0±1.3	83.7±0.4	89.7±0.5	92.0±0.4	92.9±0.4	93.4±0.3	<b>93.8±0.2</b>	<b>93.8±0.1</b>	<b>94.1±0.2</b>	94.1±0.2
	kCentGreedy	76.8±0.8	<b>86.1±0.6</b>	88.7±0.4	90.9±0.4	91.8±0.3	92.6±0.2	92.9±0.3	93.5±0.2	93.8±0.2	94.1±0.2
	GraphCut	77.2±0.6	84.9±0.5	88.0±0.4	89.8±0.3	91.1±0.4	92.3±0.3	93.2±0.2	93.5±0.2	<b>94.1±0.2</b>	94.1±0.2
	Glister	76.7±0.7	85.0±0.5	87.9±0.4	90.1±0.4	90.9±0.3	91.8±0.3	92.3±0.3	93.1±0.2	93.5±0.2	94.1±0.2
	AL(Margin)	<b>78.0±0.6</b>	<b>86.1±0.6</b>	<b>89.9±0.4</b>	<b>92.3±0.4</b>	<b>93.1±0.2</b>	<b>93.6±0.2</b>	<b>93.8±0.2</b>	<b>93.8±0.2</b>	<b>94.1±0.2</b>	94.1±0.2
CIFAR100	Random	28.3±1.2	48.9±1.0	58.0±0.7	62.6±0.5	64.8±0.5	67.3±0.4	69.2±0.3	70.9±0.3	71.9±0.3	73.5±0.2
	Margin	14.6±2.2	35.5±1.7	50.0±1.0	58.1±0.7	63.1±0.5	66.7±0.4	69.7±0.3	71.6±0.4	73.3±0.2	73.5±0.2
	Forgetting	<b>29.9±1.9</b>	52.1±1.2	59.0±0.9	63.9±0.6	67.1±0.5	68.6±0.5	69.6±0.4	71.3±0.3	72.5±0.2	73.5±0.2
	GraNd	25.7±2.0	47.2±1.4	57.2±1.1	63.8±0.9	66.6±0.6	68.5±0.5	70.2±0.3	71.9±0.3	72.8±0.2	73.5±0.2
	kCentGreedy	22.2±1.6	49.4±1.3	57.9±0.9	62.7±0.7	66.5±0.5	68.0±0.6	69.3±0.4	71.9±0.3	72.6±0.3	73.5±0.2
	GraphCut	29.9±1.5	49.1±1.1	57.1±0.8	62.4±0.5	65.7±0.6	68.0±0.4	69.2±0.4	70.8±0.3	72.5±0.2	73.5±0.2
	Glister	21.5±1.9	49.4±1.2	57.7±0.8	63.0±0.8	66.0±0.6	67.7±0.5	69.7±0.4	71.1±0.3	72.2±0.2	73.5±0.2
	AL(Margin)	28.2±1.9	<b>49.6±1.0</b>	<b>59.1±0.6</b>	<b>64.6±0.5</b>	<b>69.3±0.4</b>	<b>70.1±0.4</b>	<b>71.9±0.3</b>	<b>73.0±0.2</b>	<b>73.4±0.2</b>	73.5±0.2
ImageNet30	Random	<b>69.6±0.8</b>	80.9±0.5	85.9±0.3	90.1±0.3	91.6±0.3	93.3±0.3	93.7±0.2	94.6±0.3	94.8±0.2	95.7±0.1
	Margin	53.8±1.5	76.3±0.8	84.6±0.5	90.8±0.5	93.1±0.4	94.2±0.4	95.0±0.2	95.2±0.3	95.4±0.2	95.7±0.1
	Forgetting	63.8±1.1	81.4±0.8	88.1±0.6	90.6±0.5	93.0±0.3	93.3±0.3	93.6±0.3	94.6±0.2	95.2±0.2	95.7±0.1
	GraNd	64.3±1.1	80.0±0.8	88.6±0.6	90.9±0.4	92.2±0.3	93.0±0.4	93.8±0.3	94.5±0.2	95.2±0.1	95.7±0.1
	kCentGreedy	66.3±1.0	81.3±0.7	88.7±0.6	90.4±0.4	91.7±0.4	93.3±0.3	93.7±0.2	94.4±0.2	94.9±0.2	95.7±0.1
	GraphCut	68.3±1.2	81.7±0.6	87.3±0.5	89.2±0.3	91.9±0.3	92.8±0.3	93.5±0.2	94.1±0.3	94.9±0.2	96.1±0.1
	Glister	69.1±0.7	80.8±0.5	87.2±0.5	89.6±0.4	91.5±0.3	92.8±0.3	93.5±0.3	94.3±0.2	94.7±0.2	95.7±0.1
	AL(Margin)	69.5±1.2	<b>84.6±0.6</b>	<b>89.1±0.6</b>	<b>92.5±0.4</b>	<b>93.8±0.4</b>	<b>94.9±0.3</b>	<b>95.3±0.3</b>	<b>95.4±0.2</b>	<b>95.7±0.1</b>	95.7±0.1

Table 3: Effect of incorporating random initial set and multi-round selection into data subset selection on CIFAR10 with ResNet-18.

Select Ratio	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Margin	73.2±1.3	85.5±0.9	91.3±0.5	93.6±0.3	94.5±0.2	94.9±0.3	95.1±0.1	95.4±0.2	<b>95.5±0.2</b>	95.5±0.2
Margin + Random Init	81.2±1.2	88.4±0.7	92.1±0.5	93.9±0.4	94.7±0.3	95.1±0.2	<b>95.4±0.2</b>	<b>95.5±0.2</b>	<b>95.5±0.2</b>	95.5±0.2
Margin + Multi Round	80.1±1.0	89.2±0.5	93.0±0.3	94.3±0.4	94.8±0.3	<b>95.3±0.3</b>	<b>95.4±0.2</b>	<b>95.5±0.2</b>	<b>95.5±0.2</b>	95.5±0.2
AL(Margin)	<b>84.5±0.7</b>	<b>91.0±0.5</b>	<b>93.9±0.4</b>	<b>94.5±0.3</b>	<b>95.3±0.2</b>	<b>95.3±0.2</b>	<b>95.4±0.2</b>	<b>95.5±0.2</b>	<b>95.5±0.1</b>	95.5±0.2