SUPPLEMENTARY MATERIAL

A. Datasets

Table 2 shows a summary of the datasets utilized in our work along with their size and number of classes and budget size.

Dataset	#Classes	Train + Val	Test	Initially Labeled	Budget	Image Size
CIFAR10 [29]	10	$45000 + 5000$	10000	5000	2500	32×32
CIFAR100 [29]	100	$45000 + 5000$	10000	5000	2500	32×32
Caltech-256 [22]	256	$27607 + 3000$	2560	3060	1530	224×224
ImageNet [6]	1000	$1153047 + 128120$	50000	128120	64060	224×224
BDD100K [57]	19	$7000 + 1000$	2000	800	400	688×688
Cityscapes [5]	19	$2675 + 300$	500	300	150	688×688

Table 2. A summary of the datasets used in our experiments. CIFAR10, CIFAR100, Caltech-256 and ImageNet are datasets used for image classification, while BDD100K and Cityscapes are large scale segmentation datasets. The budget for each dataset is the number of images that can be sampled at each training iteration.

B. Hyperparameter Selection

Table 3 shows the hyperparameters found for our models through a grid search.

Experiment	d	α_1	α_2	α_3	λ_1	λ_2	β	batch size	epochs
CIFAR10	32	5×10^{-4}	5×10^{-4}	5×10^{-4}	1	1	1	64	100
CIFAR100	32	5×10^{-4}	5×10^{-4}	5×10^{-4}	1	1	1	64	100
Caltech-256	64	5×10^{-4}	5×10^{-4}	5×10^{-4}	1	10	1	64	100
ImageNet	64	10^{-1}	10^{-3}	10^{-3}	1	10	1	64	100
BDD100K	128	10^{-3}	10^{-3}	10^{-3}	1	25	1	8	100
Cityscapes	128	10^{-3}	10^{-3}	10^{-3}	1	25	1	8	100

Table 3. Hyperparameters used in our experiments for VAAL. d is the latent space dimension of VAE. α_1, α_2, and α_3 are learning rates for VAE, discriminator (D), and task module (T), respectively. λ_1 and λ_2 are the regularization parameters for transductive and adversarial terms used in Eq. (4). β is the Lagrangian parameter in Eq. (1).

Figure 6 shows the performance of our method is robust to the choice of the architecture by having consistently better performance over Core-set [43] on CIFAR100.
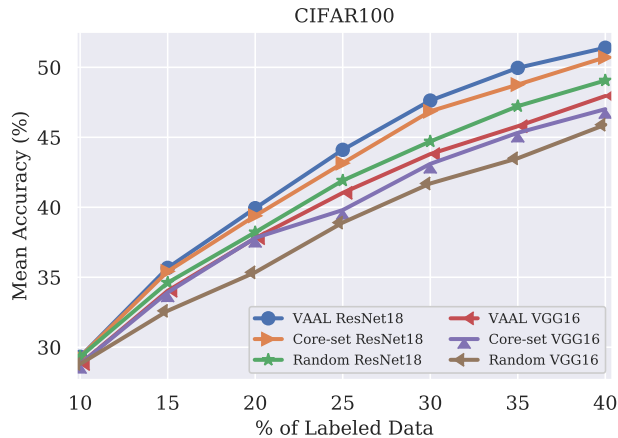

Figure 6. Performance of VAAL using ResNet18 and VGG16 on CIFAR100