

A. Experiment Details

A.1. Backbone Networks

ResNet-18 and multilayer perceptron are the backbone models in our work. The networks are separated into encoder and head and their architectures are shown in Tab. 6. In original settings for supervised contrastive learning, the training is in two phases: 1) the encoder and head are trained together using supervised contrastive loss; 2) the head is removed and a downstream classifier is trained with the encoder outputs.

A.2. Data Augmentation

For supervised contrastive learning, we applied the data augmentation approaches of horizontal and vertical *flip*, *color jitter* (*brightness* = 0.4, *contrast* = 0.4, *saturation* = 0.4, *hue* = 0.1), *gray scaling* (*probability* = 0.2), and *Gaussian blurring* (*kernelsize* = 9) on the images in *CIFAR100* and *Tiny ImageNet*.

A.3. Hyperparameters

The hyperparameters of our method are from two categories, namely the backbone networks and *OpenIncrementNN*. For the backbone networks, the hyperparameters are learning rate (lr), epochs for the t th training session (E_t), loss balance (α and λ_{dis}), and temperature in supervised contrastive loss (τ). Hyperparameter in *OpenIncrementNN* is the number of neighboring samples (K). The hyperparameters in backbone networks were searched using grid search. The hyperparameters in *OpenIncrementNN* were determined using heuristics. Their search space and values are listed in Tab. 7.

B. Extra Results

We have plotted the changes in inlier classification accuracy and AUROC for OSR on **CIFAR-100** dataset to see that they are positively correlated. We think such a correlation is due to the change in feature maps, i.e., the feature distortions. We leave the research on this problem in future work.

Block Name	Configurations
	$7 \times 7 \times 64$, stride 1, <i>Conv 2D</i>
	$1 \times 1 \times 64$, stride 1, <i>Conv 2D</i> <i>Batch Normalization</i>
Conv2_x	$3 \times 3 \times 64$, stride 2, <i>Conv 2D</i> <i>Batch Normalization</i>
	$1 \times 1 \times 64$, stride 1, <i>Conv 2D</i>
	$1 \times 1 \times 128$, stride 1, <i>Conv 2D</i> <i>Batch Normalization</i>
Conv3_x	$3 \times 3 \times 128$, stride 2, <i>Conv 2D</i> <i>Batch Normalization</i>
	$1 \times 1 \times 128$, stride 1, <i>Conv 2D</i>
Encoder	$1 \times 1 \times 256$, stride 1, <i>Conv 2D</i> <i>Batch Normalization</i>
Conv4_x	$3 \times 3 \times 256$, stride 2, <i>Conv 2D</i> <i>Batch Normalization</i>
	$1 \times 1 \times 256$, stride 1, <i>Conv 2D</i>
	$1 \times 1 \times 512$, stride 1, <i>Conv 2D</i> <i>Batch Normalization</i>
Conv5_x	$3 \times 3 \times 512$, stride 2, <i>Conv 2D</i> <i>Batch Normalization</i>
	$1 \times 1 \times 512$, stride 1, <i>Conv 2D</i>
Average Pooling	1×1 , Average Pooling
Head	<i>Flatten</i> <i>Linear</i> (512, 512) <i>Linear</i> (512, 128)

Table 6: Network architecture of *ResNet-18* trained in our paper. The encoder consists of five blocks and the head is one fully connected layer. * stands for the dimension size after flattening. It differs between different input image sizes.

	Backbone Networks			<i>OpenIncrementNN</i>	
	Parameter	Value	Search Space	Parameter	Value
CIFAR100	lr	0.001	[0.001, 0.005, 0.01]	K	10
	E_0	100	[100, 200, 300]		
	$E_{t>0}$	200	[100, 200, 300]		
	α	0.2	[0.05, 0.1, 0.2]		
	λ_{dis}	0.5	[0.2, 0.5, 0.8]		
	τ	0.05	[0.01, 0.05, 0.1]		
Tiny ImageNet	lr	0.001	[[0.01, 0.005, 0.001]]	K	20
	E_0	100	[100, 200, 300]		
	$E_{t>0}$	200	[100, 200, 300]		
	α	0.2	[0.05, 0.1, 0.2]		
	λ_{dis}	0.5	[0.2, 0.5, 0.8]		
	τ	0.05	[0.01, 0.05, 0.1]		

Table 7: Hyperparameters in the backbone networks and OpenIncrement framework that adopted in our experiments. The hyperparameters of the backbone networks are determined using grid search. The search space is given as well.

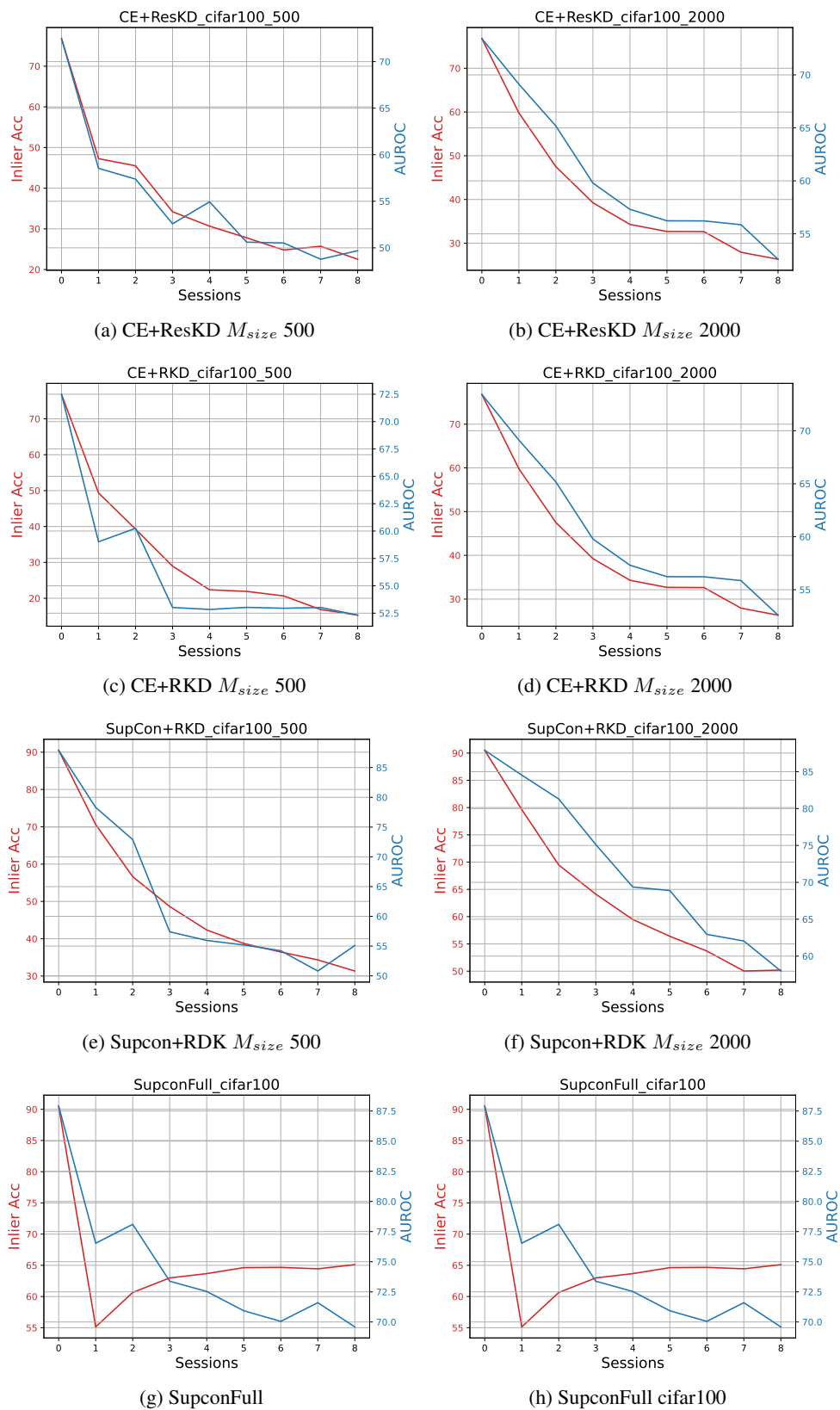


Figure 3: Joint plots on changes of inlier classification accuracy and AUROC for OSR on **CIFAR-100** dataset with the memory size of 500 and 2000. Since exemplar sizes in *Joint* settings are all the same, the last two graphs are identical.