

## A Supplementary

### A.1 Access to datasets and benchmark

The source code is available at <https://github.com/Emprime/dcic>. The datasets are available at <https://doi.org/10.5281/zenodo.7152309>.

Please check the licenses which vary between datasets and source code. You also need to credit all sources appropriately if you want to use them.

### A.2 Additional Results

#### A.2.1 More annotations are better

As stated in the main paper, the datasets Benthic, MiceBone, Pig, QualitMRI, and Treeveritsy#6 have an  $\hat{ACC}$  of lower than 85% which marks them as difficult even for humans. Thus, we call them *difficult* and the rest *easy* datasets in the following.

We include in Figure 7a all measured scores over the all calculated budgets. We see as in Figure 5a an improvement of the scores with increasing budget. For easier datasets, we see in Figure 7b that the improvements with additional annotations are lower or do not exist, supporting the observation that uncertain labels benefit more from additional annotations. If we check in Figure 7c which methods benefit the most from additional annotations, we see up to 0.15 lower relative  $KL$  and  $ECE$  for soft labels. We can confirm similar observations on a t-SNE [19] visualization of the Plankton dataset in Figure 6. We use the method Pseudolabel v2 soft for these experiments. For a budget of 10%, the clusters are not as distinct as with 100% while for a budget of 1000% we can see more linear transitions in alignment with the soft ground truth label. In agreement with [22], we see that more annotated data helps to structure the feature space while linear transitions clarify with the added benefit of soft labels.

#### A.2.2 Correlations

Extending on the mentioned correlations in the main paper, we want to highlight to important conclusion from the investigated correlations.  $F1$  balances the precision and recall and is general not the same as  $ACC$  but in our experiments we see almost identical values to  $ACC$  which we

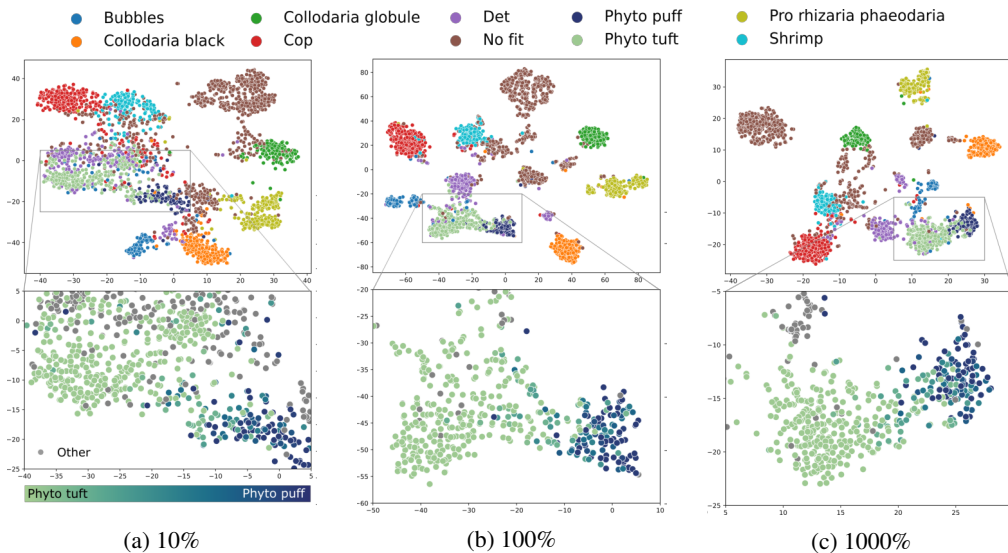


Figure 6: T-SNE evaluation on Plankton dataset on a budget of 10%, 100% and 1000% – On the top, all images are colored based on the majority class of the ground truth distribution. The lower images are zoomed in regions where we color only two classes accordingly to their actual ground truth distribution.

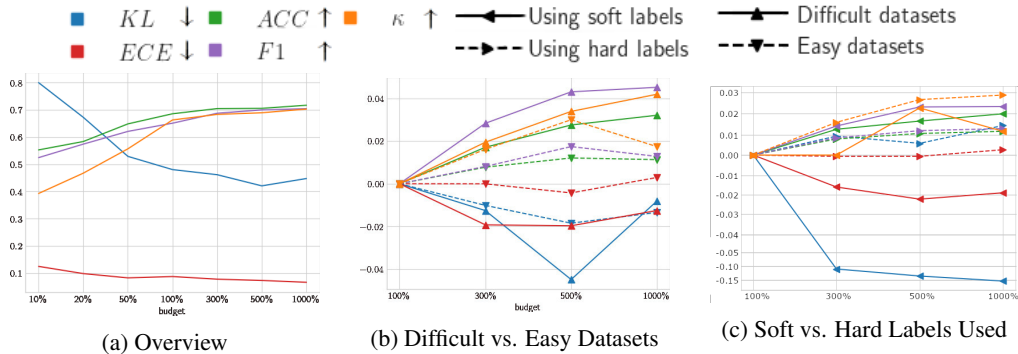


Figure 7: Impact of increased budget across all methods and datasets – The last two images show the relative changes in comparison to a budget of 100% from different subgroups of the datasets and methods respectively.

credit to the averaging per class. A high correlation also exists between  $ACC$  and  $\kappa$  which means that higher consistency in the input labels for different folds leads to better classification scores. This is a particularly useful result since consistency can be calculated without training a network or knowing anything about the ground truth distribution. This discovery allows the opportunity to roughly estimate the  $ACC$  of dataset without the knowledge of any labels.

### A.3 Further implementation details and insights

#### A.4 Details about $\hat{ACC}$

In this section, we describe in detail why  $\hat{ACC}$  measures the human performance on the baseline method.  $\hat{ACC}$  measures the impact of the initialized labels against the ground truth for the baseline method. This is special for the baseline because it does not change the initialised labels. The ground truth is based on human annotations and the initialised labels on the ground truth and thus also human annotations.

The difference is the following: For example, if we have 20 annotations for an image  $x$  and the resulting ground truth distribution for the image would be  $l_x = [0.8, 0.05, 0.15]$  for three classes.. The initialized labels are sampled from  $l_x$ . This means for one annotation per image ( $m = 1$ ) defined by an initialization scheme, the initialized label would be  $[1\ 0\ 0]$  in 80% of the cases,  $[0\ 1\ 0]$  in 5% of the cases and  $[0\ 0\ 1]$  in 15% of the cases. In this case, the initialized label often represents the majority vote, but in about 20% of the cases it would be a hard vote for a clear minority. For  $\hat{ACC}$  we only look at the majority vote of the ground truth and thus the expected result would be class 0 for image  $x$  even if we feed in 20% of the cases a different training label for image  $x$  in the training. This would of course negatively impact the classification performance as it would confuse the network.

An important detail is that we normally do not use the initialized label directly for the evaluation (it is not allowed to initialize labels on the test set). However, the baseline is designed to forward the sampled initialized labels in the test stage which effectively means that we sample one ( $m=1$ ) or multiple human(s) ( $m>1$ ) from the ground truth distribution. By doing so, the  $\hat{ACC}$  measures the performance of the sampled annotation against the complete ground truth.

In short,  $\hat{ACC}$  measure the impact of sample human annotation against the complete ground truth which we interpret as the performance of a human annotator against the combined ground truth.

#### A.5 About the improvement of Semi- and Self-supervised methods

In this section, we describe how Semi- and Self-supervised methods can be used to improve the labels in the first phase. When we talk about self-supervised methods, we mean the self-supervised training on a proxy task (e.g. maximizing agreement in SimCLR) and then a fine-tuning to the given classification task.

The improvement can be credited to two properties. Firstly, in many experiments the labels are initialized only on a small portion of the dataset and thus predicting a label on the rest of the unlabeled data creates / improves them in contrast to the previously unknown state. Secondly, when the network generalizes well, errors or ambiguities in the training data should be mitigated. For example, the network should predict a similar probability distribution for similar images even if the used training labels diverge. As seen in Figure 5a, these two benefits lead to a lower  $KL$  score in the second phase in comparison to the first phase  $\hat{KL}$ .

### A.5.1 Self-supervised methods

We investigated the four self-supervised methods BYOL [21], MOCOv2 [13], SimCLR [11] and SWAV [10]. On large datasets like ImageNet [33] these methods have been shown to be comparable to or even surpass semi-supervised methods like FixMatch [67]. However, we assume all self-supervised methods to be inferior to most other benchmarked methods due to three main reasons. Firstly, the self-supervised methods are more difficult to tune. The models learn in the first phase without any supervision and thus depend on properly selected image augmentations [21]. For better comparability, we enforced the same set of hyperparameters including image augmentations across all datasets. This was problematic for some methods on specific datasets (e.g., BYOL on CIFAR-10H). Secondly, self-supervised methods are typically trained on large datasets. We trained only on a few thousand images and restricted the batch size for a comparable hardware usage. These two restrictions are often not considered in the literature and therefore might negatively impact the performance. Thirdly, many of the compared methods are pretrained on ImageNet. If we compare to not-pretrained methods such as Pseudo v2 not, we see that the self-supervised method is performing comparable or better.

### A.5.2 Hyperparameter optimization

We used the public framework Ray Tune <sup>4</sup> for the hyperparameter optimization of the evaluation phase. As stated in the main paper we determined them by applying Hyperopt [7] with 100 search trials across the same grid of parameters for all datasets. The target was the minimization of  $KL$  between  $\Phi(x)$  and  $l_x$  for the baseline experiment with exactly ten annotations per image across one fold. Below we show the used parameter grid.

```
config={
  'weights': weights,
  'batch_size': tune.choice([8,32,128]),
  'epochs': tune.choice([20]),
  'lr': tune.choice([1e-1,3e-1,1e-4,1e-5]),
  'dropout': tune.choice([0, 0.5]),
  'network': tune.choice(['wideresnet28-10', 'resnet50v2_large',
    'densenet121', 'inceptionv2', 'resnet50v2', 'wideresnet16-8' ]),
  'augmentation': tune.choice([0,1,2,3]),
  'opt': tune.choice(['sgd', 'adam', 'sgdw', 'sgdwr']),
  'input_upsampling': tune.choice([True, False]),
  'weight_decay': tune.choice([5e-4, 1e-3]),
},
```

As initialization, we used a heuristically determined set of hyperparameters from our preliminary studies. The details about these parameters can be found in the source code.

## A.6 Detailed dataset descriptions

In this section, we describe in detail from where and how we collected the raw data, how we preprocessed the data and the labels, and give additional information per dataset. For all datasets, we split the data into five folds. We ensured during the splitting that the class distribution and the total size are similar between the folds. Moreover, we ensured that very similar images, e.g. consecutive slices in MiceBone are in the same fold. In multiple cases, we simplified the class labels due to very long tail distributions or missing data. In these cases, we give the mapping as 'original\_label' : 'our\_label'. If our label is None, we ignored the class completely.

<sup>4</sup><https://docs.ray.io/en/latest/tune/index.html>

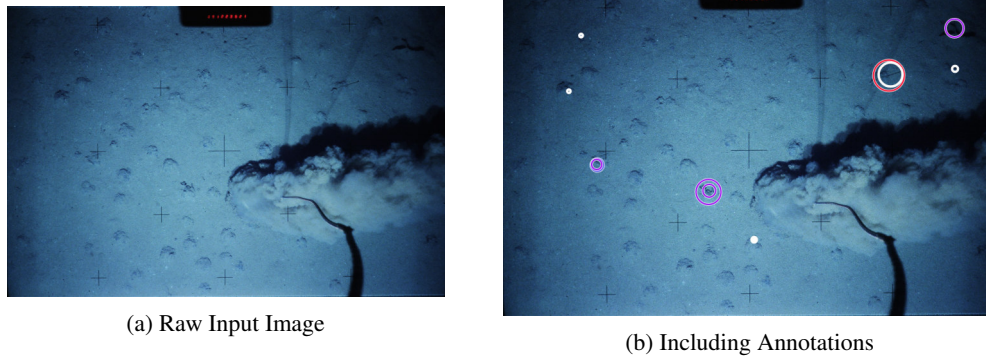


Figure 8: Raw example image without and with annotations. Every circle represents one annotation.

### A.6.1 Benthic

The raw data was provided to us by the authors of [65, 37]. In their studies, multiple annotators labeled any objects they could detect on seafloor images. One example image without and with annotations is given in Figure 8.

We see that some items are only annotated once while others have multiple annotations. We sorted all annotations in descending order of their size. We merged all overlapping annotations with nearby centers. If at least three annotations were found, we used a cutout region around the largest found annotation as an image and the average across all annotations as the ground truth label. The total number of annotations and the agreement with the majority vote per image can be seen as a histogram in Figure 9. In agreement with the original authors, we simplified the original class with the following mapping:

```

"Anemone": "other_fauna",
"Coral": "coral",
"Crustacean": "crustacean",
"Epifauna": "other_fauna",
"Ipynops fish": "other_fauna",
"Jellyfish": "other_fauna",
"Litter": None,
"Mollusc": "other_fauna",
"Octopus / squid": None,
"Ophiuroid": "star",
"Other fauna": "other_fauna",
"Other fish": "other_fauna",
"Polychaete (sessile)": "worm",
"Sea Urchin": "other_fauna",
"Sea cucumber": "cucumber",
"Small encrusting": "encrusting",
"Spiral worm": "worm",
"Sponge": "sponge",
"Stalk (no head)": "other_fauna",
"Stalked crinoid": "other_fauna",
"Stalked sponge": "other_fauna",
"Starfish (not ophiuroid)": "star",
"polychaete (mobile)": "worm",
'lost+found': None,

```

### A.6.2 CIFAR-10H

This dataset was originally described and created by [54]. The total number of annotations and the agreement with the majority vote per image can be seen as a histogram in Figure 10.

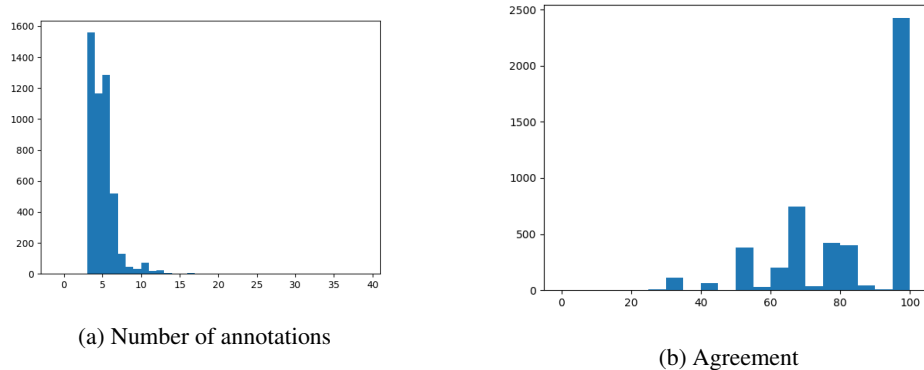


Figure 9: Histograms for the Benthic Dataset

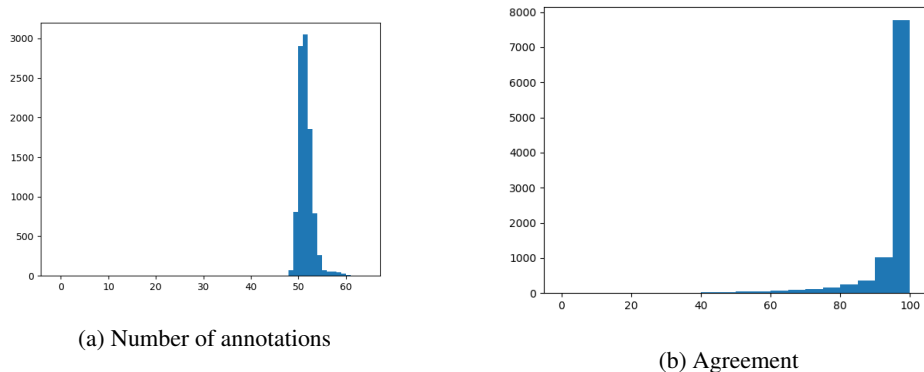


Figure 10: Histograms for the CIFAR-10H Dataset

### A.6.3 MiceBone

We used the preprocessing described in [64] and reproduced their description here with the authors permission:

The Mice Bone dataset is based on the raw data which was published in [60]. The raw data are 3D scans from collagen fibers in mice bones. The three proposed classes are similar ("g") and dissimilar ("ug") collagen fiber orientations and not relevant ("nr") regions due to noise or background. We used the given segmentations to cut image regions from the original 2D image slices which mainly consist of one class.

We collected the annotations with four human workers. Three of them were hired workers which replied to a local job advertisement and were paid 11€ per hour. The last worker was a domain expert who desired no further payment than their contribution to science. All workers received reference images as guidance from the domain expert and oral explanations of the different classes. All reference images are shown in Figure 11 with the written explanations. The annotation process was structured in a partial first annotation of the data and multiple following annotations. The first partial annotation was used to get familiar with the annotation process and the classes. No results from this first part were used in this publication. We checked that all annotators achieved a similar classification performance than the domain expert with himself after this first part. All annotators annotated as much as possible during a fixed time frame of 15 hours. We either annotated the full dataset or subset of 10% multiple times. The total number of annotations and the agreement with the majority vote per image can be seen as a histogram in Figure 12.

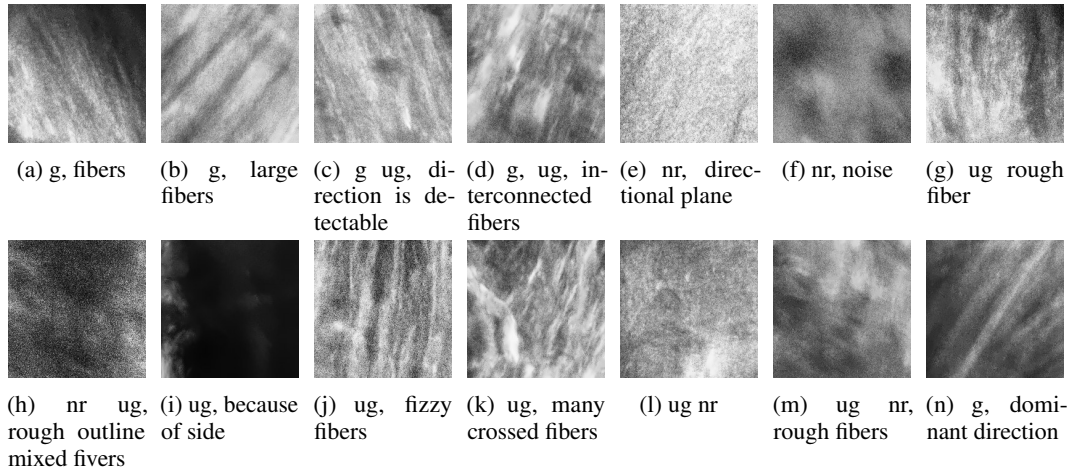


Figure 11: Reference images for the annotation of the MiceBone Dataset

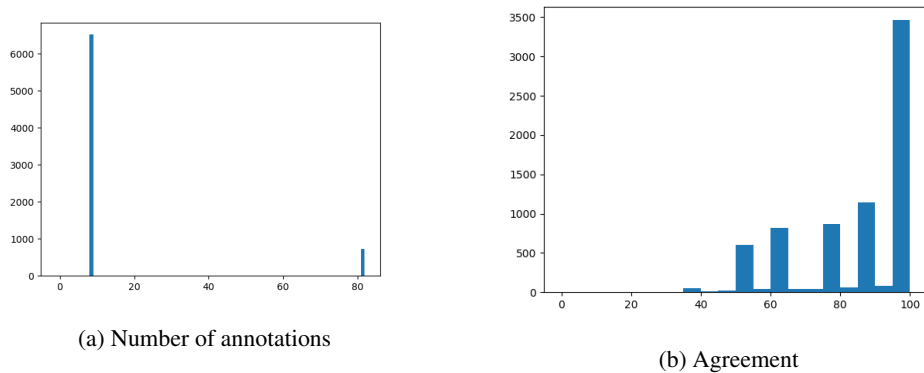


Figure 12: Histograms for the MiceBone Dataset

## A.7 Pig

The raw data was collected at European farms for pigs. The animals were captured via a camera during their arrival and departure. We only received and worked with cropped regions of the tail provided to us by the University of Helsinki to protect data privacy and personal interest of the farms. We created an annotation catalog for the four classes, ‘no injury’, ‘shortened tail but healed’, ‘fresh wound at tail’ and ‘tail not visible’ as instruction. If annotators were uncertain they should decide themselves for the more probable class and in edge cases for the more severe class. We tracked the consistency between the annotators and it stayed at around 50% agreement over all annotations and annotators. The data was annotated in multiple rounds and a subset of 20% was annotated more often. The 6 annotators were domain experts with one exception, a hired worker only for this task which was trained and instructed as the others. This hired worker was paid 11€ per hour and the domain expert did the work as part of their daily job.

### A.7.1 Plankton

We used the preprocessing described in [64] and reproduced their description here with the authors’ permission.

The Plankton dataset was introduced in [61]. The dataset contains 10 plankton classes and has multiple labels per image due to the help of citizen scientists. In contrast to [61], we include ambiguous images in the training and validation set and do not enforce a class balance.

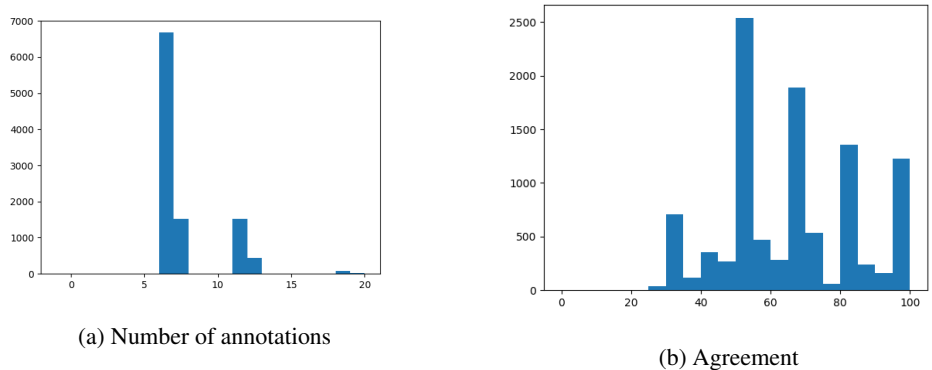


Figure 13: Histograms for the Pig Dataset

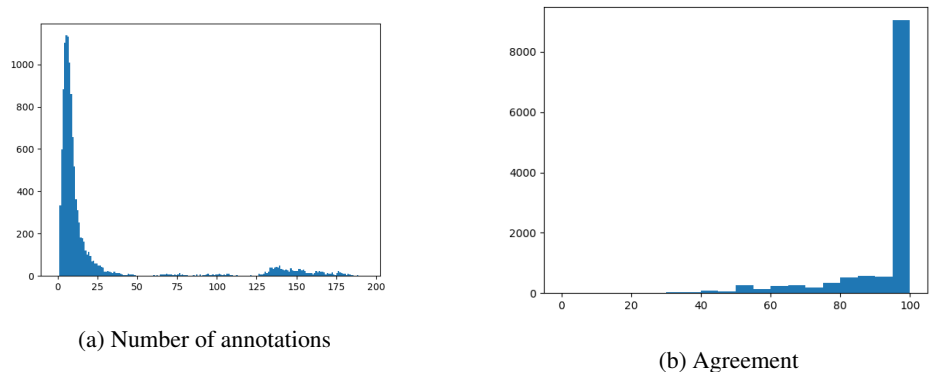


Figure 14: Histograms for the Plankton Dataset

The mentioned citizen scientist used the platform PlanktonID<sup>5</sup>. The total number of annotations and the agreement with the majority vote per image can be seen as a histogram in Figure 14.

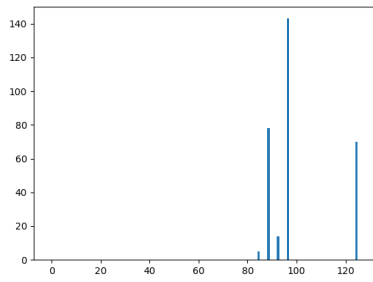
### A.7.2 QualityMRI

The QualityMRI dataset consists of human magnetic resonance images (MRI) with a varying quality and multiple subjective quality ratings gathered in tests with radiologists. It was introduced and evaluated in [51, 69]. We received as raw data 70 images with around 31 annotations and 240 images with around 24 annotations. The annotations had a score from 1 to 5. We simplified these labels into a continuous score between 0 and 1. We converted the given annotations into 4 new annotations for this continuous score. For example, we interpreted an original annotation of 1 as 4 annotations of 0 and an original annotation of 2 as 3 annotations 0 and one annotation 1. The total number of annotations and the agreement with the majority vote per image can be seen as histogram in Figure 15.

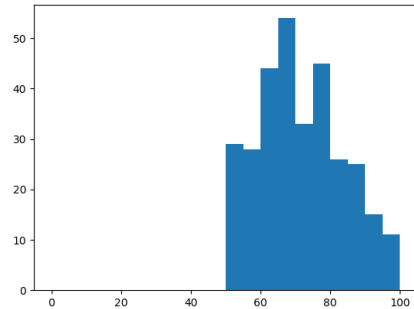
### A.7.3 Synthetic

This dataset consists of images that contain one blue, red, or green circle or ellipse on a black background. We used a hue value of 0 (red), 120 (green) or 240 (blue) and the main axis ratio of 1 (circle) and 2 (ellipse). The combination of all colors and ratios leads to six distinct classes. We generated 1200 images (200 per class) of all classes per fold. The position and scale were determined randomly. Additionally, we generated 1800 images with a random hue and main axis ratio. Based on the random hue and ratio, we determined the class distribution of the six main classes. For example, an image with the hue 60 and an axis ratio of 1.3 would have marginal probabilities of 50% red, 50%

<sup>5</sup><https://planktonid.geomar.de>

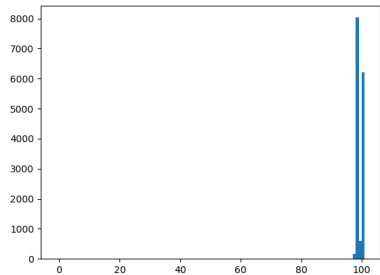


(a) Number of annotations

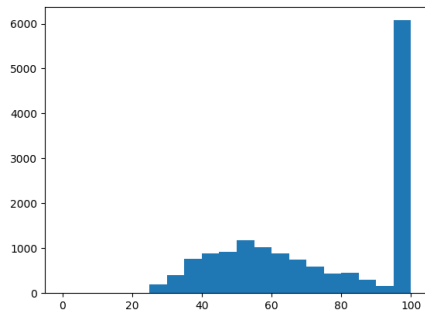


(b) Agreement

Figure 15: Histograms for the QualityMRI Dataset



(a) Number of annotations



(b) Agreement

Figure 16: Histograms for the Synthetic Dataset

green, 30% circle, and 70% ellipse. When we assume an independent combination, we arrive at the following ground truth distribution: red circle 0.15, red ellipse 0.35, blue circle 0.15, blue ellipse 0.35, green circle 0.0, green ellipse 0.0. We generated 100 annotations based on these probability distributions for each image.

The total number of annotations and the agreement with the majority vote per image can be seen as a histogram in Figure 16.

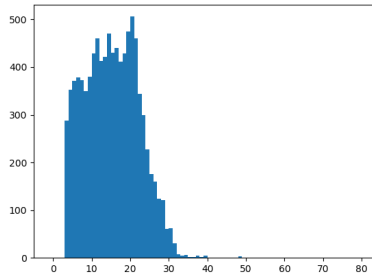
#### A.7.4 Treeversity#1 and Treeversity#6

Both datasets are based on a publicly available crowdsourced dataset of plant images from the Arnold Arboretum of Harvard University<sup>6</sup>. The original task was to tag plant images with up to 21 original classes. The original crowd source page can be found at <https://www.zooniverse.org/projects/friedmaw/treeversity/about/research>. We removed the original photographer and location at the bottom of each image to prevent any information leakage. We used all tags from the original research and interpreted them as annotations. We simplified the original 22 tags to 6 classes as shown below.

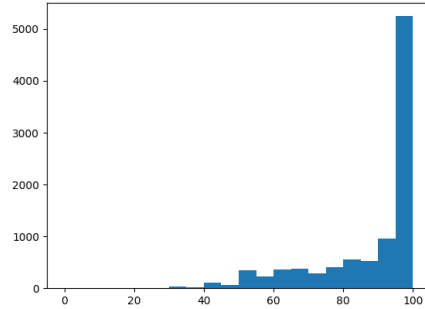
```
'FLWR': 'flower',
'FLOWERINFLORESCENCE': 'flower',
'LEAFNEEDLE': 'leaf',
'LEAF': 'leaf', 'LF': 'leaf',
'BARK': 'bark', 'BRK': 'bark',
'BUDBREAK': 'bud',
```

<sup>6</sup><https://arboretum.harvard.edu/research/data-resources/>





(a) Number of annotations



(b) Agreement

Figure 17: Histograms for the TreeVersity#1 Dataset

```

'BUD': 'bud', 'BD': 'bud',
'FRUIT': 'fruit',
'FRUITSEED': 'fruit', 'FRTSD': 'fruit',
'WHOLEPLANT': 'whole_plant', 'WHLPLNT': 'whole_plant',
'TRUNK': 'bark', 'TRNK': 'bark',
'AUTUMNLEAFCOLOR': 'leaf', 'TMNLFCLR': 'leaf',
'CONE': 'fruit', 'CN': 'fruit',
'LEAFLESS': None, 'LFLSS': None,
'PLNTHLTHSS': None, 'PLANTHEALTHISSUE': None,
'BRNCH': 'bark', 'BRANCH': 'bark',
'NSCTSPDRMT': None, 'INSECTSPIDERMITE': None,
'PRICKLESPINETHORN': None, 'PRCKLSPNTHRN': None
'LCHNMSSFNGS': None, 'LICHENMOSSFUNGUS': None
'POORIMAGEQUALITYNOFEATURES': None
'ROOT': 'bark'

```

We used a simplification due to overlapping classes because of the original tagging and small classes (e.g. health issues with 8 tags). We counted multiple tags which we simplified to the same annotation class only once per image and user and used only images with at least three different annotations. We saw that most people gave only one original tag per image (approx. 54%) and the majority did not give more than three (approx. 98%). Due to the fact that tags are not the same as class annotations, we investigate two subsets. In Treeversity#1, we restrict the annotations based on original tags where exactly one tag was given. In Treeversity#6, we do not use this restriction. The total number of annotations and the agreement with the majority vote per image can be seen as a histogram in Figure 17 and Figure 18.

### A.7.5 Turkey

The turkey dataset with images of turkeys and their injuries [73, 74]. We use three classes head injury, plumage injury, or no injury. All injuries from the neck upwards to the eyes were counted as head injuries and the rest as plumage injuries. We used the original cross-analysis results between domain experts of the original papers. In addition, we collected additional annotations with a similar protocol as for the MiceBone dataset. The main difference is that reference images were not given but the distinction were discussed with all annotators and the domain expert multiple times. The total number of annotations and the agreement with the majority vote per image can be seen as a histogram in Figure 19.

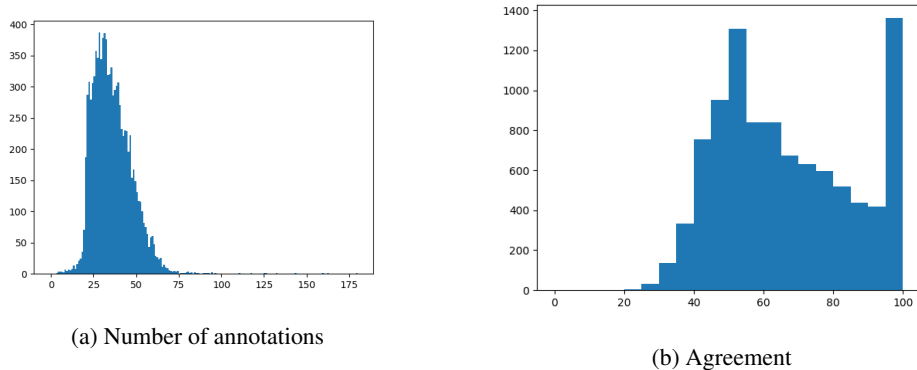


Figure 18: Histograms for the TreeVersity#6 Dataset

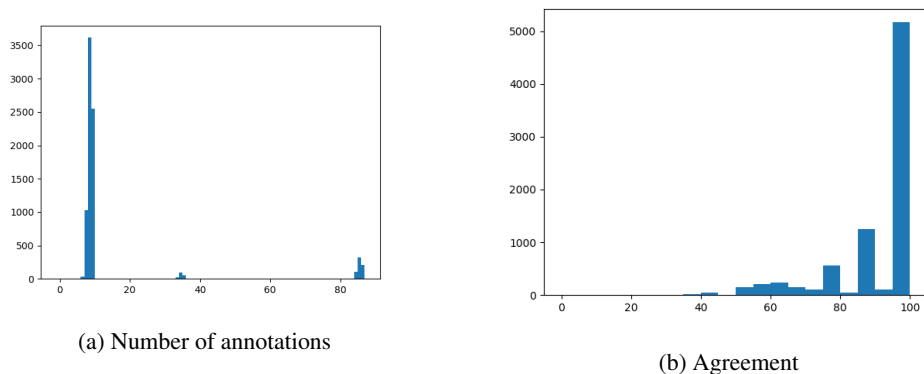


Figure 19: Histograms for the Turkey Dataset

## A.8 Full result tables

We give the full result tables including Median and Mean  $\pm$  SEM for all figures in the main paper. Additionally, we create an overview for  $ACC$  and  $KL$  for a budget of 10%, 100%, and 1000%.

Table 4: Full table to Figure 7a

Budget	10%		20%		50%		100%		300%		500%		1000%	
	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM
$KL$	0.80	0.91 $\pm$ 0.03	0.67	0.83 $\pm$ 0.03	0.53	0.64 $\pm$ 0.02	0.48	0.60 $\pm$ 0.02	0.46	0.55 $\pm$ 0.04	0.42	0.54 $\pm$ 0.03	0.45	0.56 $\pm$ 0.03
$ECE$	0.12	0.15 $\pm$ 0.00	0.10	0.13 $\pm$ 0.00	0.08	0.11 $\pm$ 0.00	0.09	0.11 $\pm$ 0.00	0.08	0.09 $\pm$ 0.01	0.07	0.09 $\pm$ 0.00	0.07	0.09 $\pm$ 0.01
$ACC$	0.55	0.56 $\pm$ 0.01	0.58	0.60 $\pm$ 0.01	0.65	0.66 $\pm$ 0.01	0.69	0.68 $\pm$ 0.01	0.70	0.70 $\pm$ 0.01	0.71	0.71 $\pm$ 0.01	0.72	0.71 $\pm$ 0.01
$F1$	0.52	0.53 $\pm$ 0.01	0.57	0.57 $\pm$ 0.01	0.62	0.63 $\pm$ 0.01	0.65	0.65 $\pm$ 0.01	0.69	0.68 $\pm$ 0.01	0.70	0.70 $\pm$ 0.01	0.70	0.70 $\pm$ 0.01
$\kappa$	0.39	0.39 $\pm$ 0.01	0.47	0.45 $\pm$ 0.01	0.56	0.53 $\pm$ 0.01	0.66	0.65 $\pm$ 0.01	0.68	0.69 $\pm$ 0.01	0.69	0.70 $\pm$ 0.01	0.70	0.71 $\pm$ 0.01

Table 5: Full table to Figure 7b

Budget	100%		300%		500%		1000%	
	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM
Results for <i>KL</i>								
Difficult datasets	0.00	0.00 $\pm$ 0.02	-0.01	-0.03 $\pm$ 0.06	-0.04	-0.04 $\pm$ 0.04	-0.01	-0.06 $\pm$ 0.04
Easy datasets	0.00	-0.00 $\pm$ 0.01	-0.01	-0.06 $\pm$ 0.03	-0.02	-0.08 $\pm$ 0.03	-0.01	-0.02 $\pm$ 0.02
Results for <i>ECE</i>								
Difficult datasets	0.00	0.00 $\pm$ 0.00	-0.02	-0.02 $\pm$ 0.01	-0.02	-0.02 $\pm$ 0.01	-0.01	-0.02 $\pm$ 0.01
Easy datasets	0.00	0.00 $\pm$ 0.00	-0.00	-0.01 $\pm$ 0.01	-0.00	-0.01 $\pm$ 0.00	0.00	-0.00 $\pm$ 0.00
Results for <i>ACC</i>								
Difficult datasets	0.00	-0.00 $\pm$ 0.00	0.02	0.02 $\pm$ 0.01	0.03	0.02 $\pm$ 0.01	0.03	0.04 $\pm$ 0.01
Easy datasets	0.00	0.00 $\pm$ 0.00	0.01	0.01 $\pm$ 0.01	0.01	0.02 $\pm$ 0.00	0.01	0.01 $\pm$ 0.00
Results for <i>F1</i>								
Difficult datasets	0.00	-0.00 $\pm$ 0.00	0.03	0.02 $\pm$ 0.01	0.04	0.04 $\pm$ 0.01	0.05	0.05 $\pm$ 0.01
Easy datasets	0.00	-0.00 $\pm$ 0.00	0.01	0.02 $\pm$ 0.01	0.02	0.03 $\pm$ 0.01	0.01	0.02 $\pm$ 0.00
Results for $\kappa$								
Difficult datasets	0.00	-0.00 $\pm$ 0.00	0.02	0.02 $\pm$ 0.01	0.03	0.04 $\pm$ 0.01	0.04	0.06 $\pm$ 0.01
Easy datasets	0.00	-0.00 $\pm$ 0.00	0.02	0.03 $\pm$ 0.01	0.03	0.04 $\pm$ 0.01	0.02	0.04 $\pm$ 0.01

Table 6: Full table to Figure 7c

Budget	100%		300%		500%		1000%	
	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM
Results for <i>KL</i>								
Using soft labels	0.00	-0.00 $\pm$ 0.02	-0.11	-0.22 $\pm$ 0.05	-0.13	-0.25 $\pm$ 0.04	-0.16	-0.23 $\pm$ 0.04
Using hard labels	0.00	0.00 $\pm$ 0.01	0.01	0.06 $\pm$ 0.05	0.01	0.06 $\pm$ 0.03	0.02	0.07 $\pm$ 0.02
Results for <i>ECE</i>								
Using soft labels	-0.00	-0.00 $\pm$ 0.00	-0.02	-0.04 $\pm$ 0.01	-0.03	-0.04 $\pm$ 0.01	-0.03	-0.04 $\pm$ 0.01
Using hard labels	0.00	0.00 $\pm$ 0.00	-0.00	-0.00 $\pm$ 0.00	-0.00	-0.00 $\pm$ 0.00	0.00	-0.00 $\pm$ 0.00
Results for <i>ACC</i>								
Using soft labels	0.00	0.00 $\pm$ 0.01	0.02	0.03 $\pm$ 0.01	0.02	0.03 $\pm$ 0.01	0.03	0.03 $\pm$ 0.01
Using hard labels	0.00	0.00 $\pm$ 0.00	0.01	0.00 $\pm$ 0.00	0.01	0.02 $\pm$ 0.00	0.01	0.02 $\pm$ 0.00
Results for <i>F1</i>								
Using soft labels	0.00	0.00 $\pm$ 0.01	0.02	0.04 $\pm$ 0.01	0.04	0.04 $\pm$ 0.01	0.03	0.04 $\pm$ 0.01
Using hard labels	0.00	-0.00 $\pm$ 0.00	0.01	0.01 $\pm$ 0.01	0.02	0.03 $\pm$ 0.00	0.02	0.03 $\pm$ 0.01
Results for $\kappa$								
Using soft labels	0.00	0.00 $\pm$ 0.01	0.01	0.03 $\pm$ 0.01	0.03	0.04 $\pm$ 0.01	0.01	0.04 $\pm$ 0.01
Using hard labels	0.00	-0.00 $\pm$ 0.00	0.02	0.02 $\pm$ 0.01	0.03	0.04 $\pm$ 0.01	0.04	0.06 $\pm$ 0.01

Table 7: Full table to Figure 5a

Budget	10%		20%		50%		100%		300%		500%		1000%	
	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM
<i>KL</i>	0.80	0.91 $\pm$ 0.03	0.67	0.83 $\pm$ 0.03	0.53	0.64 $\pm$ 0.02	0.48	0.60 $\pm$ 0.02	0.46	0.55 $\pm$ 0.04	0.42	0.54 $\pm$ 0.03	0.45	0.56 $\pm$ 0.03
<i>KL</i>	1.48	1.70 $\pm$ 0.05	1.21	1.52 $\pm$ 0.05	0.89	1.21 $\pm$ 0.04	0.74	1.08 $\pm$ 0.04	0.72	1.02 $\pm$ 0.06	0.73	0.96 $\pm$ 0.06	0.65	0.95 $\pm$ 0.07
<i>ACC</i>	0.55	0.56 $\pm$ 0.01	0.58	0.60 $\pm$ 0.01	0.65	0.66 $\pm$ 0.01	0.69	0.68 $\pm$ 0.01	0.70	0.70 $\pm$ 0.01	0.71	0.71 $\pm$ 0.01	0.72	0.71 $\pm$ 0.01
<i>ACC</i>	0.54	0.54 $\pm$ 0.01	0.57	0.58 $\pm$ 0.01	0.65	0.63 $\pm$ 0.01	0.67	0.66 $\pm$ 0.01	0.70	0.69 $\pm$ 0.01	0.71	0.70 $\pm$ 0.01	0.72	0.71 $\pm$ 0.01
<i>ECE</i>	0.12	0.15 $\pm$ 0.00	0.10	0.13 $\pm$ 0.00	0.08	0.11 $\pm$ 0.00	0.09	0.11 $\pm$ 0.00	0.08	0.09 $\pm$ 0.01	0.07	0.09 $\pm$ 0.00	0.07	0.09 $\pm$ 0.01
<i>ECE</i>	0.22	0.23 $\pm$ 0.01	0.19	0.21 $\pm$ 0.01	0.15	0.17 $\pm$ 0.01	0.12	0.15 $\pm$ 0.00	0.09	0.13 $\pm$ 0.01	0.08	0.12 $\pm$ 0.01	0.09	0.13 $\pm$ 0.01

Table 8: Full table to Figure 5b

Budget	10-0.10		05-0.20		02-0.50		01-1.00	
	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM
Results for <i>KL</i>								
Pseudo v2 hard	0.84	0.95 $\pm$ 0.12	0.77	0.98 $\pm$ 0.15	0.61	0.67 $\pm$ 0.05	0.44	0.51 $\pm$ 0.05
Pseudo v2 soft	0.45	0.56 $\pm$ 0.07	0.40	0.47 $\pm$ 0.05	0.55	0.52 $\pm$ 0.05	0.43	0.52 $\pm$ 0.05
Results for <i>ACC</i>								
Pseudo v2 hard	0.65	0.63 $\pm$ 0.03	0.64	0.64 $\pm$ 0.03	0.71	0.67 $\pm$ 0.03	0.75	0.71 $\pm$ 0.03
Pseudo v2 soft	0.67	0.64 $\pm$ 0.03	0.66	0.68 $\pm$ 0.03	0.72	0.69 $\pm$ 0.02	0.72	0.71 $\pm$ 0.03
Results for <i>ECE</i>								
Pseudo v2 hard	0.12	0.13 $\pm$ 0.02	0.13	0.14 $\pm$ 0.02	0.10	0.11 $\pm$ 0.02	0.08	0.10 $\pm$ 0.01
Pseudo v2 soft	0.07	0.09 $\pm$ 0.01	0.04	0.07 $\pm$ 0.01	0.07	0.09 $\pm$ 0.01	0.10	0.10 $\pm$ 0.01

Table 9: Full table to Figure 5c

Budget	10%		20%		50%		100%		300%		500%		1000%	
	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM
Results for <i>KL</i>														
Baseline	0.83	1.39 $\pm$ 0.25	0.63	0.98 $\pm$ 0.17	0.52	0.70 $\pm$ 0.08	0.52	0.69 $\pm$ 0.09	0.34	0.37 $\pm$ 0.04	0.37	0.39 $\pm$ 0.04	0.29	0.33 $\pm$ 0.04
Pseudo v2 soft	0.77	0.99 $\pm$ 0.20	0.72	0.90 $\pm$ 0.13	0.61	0.64 $\pm$ 0.06	0.43	0.52 $\pm$ 0.05	0.37	0.39 $\pm$ 0.04	0.34	0.36 $\pm$ 0.03	0.30	0.33 $\pm$ 0.04
DivideMix	0.59	0.60 $\pm$ 0.05	0.53	0.55 $\pm$ 0.04	0.45	0.51 $\pm$ 0.04	0.44	0.52 $\pm$ 0.04	0.47	0.56 $\pm$ 0.06	0.41	0.54 $\pm$ 0.06	0.45	0.64 $\pm$ 0.08
ELR+	0.78	0.80 $\pm$ 0.07	0.57	0.59 $\pm$ 0.06	0.45	0.51 $\pm$ 0.07	0.47	0.52 $\pm$ 0.08	0.48	0.55 $\pm$ 0.09	0.44	0.70 $\pm$ 0.15	0.50	0.57 $\pm$ 0.06
Results for <i>ACC</i>														
Baseline	0.52	0.57 $\pm$ 0.04	0.60	0.63 $\pm$ 0.03	0.65	0.68 $\pm$ 0.03	0.71	0.71 $\pm$ 0.03	0.72	0.74 $\pm$ 0.03	0.75	0.74 $\pm$ 0.03	0.75	0.74 $\pm$ 0.03
Pseudo v2 soft	0.61	0.60 $\pm$ 0.02	0.63	0.63 $\pm$ 0.03	0.69	0.68 $\pm$ 0.03	0.72	0.71 $\pm$ 0.03	0.73	0.74 $\pm$ 0.02	0.77	0.74 $\pm$ 0.03	0.78	0.75 $\pm$ 0.03
DivideMix	0.62	0.62 $\pm$ 0.03	0.68	0.66 $\pm$ 0.03	0.72	0.72 $\pm$ 0.03	0.73	0.73 $\pm$ 0.03	0.74	0.74 $\pm$ 0.03	0.76	0.75 $\pm$ 0.03	0.75	0.74 $\pm$ 0.03
ELR+	0.53	0.53 $\pm$ 0.04	0.59	0.60 $\pm$ 0.04	0.66	0.65 $\pm$ 0.04	0.72	0.68 $\pm$ 0.04	0.73	0.67 $\pm$ 0.04	0.72	0.68 $\pm$ 0.04	0.73	0.69 $\pm$ 0.04
Results for <i>ECE</i>														
Baseline	0.16	0.19 $\pm$ 0.03	0.14	0.17 $\pm$ 0.02	0.12	0.14 $\pm$ 0.02	0.13	0.13 $\pm$ 0.01	0.04	0.07 $\pm$ 0.01	0.05	0.07 $\pm$ 0.01	0.03	0.07 $\pm$ 0.01
Pseudo v2 soft	0.11	0.14 $\pm$ 0.02	0.11	0.14 $\pm$ 0.02	0.11	0.11 $\pm$ 0.01	0.10	0.10 $\pm$ 0.01	0.06	0.08 $\pm$ 0.01	0.06	0.08 $\pm$ 0.01	0.05	0.08 $\pm$ 0.01
DivideMix	0.12	0.12 $\pm$ 0.01	0.09	0.11 $\pm$ 0.01	0.08	0.10 $\pm$ 0.01	0.06	0.09 $\pm$ 0.01	0.06	0.09 $\pm$ 0.01	0.05	0.09 $\pm$ 0.02	0.05	0.09 $\pm$ 0.02
ELR+	0.09	0.15 $\pm$ 0.02	0.07	0.10 $\pm$ 0.02	0.06	0.09 $\pm$ 0.01	0.05	0.09 $\pm$ 0.02	0.05	0.09 $\pm$ 0.02	0.05	0.09 $\pm$ 0.02	0.05	0.09 $\pm$ 0.02

Table 10: Improvements in KL across methods

Budget	10%		100%		1000%	
	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM
Baseline	-0.00	-0.00 $\pm$ 0.07	0.00	-0.00 $\pm$ 0.03	0.00	0.00 $\pm$ 0.01
ELR+	-0.16	-0.59 $\pm$ 0.22	-0.13	-0.17 $\pm$ 0.06	0.15	0.24 $\pm$ 0.06
Het	-0.06	-0.02 $\pm$ 0.08	-0.01	0.01 $\pm$ 0.02	0.22	0.30 $\pm$ 0.05
SGNP	-0.26	-0.59 $\pm$ 0.27	0.00	-0.15 $\pm$ 0.08	0.20	0.26 $\pm$ 0.04
DivideMix	-0.21	-0.78 $\pm$ 0.25	-0.03	-0.17 $\pm$ 0.08	0.16	0.31 $\pm$ 0.07
Fixmatch	-0.22	-0.43 $\pm$ 0.14	-0.05	-0.17 $\pm$ 0.05	N/A	
Fixmatch + DC3	-0.29	-0.48 $\pm$ 0.17	-0.06	-0.13 $\pm$ 0.04	N/A	
Mean	-0.14	-0.59 $\pm$ 0.22	-0.12	-0.22 $\pm$ 0.08	N/A	
Mean+DC3	-0.20	-0.51 $\pm$ 0.17	-0.08	-0.15 $\pm$ 0.07	N/A	
$\pi$	-0.33	-0.63 $\pm$ 0.19	-0.06	-0.17 $\pm$ 0.05	N/A	
$\pi$ +DC3	-0.32	-0.62 $\pm$ 0.20	-0.10	-0.22 $\pm$ 0.06	N/A	
Pseudo v1	-0.25	-0.47 $\pm$ 0.19	-0.07	-0.17 $\pm$ 0.04	N/A	
Pseudo v1 + DC3	-0.24	-0.56 $\pm$ 0.22	-0.07	0.05 $\pm$ 0.23	N/A	
Pseudo v2 hard	-0.30	-0.39 $\pm$ 0.16	-0.03	-0.19 $\pm$ 0.08	0.19	0.28 $\pm$ 0.04
Pseudo v2 soft	-0.28	-0.40 $\pm$ 0.19	-0.04	-0.17 $\pm$ 0.06	0.01	0.00 $\pm$ 0.02
Pseudo v2 not	0.12	-0.11 $\pm$ 0.34	0.10	0.22 $\pm$ 0.19	0.28	0.46 $\pm$ 0.13
BYOL	-0.20	-0.47 $\pm$ 0.21	-0.08	0.02 $\pm$ 0.09	N/A	
MOCOv2	-0.29	-0.63 $\pm$ 0.22	-0.08	-0.13 $\pm$ 0.10	N/A	
SimCLR	-0.07	-0.40 $\pm$ 0.22	0.01	0.04 $\pm$ 0.06	N/A	
SWAV	-0.05	-0.43 $\pm$ 0.22	0.19	0.17 $\pm$ 0.08	N/A	

Table 11: Improvements in ACC across methods

Budget	10%		100%		1000%	
	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM	Median	Mean $\pm$ SEM
Baseline	0.00	0.00 $\pm$ 0.01	0.00	0.00 $\pm$ 0.01	-0.00	-0.00 $\pm$ 0.01
ELR+	-0.02	-0.03 $\pm$ 0.01	0.01	-0.03 $\pm$ 0.03	-0.01	-0.05 $\pm$ 0.02
Het	-0.06	-0.10 $\pm$ 0.02	-0.00	-0.00 $\pm$ 0.01	-0.00	-0.01 $\pm$ 0.01
SGNP	-0.03	-0.05 $\pm$ 0.02	-0.01	-0.01 $\pm$ 0.01	-0.01	-0.01 $\pm$ 0.01
DivideMix	0.05	0.05 $\pm$ 0.02	0.03	0.02 $\pm$ 0.02	0.00	-0.00 $\pm$ 0.01
Fixmatch	0.05	0.05 $\pm$ 0.01	0.01	0.01 $\pm$ 0.01	N/A	
Fixmatch + DC3	0.03	0.02 $\pm$ 0.01	-0.00	-0.00 $\pm$ 0.01	N/A	
Mean	0.02	0.03 $\pm$ 0.02	0.02	-0.01 $\pm$ 0.02	N/A	
Mean+DC3	0.05	0.03 $\pm$ 0.02	0.00	0.00 $\pm$ 0.01	N/A	
$\pi$	0.05	0.04 $\pm$ 0.02	0.02	0.01 $\pm$ 0.01	N/A	
$\pi$ +DC3	0.04	0.03 $\pm$ 0.02	0.02	0.02 $\pm$ 0.01	N/A	
Pseudo v1	0.07	0.03 $\pm$ 0.02	0.01	0.03 $\pm$ 0.01	N/A	
Pseudo v1 + DC3	0.03	0.02 $\pm$ 0.02	0.02	0.03 $\pm$ 0.01	N/A	
Pseudo v2 hard	0.05	0.04 $\pm$ 0.03	0.01	0.00 $\pm$ 0.01	-0.03	-0.02 $\pm$ 0.02
Pseudo v2 soft	0.01	0.04 $\pm$ 0.03	-0.00	-0.00 $\pm$ 0.01	-0.00	0.01 $\pm$ 0.01
Pseudo v2 not	-0.18	-0.18 $\pm$ 0.04	-0.07	-0.14 $\pm$ 0.03	-0.12	-0.17 $\pm$ 0.03
BYOL	-0.05	-0.06 $\pm$ 0.04	-0.11	-0.14 $\pm$ 0.03	N/A	
MOCOv2	-0.04	-0.05 $\pm$ 0.02	-0.10	-0.12 $\pm$ 0.02	N/A	
SimCLR	0.01	0.01 $\pm$ 0.01	-0.04	-0.04 $\pm$ 0.01	N/A	
SWAV	-0.04	-0.06 $\pm$ 0.02	-0.12	-0.13 $\pm$ 0.02	N/A	

Table 12: Details 10% Budget KL

Dataset	Benthic	CIFAR10H	MiceBone	Pig	Plankton	QualityMRI	Synthetic	Treeversity#1	Treeversity#6	Turkey
Baseline	1.17 ± 0.04	0.41 ± 0.02	0.55 ± 0.06	0.75 ± 0.05	0.34 ± 0.02	1.73 ± 0.48	0.08 ± 0.01	0.49 ± 0.04	1.02 ± 0.03	0.40 ± 0.08
ELR+	0.70 ± 0.01	0.29 ± 0.02	0.29 ± 0.01	0.62 ± 0.09	0.24 ± 0.03	1.44 ± 0.83	0.18 ± 0.02	0.46 ± 0.01	0.47 ± 0.05	0.52 ± 0.05
Het	1.16 ± 0.15	0.39 ± 0.01	0.62 ± 0.18	0.70 ± 0.10	0.34 ± 0.03	N/A	0.10 ± 0.01	0.46 ± 0.01	0.99 ± 0.02	0.52 ± 0.07
SGNP	1.11 ± 0.05	0.38 ± 0.01	0.56 ± 0.14	0.77 ± 0.07	0.33 ± 0.02	0.25 ± 0.14	0.10 ± 0.00	0.46 ± 0.01	1.07 ± 0.05	0.42 ± 0.08
DivideMix	0.87 ± 0.07	0.36 ± 0.02	0.38 ± 0.05	0.95 ± 0.04	0.34 ± 0.01	0.46 ± 0.26	0.33 ± 0.00	0.47 ± 0.01	0.62 ± 0.05	0.43 ± 0.08
Fixmatch	0.88 ± 0.24	0.37 ± 0.02	0.41 ± 0.01	0.72 ± 0.05	0.27 ± 0.02	0.99 ± 0.30	0.11 ± 0.02	0.48 ± 0.06	0.51 ± 0.01	0.48 ± 0.05
Fixmatch + DC3	0.82 ± 0.02	0.34 ± 0.03	0.54 ± 0.12	0.71 ± 0.04	0.27 ± 0.02	1.41 ± 0.15	0.18 ± 0.15	0.44 ± 0.01	0.46 ± 0.01	0.47 ± 0.03
Mean	0.80 ± 0.06	0.28 ± 0.02	0.38 ± 0.02	0.64 ± 0.03	0.32 ± 0.01	0.35 ± 0.11	0.09 ± 0.01	0.61 ± 0.01	0.52 ± 0.03	0.75 ± 0.07
Mean+DC3	0.73 ± 0.02	0.28 ± 0.04	0.37 ± 0.00	0.71 ± 0.11	0.30 ± 0.02	1.42 ± 0.93	0.11 ± 0.02	0.51 ± 0.01	0.48 ± 0.02	0.55 ± 0.10
$\pi$	0.71 ± 0.03	0.33 ± 0.02	0.38 ± 0.00	0.83 ± 0.18	0.30 ± 0.02	0.98 ± 0.04	0.08 ± 0.01	0.52 ± 0.01	0.51 ± 0.03	0.55 ± 0.01
$\pi$ +DC3	0.72 ± 0.06	0.30 ± 0.02	0.39 ± 0.04	0.67 ± 0.06	0.29 ± 0.01	0.88 ± 0.34	0.08 ± 0.00	0.48 ± 0.01	0.49 ± 0.00	0.44 ± 0.08
Pseudo v1	1.10 ± 0.13	0.35 ± 0.02	0.33 ± 0.01	0.71 ± 0.07	0.32 ± 0.02	1.25 ± 0.35	0.05 ± 0.00	0.42 ± 0.01	0.42 ± 0.01	0.33 ± 0.03
Pseudo v1 + DC3	1.10 ± 0.12	0.33 ± 0.03	0.34 ± 0.02	2.87 ± 3.09	0.31 ± 0.01	1.22 ± 0.51	0.06 ± 0.00	0.41 ± 0.01	0.43 ± 0.00	0.40 ± 0.04
Pseudo v2 hard	0.97 ± 0.10	0.43 ± 0.02	0.45 ± 0.10	0.67 ± 0.04	0.31 ± 0.03	0.29 ± 0.05	0.10 ± 0.01	0.46 ± 0.03	0.99 ± 0.09	0.39 ± 0.05
Pseudo v2 soft	1.00 ± 0.08	0.41 ± 0.02	0.40 ± 0.08	0.70 ± 0.06	0.32 ± 0.06	0.62 ± 0.16	0.10 ± 0.01	0.46 ± 0.01	0.83 ± 0.13	0.37 ± 0.08
Pseudo v2 not	1.05 ± 0.02	1.12 ± 0.04	0.46 ± 0.10	0.79 ± 0.09	2.70 ± 0.89	0.12 ± 0.02	1.00 ± 0.57	0.68 ± 0.17	0.67 ± 0.11	0.57 ± 0.06
BYOL	0.74 ± 0.03	1.61 ± 0.22	0.46 ± 0.02	0.64 ± 0.01	0.51 ± 0.05	1.07 ± 0.18	0.36 ± 0.01	0.82 ± 0.09	0.58 ± 0.03	0.34 ± 0.04
MOCOv2	0.91 ± 0.05	0.98 ± 0.02	0.37 ± 0.04	0.56 ± 0.01	0.52 ± 0.01	0.29 ± 0.11	0.13 ± 0.01	0.80 ± 0.03	0.61 ± 0.01	0.42 ± 0.08
SimCLR	1.19 ± 0.10	0.36 ± 0.02	0.48 ± 0.05	0.75 ± 0.09	0.29 ± 0.02	1.83 ± 1.01	0.38 ± 0.03	0.68 ± 0.00	0.94 ± 0.04	0.43 ± 0.06
SWAV	1.56 ± 0.16	0.70 ± 0.01	0.56 ± 0.09	1.33 ± 0.98	0.38 ± 0.01	1.90 ± 0.31	0.36 ± 0.03	0.71 ± 0.02	0.61 ± 0.03	0.51 ± 0.04

Table 13: Details 10% Budget Acc

Dataset	Benthic	CIFAR10H	MiceBone	Pig	Plankton	QualityMRI	Synthetic	Treeversity#1	Treeversity#6	Turkey
Baseline	0.64 ± 0.01	0.91 ± 0.00	0.62 ± 0.09	0.36 ± 0.04	0.90 ± 0.01	0.67 ± 0.04	0.88 ± 0.00	0.80 ± 0.02	0.57 ± 0.05	0.76 ± 0.03
ELR+	0.69 ± 0.01	0.89 ± 0.01	0.56 ± 0.05	0.39 ± 0.01	0.92 ± 0.01	0.50 ± 0.00	0.92 ± 0.01	0.80 ± 0.00	0.75 ± 0.02	0.36 ± 0.02
Het	0.63 ± 0.01	0.90 ± 0.01	0.64 ± 0.01	0.36 ± 0.01	0.90 ± 0.01	N/A	0.84 ± 0.02	0.80 ± 0.01	0.60 ± 0.01	0.72 ± 0.01
SGNP	0.64 ± 0.00	0.90 ± 0.01	0.61 ± 0.03	0.33 ± 0.07	0.91 ± 0.01	0.63 ± 0.11	0.84 ± 0.03	0.80 ± 0.01	0.56 ± 0.03	0.75 ± 0.05
DivideMix	0.69 ± 0.04	0.90 ± 0.01	0.66 ± 0.02	0.43 ± 0.02	0.92 ± 0.01	0.52 ± 0.02	0.93 ± 0.00	0.82 ± 0.01	0.77 ± 0.02	0.68 ± 0.02
Fixmatch	0.69 ± 0.04	0.91 ± 0.01	0.72 ± 0.02	0.29 ± 0.05	0.92 ± 0.01	0.72 ± 0.06	0.83 ± 0.03	0.78 ± 0.03	0.64 ± 0.02	0.69 ± 0.06
Fixmatch + DC3	0.69 ± 0.01	0.90 ± 0.00	0.59 ± 0.08	0.33 ± 0.03	0.92 ± 0.01	0.63 ± 0.07	0.81 ± 0.13	0.79 ± 0.01	0.67 ± 0.01	0.76 ± 0.03
Mean	0.69 ± 0.01	0.90 ± 0.01	0.64 ± 0.03	0.39 ± 0.02	0.91 ± 0.01	0.64 ± 0.06	0.90 ± 0.01	0.71 ± 0.01	0.66 ± 0.01	0.52 ± 0.01
Mean+DC3	0.71 ± 0.01	0.90 ± 0.02	0.66 ± 0.00	0.35 ± 0.03	0.91 ± 0.01	0.62 ± 0.07	0.89 ± 0.01	0.75 ± 0.00	0.66 ± 0.01	0.66 ± 0.06
$\pi$	0.71 ± 0.02	0.90 ± 0.00	0.65 ± 0.01	0.34 ± 0.01	0.91 ± 0.01	0.71 ± 0.01	0.92 ± 0.01	0.75 ± 0.01	0.66 ± 0.00	0.63 ± 0.02
$\pi$ +DC3	0.70 ± 0.02	0.90 ± 0.01	0.64 ± 0.05	0.37 ± 0.04	0.92 ± 0.00	0.68 ± 0.06	0.91 ± 0.01	0.77 ± 0.00	0.66 ± 0.00	0.69 ± 0.03
Pseudo v1	0.65 ± 0.05	0.91 ± 0.00	0.71 ± 0.02	0.34 ± 0.08	0.91 ± 0.01	0.68 ± 0.05	0.92 ± 0.01	0.80 ± 0.01	0.69 ± 0.01	0.80 ± 0.04
Pseudo v1 + DC3	0.64 ± 0.04	0.90 ± 0.02	0.73 ± 0.01	0.30 ± 0.04	0.91 ± 0.01	0.75 ± 0.03	0.90 ± 0.02	0.82 ± 0.01	0.69 ± 0.00	0.79 ± 0.04
Pseudo v2 hard	0.66 ± 0.02	0.84 ± 0.01	0.64 ± 0.05	0.35 ± 0.01	0.89 ± 0.02	0.73 ± 0.09	0.87 ± 0.01	0.80 ± 0.01	0.60 ± 0.03	0.78 ± 0.04
Pseudo v2 soft	0.65 ± 0.02	0.85 ± 0.01	0.68 ± 0.01	0.31 ± 0.04	0.88 ± 0.04	0.74 ± 0.06	0.87 ± 0.01	0.79 ± 0.01	0.58 ± 0.04	0.71 ± 0.05
Pseudo v2 not	0.58 ± 0.01	0.59 ± 0.04	0.64 ± 0.04	0.33 ± 0.05	0.39 ± 0.14	0.60 ± 0.06	0.62 ± 0.10	0.69 ± 0.07	0.53 ± 0.10	0.68 ± 0.05
BYOL	0.64 ± 0.01	0.38 ± 0.06	0.54 ± 0.03	0.34 ± 0.02	0.80 ± 0.02	0.50 ± 0.02	0.68 ± 0.00	0.59 ± 0.03	0.59 ± 0.02	0.62 ± 0.04
MOCOv2	0.58 ± 0.01	0.64 ± 0.01	0.52 ± 0.03	0.35 ± 0.01	0.79 ± 0.01	0.56 ± 0.02	0.89 ± 0.02	0.61 ± 0.01	0.58 ± 0.01	0.40 ± 0.05
SimCLR	0.58 ± 0.03	0.87 ± 0.01	0.65 ± 0.05	0.35 ± 0.07	0.88 ± 0.01	0.63 ± 0.11	0.79 ± 0.00	0.69 ± 0.01	0.54 ± 0.00	0.66 ± 0.05
SWAV	0.24 ± 0.02	0.74 ± 0.01	0.53 ± 0.04	0.35 ± 0.04	0.86 ± 0.01	0.60 ± 0.08	0.66 ± 0.03	0.65 ± 0.01	0.60 ± 0.03	0.56 ± 0.10

Table 14: Details 100% Budget Kl

Dataset	Benthic	CIFAR10H	MiceBone	Pig	Plankton	QualityMRI	Synthetic	Treeversity#1	Treeversity#6	Turkey
Baseline	1.17 ± 0.04	0.41 ± 0.02	0.55 ± 0.06	0.75 ± 0.05	0.34 ± 0.02	1.73 ± 0.48	0.08 ± 0.01	0.49 ± 0.04	1.02 ± 0.03	0.40 ± 0.08
ELR+	0.70 ± 0.01	0.29 ± 0.02	0.29 ± 0.01	0.62 ± 0.09	0.24 ± 0.03	1.44 ± 0.83	0.18 ± 0.02	0.46 ± 0.01	0.47 ± 0.05	0.52 ± 0.05
Het	1.16 ± 0.15	0.39 ± 0.01	0.62 ± 0.18	0.70 ± 0.10	0.34 ± 0.03	N/A	0.10 ± 0.01	0.46 ± 0.01	0.99 ± 0.02	0.52 ± 0.07
SGNP	1.11 ± 0.05	0.38 ± 0.01	0.56 ± 0.14	0.77 ± 0.07	0.33 ± 0.02	0.25 ± 0.14	0.10 ± 0.00	0.46 ± 0.01	1.07 ± 0.05	0.42 ± 0.08
DivideMix	0.87 ± 0.07	0.36 ± 0.02	0.38 ± 0.05	0.95 ± 0.04	0.34 ± 0.01	0.46 ± 0.26	0.33 ± 0.00	0.47 ± 0.01	0.62 ± 0.05	0.43 ± 0.08
Fixmatch	0.88 ± 0.24	0.37 ± 0.02	0.41 ± 0.01	0.72 ± 0.05	0.27 ± 0.02	0.99 ± 0.30	0.11 ± 0.02	0.48 ± 0.06	0.51 ± 0.01	0.48 ± 0.05
Fixmatch + DC3	0.82 ± 0.02	0.34 ± 0.03	0.54 ± 0.12	0.71 ± 0.04	0.27 ± 0.02	1.41 ± 0.15	0.18 ± 0.15	0.44 ± 0.01	0.46 ± 0.01	0.47 ± 0.03
Mean	0.80 ± 0.06	0.28 ± 0.02	0.38 ± 0.02	0.64 ± 0.03	0.32 ± 0.01	0.35 ± 0.11	0.09 ± 0.01	0.61 ± 0.01	0.52 ± 0.03	0.75 ± 0.07
Mean+DC3	0.73 ± 0.02	0.28 ± 0.04	0.37 ± 0.00	0.71 ± 0.11	0.30 ± 0.02	1.42 ± 0.93	0.11 ± 0.02	0.51 ± 0.01	0.48 ± 0.02	0.55 ± 0.10
$\pi$	0.71 ± 0.03	0.33 ± 0.02	0.38 ± 0.00	0.83 ± 0.18	0.30 ± 0.02	0.98 ± 0.04	0.08 ± 0.01	0.52 ± 0.01	0.51 ± 0.03	0.55 ± 0.01
$\pi$ +DC3	0.72 ± 0.06	0.30 ± 0.02	0.39 ± 0.04	0.67 ± 0.06	0.29 ± 0.01	0.88 ± 0.34	0.08 ± 0.00	0.48 ± 0.01	0.49 ± 0.00	0.44 ± 0.08
Pseudo v1	1.10 ± 0.13	0.35 ± 0.02	0.33 ± 0.01	0.71 ± 0.07	0.32 ± 0.02	1.25 ± 0.35	0.05 ± 0.00	0.42 ± 0.01	0.42 ± 0.01	0.33 ± 0.03
Pseudo v1 + DC3	1.10 ± 0.12	0.33 ± 0.03	0.34 ± 0.02	2.87 ± 3.09	0.31 ± 0.01	1.22 ± 0.51	0.06 ± 0.00	0.41 ± 0.01	0.43 ± 0.00	0.40 ± 0.04
Pseudo v2 hard	0.97 ± 0.10	0.43 ± 0.02	0.45 ± 0.10	0.67 ± 0.04	0.31 ± 0.03	0.29 ± 0.05	0.10 ± 0.01	0.46 ± 0.03	0.99 ± 0.09	0.39 ± 0.05
Pseudo v2 soft	1.00 ± 0.08	0.41 ± 0.02	0.40 ± 0.08	0.70 ± 0.06	0.32 ± 0.06	0.62 ± 0.16	0.10 ± 0.01	0.46 ± 0.01	0.83 ± 0.13	0.37 ± 0.08
Pseudo v2 not	1.05 ± 0.02	1.12 ± 0.04	0.46 ± 0.10	0.79 ± 0.09	2.70 ± 0.89	0.12 ± 0.02	1.00 ± 0.57	0.68 ± 0.17	0.67 ± 0.11	0.57 ± 0.06
BYOL	0.74 ± 0.03	1.61 ± 0.22	0.46 ± 0.02	0.64 ± 0.01	0.51 ± 0.05	1.07 ± 0.18	0.36 ± 0.01	0.82 ± 0.09	0.58 ± 0.03	0.34 ± 0.04
MOCOv2	0.91 ± 0.05	0.98 ± 0.02	0.37 ± 0.04	0.56 ± 0.01	0.52 ± 0.01	0.29 ± 0.11	0.13 ± 0.01	0.80 ± 0.03	0.61 ± 0.01	0.42 ± 0.08
SimCLR	1.19 ± 0.10	0.36 ± 0.02	0.48 ± 0.05	0.75 ± 0.09	0.29 ± 0.02	1.83 ± 1.01	0.38 ± 0.03	0.68 ± 0.00	0.94 ± 0.04	0.43 ± 0.06
SWAV	1.56 ± 0.16	0.70 ± 0.01	0.56 ± 0.09	1.33 ± 0.98	0.38 ± 0.01	1.90 ± 0.31	0.36 ± 0.03	0.71 ± 0.02	0.61 ± 0.03	0.51 ± 0.04

Table 15: Details 100% Budget Acc

Dataset	Benthic	CIFAR10H	MiceBone	Pig	Plankton	QualityMRI	Synthetic	Treversity#1	Treversity#6	Turkey
Baseline	0.64 ± 0.01	0.91 ± 0.00	0.62 ± 0.09	0.36 ± 0.04	0.90 ± 0.01	0.67 ± 0.04	0.88 ± 0.00	0.80 ± 0.02	0.57 ± 0.05	0.76 ± 0.03
ELR+	0.69 ± 0.01	0.89 ± 0.01	0.56 ± 0.05	0.39 ± 0.01	0.92 ± 0.01	0.50 ± 0.00	0.92 ± 0.01	0.80 ± 0.00	0.75 ± 0.02	0.36 ± 0.02
Het	0.63 ± 0.01	0.90 ± 0.01	0.64 ± 0.01	0.36 ± 0.01	0.90 ± 0.01	N/A	0.84 ± 0.02	0.80 ± 0.01	0.60 ± 0.01	0.72 ± 0.01
SGNP	0.64 ± 0.00	0.90 ± 0.01	0.61 ± 0.03	0.33 ± 0.07	0.91 ± 0.01	0.63 ± 0.11	0.84 ± 0.03	0.80 ± 0.01	0.56 ± 0.03	0.75 ± 0.05
DivideMix	0.69 ± 0.04	0.90 ± 0.01	0.66 ± 0.02	0.43 ± 0.02	0.92 ± 0.01	0.52 ± 0.02	0.93 ± 0.00	0.82 ± 0.01	0.77 ± 0.02	0.68 ± 0.02
Fixmatch	0.69 ± 0.04	0.91 ± 0.01	0.72 ± 0.02	0.29 ± 0.05	0.92 ± 0.01	0.72 ± 0.06	0.83 ± 0.03	0.78 ± 0.03	0.64 ± 0.02	0.69 ± 0.06
Fixmatch + DC3	0.69 ± 0.01	0.90 ± 0.00	0.59 ± 0.08	0.33 ± 0.03	0.92 ± 0.01	0.63 ± 0.07	0.81 ± 0.13	0.79 ± 0.01	0.67 ± 0.01	0.76 ± 0.03
Mean	0.69 ± 0.01	0.90 ± 0.01	0.64 ± 0.03	0.39 ± 0.02	0.91 ± 0.01	0.64 ± 0.06	0.90 ± 0.01	0.71 ± 0.01	0.66 ± 0.01	0.52 ± 0.01
Mean+DC3	0.71 ± 0.01	0.90 ± 0.02	0.66 ± 0.00	0.35 ± 0.03	0.91 ± 0.01	0.62 ± 0.07	0.89 ± 0.01	0.75 ± 0.00	0.66 ± 0.01	0.66 ± 0.06
$\pi$	0.71 ± 0.02	0.90 ± 0.00	0.65 ± 0.01	0.34 ± 0.01	0.91 ± 0.01	0.71 ± 0.01	0.92 ± 0.01	0.75 ± 0.01	0.66 ± 0.00	0.63 ± 0.02
$\pi$ +DC3	0.70 ± 0.02	0.90 ± 0.01	0.64 ± 0.05	0.37 ± 0.04	0.92 ± 0.00	0.68 ± 0.06	0.91 ± 0.01	0.77 ± 0.00	0.66 ± 0.00	0.69 ± 0.03
Pseudo v1	0.65 ± 0.05	0.91 ± 0.00	0.71 ± 0.02	0.34 ± 0.08	0.91 ± 0.01	0.68 ± 0.05	0.92 ± 0.01	0.80 ± 0.01	0.69 ± 0.01	0.80 ± 0.04
Pseudo v1 + DC3	0.64 ± 0.04	0.90 ± 0.02	0.73 ± 0.01	0.30 ± 0.04	0.91 ± 0.01	0.75 ± 0.03	0.90 ± 0.02	0.82 ± 0.01	0.69 ± 0.00	0.79 ± 0.04
Pseudo v2 hard	0.66 ± 0.02	0.84 ± 0.01	0.64 ± 0.05	0.35 ± 0.01	0.89 ± 0.02	0.73 ± 0.09	0.87 ± 0.01	0.80 ± 0.01	0.60 ± 0.03	0.78 ± 0.04
Pseudo v2 soft	0.65 ± 0.02	0.85 ± 0.01	0.68 ± 0.01	0.31 ± 0.04	0.88 ± 0.04	0.74 ± 0.06	0.87 ± 0.01	0.79 ± 0.01	0.58 ± 0.04	0.71 ± 0.05
Pseudo v2 not	0.58 ± 0.01	0.59 ± 0.04	0.64 ± 0.04	0.33 ± 0.05	0.39 ± 0.14	0.60 ± 0.06	0.62 ± 0.10	0.69 ± 0.07	0.53 ± 0.10	0.68 ± 0.05
BYOL	0.64 ± 0.01	0.38 ± 0.06	0.54 ± 0.03	0.34 ± 0.02	0.80 ± 0.02	0.50 ± 0.02	0.68 ± 0.00	0.59 ± 0.03	0.59 ± 0.02	0.62 ± 0.04
MOCov2	0.58 ± 0.01	0.64 ± 0.01	0.52 ± 0.03	0.35 ± 0.01	0.79 ± 0.01	0.56 ± 0.02	0.89 ± 0.02	0.61 ± 0.01	0.58 ± 0.01	0.40 ± 0.05
SimCLR	0.58 ± 0.03	0.87 ± 0.01	0.65 ± 0.05	0.35 ± 0.07	0.88 ± 0.01	0.63 ± 0.11	0.79 ± 0.00	0.69 ± 0.01	0.54 ± 0.00	0.66 ± 0.05
SWAV	0.24 ± 0.02	0.74 ± 0.01	0.53 ± 0.04	0.35 ± 0.04	0.86 ± 0.01	0.60 ± 0.08	0.66 ± 0.03	0.65 ± 0.01	0.60 ± 0.03	0.56 ± 0.10

Table 16: Details 1000% Budget K1

Dataset	Benthic	CIFAR10H	MiceBone	Pig	Plankton	QualityMRI	Synthetic	Treversity#1	Treversity#6	Turkey
Baseline	0.75 ± 0.03	0.26 ± 0.02	0.23 ± 0.02	0.54 ± 0.04	0.20 ± 0.02	0.31 ± 0.14	0.04 ± 0.00	0.36 ± 0.02	0.33 ± 0.01	0.24 ± 0.06
ELR+	0.75 ± 0.05	0.28 ± 0.02	0.34 ± 0.04	0.73 ± 0.02	0.27 ± 0.03	1.26 ± 0.56	0.49 ± 0.06	0.45 ± 0.01	0.64 ± 0.10	0.50 ± 0.13
Het	1.19 ± 0.05	0.38 ± 0.02	0.47 ± 0.03	0.87 ± 0.33	0.36 ± 0.04	N/A	0.26 ± 0.03	0.48 ± 0.03	1.25 ± 0.07	0.39 ± 0.03
SGNP	1.15 ± 0.04	0.36 ± 0.02	0.48 ± 0.13	1.02 ± 0.26	0.35 ± 0.04	0.25 ± 0.18	0.21 ± 0.03	0.49 ± 0.01	1.06 ± 0.04	0.47 ± 0.10
DivideMix	0.83 ± 0.05	0.37 ± 0.03	0.35 ± 0.05	1.04 ± 0.25	0.34 ± 0.01	1.34 ± 0.84	0.43 ± 0.04	0.46 ± 0.01	0.77 ± 0.03	0.41 ± 0.07
Pseudo v2 hard	0.95 ± 0.06	0.42 ± 0.01	0.45 ± 0.09	0.67 ± 0.02	0.43 ± 0.17	1.02 ± 0.27	0.20 ± 0.09	0.48 ± 0.04	1.00 ± 0.05	0.42 ± 0.10
Pseudo v2 soft	0.70 ± 0.02	0.43 ± 0.02	0.22 ± 0.02	0.61 ± 0.07	0.25 ± 0.01	1.12 ± 0.04	0.06 ± 0.01	0.37 ± 0.02	0.35 ± 0.01	0.21 ± 0.02
Pseudo v2 not	1.04 ± 0.13	1.04 ± 0.03	0.31 ± 0.05	0.57 ± 0.01	2.42 ± 0.98	0.16 ± 0.03	0.72 ± 0.32	0.66 ± 0.07	0.60 ± 0.09	0.35 ± 0.04

Table 17: Details 1000% Budget Acc

Dataset	Benthic	CIFAR10H	MiceBone	Pig	Plankton	QualityMRI	Synthetic	Treversity#1	Treversity#6	Turkey
Baseline	0.68 ± 0.01	0.91 ± 0.01	0.69 ± 0.07	0.42 ± 0.04	0.92 ± 0.01	0.56 ± 0.01	0.91 ± 0.01	0.82 ± 0.01	0.76 ± 0.01	0.75 ± 0.07
ELR+	0.68 ± 0.02	0.90 ± 0.01	0.64 ± 0.05	0.38 ± 0.00	0.92 ± 0.01	0.50 ± 0.00	0.93 ± 0.02	0.81 ± 0.01	0.79 ± 0.01	0.38 ± 0.04
Het	0.66 ± 0.02	0.91 ± 0.01	0.68 ± 0.02	0.39 ± 0.04	0.91 ± 0.01	N/A	0.91 ± 0.01	0.82 ± 0.01	0.68 ± 0.02	0.79 ± 0.05
SGNP	0.66 ± 0.01	0.91 ± 0.01	0.68 ± 0.02	0.36 ± 0.04	0.90 ± 0.02	0.68 ± 0.08	0.92 ± 0.01	0.81 ± 0.01	0.70 ± 0.01	0.74 ± 0.06
DivideMix	0.72 ± 0.01	0.90 ± 0.01	0.68 ± 0.04	0.40 ± 0.02	0.92 ± 0.01	0.57 ± 0.08	0.94 ± 0.00	0.82 ± 0.01	0.78 ± 0.01	0.69 ± 0.01
Pseudo v2 hard	0.67 ± 0.01	0.85 ± 0.01	0.68 ± 0.07	0.34 ± 0.06	0.85 ± 0.08	0.73 ± 0.08	0.87 ± 0.04	0.80 ± 0.01	0.66 ± 0.05	0.74 ± 0.05
Pseudo v2 soft	0.68 ± 0.01	0.84 ± 0.01	0.73 ± 0.03	0.39 ± 0.08	0.90 ± 0.01	0.70 ± 0.05	0.91 ± 0.01	0.81 ± 0.01	0.76 ± 0.01	0.79 ± 0.07
Pseudo v2 not	0.57 ± 0.05	0.62 ± 0.02	0.64 ± 0.08	0.40 ± 0.04	0.35 ± 0.13	0.56 ± 0.05	0.67 ± 0.09	0.68 ± 0.05	0.57 ± 0.09	0.67 ± 0.01

## A.9 Raw Data

In the following, we provide the raw data aggregated across the three slices as mean  $\pm$  standard deviation (STD). The content is best viewed digitally. Some results are missing (N/A) due to hardware restrictions, degenerated training, or not applicable methods (e.g., the method Het does not work with only two classes as in QualityMRI).







Method	m	n	F1	ACC	KL	ECE	F1	ACC	KL	ECE	$\kappa$	b	$\hat{b}$
BYOL	01	01	0.5907 + 0.0035	0.5600 + 0.0477	0.5300 + 0.0460	0.1084 + 0.0178	0.5154 + 0.0629	0.5684 + 0.0662	1.0798 + 0.2157	0.2017 + 0.0318	0.4100 + 0.0375	0.0999 + 0.0000	0.0999 + 0.0000
BYOL	01	020	0.5085 + 0.0245	0.5536 + 0.0260	0.5356 + 0.0288	0.1300 + 0.0269	0.5176 + 0.0325	0.5689 + 0.0197	1.0211 + 0.0441	0.2048 + 0.0276	0.4847 + 0.0596	0.2000 + 0.0000	0.2000 + 0.0000
BYOL	01	050	0.5871 + 0.0271	0.6071 + 0.0219	0.4270 + 0.0290	0.0920 + 0.0213	0.5418 + 0.0218	0.5905 + 0.0105	0.8816 + 0.0470	0.1613 + 0.0232	0.5125 + 0.0473	0.4999 + 0.0000	0.4999 + 0.0000
BYOL	01	100	0.5422 + 0.0175	0.5444 + 0.0256	0.4568 + 0.0241	0.0802 + 0.0179	0.5284 + 0.0170	0.5647 + 0.0259	1.0331 + 0.0772	0.1688 + 0.0112	0.5552 + 0.0664	1.0000 + 0.0000	1.0000 + 0.0000
DivideMix	01	010	0.5085 + 0.0956	0.4871 + 0.0991	0.4514 + 0.0687	0.1050 + 0.0661	0.5195 + 0.0671	0.5389 + 0.0752	1.0036 + 0.0859	0.0636 + 0.0416	0.3707 + 0.1178	0.0999 + 0.0000	0.0999 + 0.0000
DivideMix	01	020	0.5935 + 0.0336	0.6296 + 0.0346	0.4212 + 0.0605	0.0468 + 0.0055	0.6408 + 0.0320	0.6808 + 0.0402	1.0088 + 0.0342	0.0320 + 0.0041	0.6533 + 0.0391	0.2000 + 0.0000	0.2000 + 0.0000
DivideMix	01	050	0.6029 + 0.0276	0.6265 + 0.0323	0.5782 + 0.0448	0.0820 + 0.0348	0.6328 + 0.0402	0.6697 + 0.0546	1.0429 + 0.0503	0.0390 + 0.0357	0.6752 + 0.0522	0.4999 + 0.0000	0.4999 + 0.0000
DivideMix	01	100	0.6399 + 0.0136	0.6599 + 0.0217	0.3800 + 0.0482	0.0658 + 0.0269	0.6664 + 0.0176	0.6923 + 0.0308	1.0372 + 0.0599	0.0524 + 0.0106	0.7217 + 0.0324	1.0000 + 0.0000	1.0000 + 0.0000
DivideMix	02	050	0.6374 + 0.0356	0.6625 + 0.0294	0.5073 + 0.0141	0.0967 + 0.0256	0.6897 + 0.0307	0.7123 + 0.0246	1.0611 + 0.0286	0.1158 + 0.0276	0.7170 + 0.0424	0.9998 + 0.0000	0.9998 + 0.0000
DivideMix	03	010	0.6878 + 0.0128	0.6999 + 0.0165	0.3035 + 0.0438	0.0431 + 0.0079	0.6883 + 0.0273	0.7146 + 0.0229	1.0428 + 0.0259	0.0319 + 0.0086	0.7613 + 0.0316	3.0000 + 0.0000	3.0000 + 0.0000
DivideMix	03	020	0.5773 + 0.0403	0.5919 + 0.0406	0.4718 + 0.0164	0.1085 + 0.0214	0.6097 + 0.0470	0.6527 + 0.0470	1.0605 + 0.0425	0.1130 + 0.0123	0.6610 + 0.0496	0.9998 + 0.0001	0.9998 + 0.0001
DivideMix	03	050	0.6930 + 0.0039	0.7174 + 0.0099	0.3352 + 0.0235	0.0441 + 0.0131	0.6924 + 0.0120	0.7088 + 0.0091	1.0134 + 0.0154	0.0244 + 0.0034	0.7585 + 0.0232	5.0000 + 0.0000	5.0000 + 0.0000
DivideMix	03	100	0.6024 + 0.0353	0.6103 + 0.0405	0.4647 + 0.1154	0.0851 + 0.0404	0.6313 + 0.0321	0.6532 + 0.0510	1.0565 + 0.1302	0.1131 + 0.0310	0.6209 + 0.0614	0.9993 + 0.0004	0.9993 + 0.0004
DivideMix	03	100	0.6618 + 0.0309	0.6841 + 0.0410	0.3532 + 0.0484	0.0412 + 0.0108	0.6999 + 0.0102	0.7195 + 0.0130	1.0350 + 0.0117	0.0313 + 0.0059	0.7576 + 0.0383	10.0000 + 0.0000	10.0000 + 0.0000
DivideMix	03	100	0.2765 + 0.0026	0.3333 + 0.0000	0.9533 + 0.1075	0.2603 + 0.0159	0.2765 + 0.0026	0.3333 + 0.0000	0.9666 + 0.1082	0.2576 + 0.0198	N/A	0.9999 + 0.0000	0.9999 + 0.0000
DivideMix	03	100	0.4443 + 0.0147	0.4628 + 0.0306	0.3270 + 0.0251	0.0585 + 0.0256	0.4830 + 0.0301	0.5077 + 0.0658	1.0320 + 0.0127	0.0447 + 0.0182	0.5424 + 0.1199	0.2000 + 0.0000	0.2000 + 0.0000
DivideMix	03	100	0.4790 + 0.0158	0.5140 + 0.0247	0.3315 + 0.0128	0.0718 + 0.0046	0.5347 + 0.0081	0.5751 + 0.0458	1.0394 + 0.0059	0.0597 + 0.0153	0.6611 + 0.0302	0.4999 + 0.0000	0.4999 + 0.0000
DivideMix	03	100	0.5775 + 0.0456	0.5647 + 0.0512	0.2945 + 0.0150	0.0475 + 0.0071	0.6060 + 0.0263	0.6110 + 0.0227	1.0420 + 0.0358	0.0563 + 0.0326	0.6586 + 0.0708	1.0000 + 0.0000	1.0000 + 0.0000
DivideMix	03	100	0.4795 + 0.0095	0.4809 + 0.1037	0.3456 + 0.0719	0.0572 + 0.1152	0.5307 + 0.1280	0.5307 + 0.1280	0.5306 + 0.0888	0.0946 + 0.0126	0.4608 + 0.0229	0.9998 + 0.0000	0.9998 + 0.0000
DivideMix	03	100	0.5466 + 0.0547	0.5405 + 0.0356	0.3320 + 0.0356	0.0599 + 0.0137	0.5773 + 0.0179	0.5942 + 0.0759	1.0345 + 0.0051	0.0666 + 0.0258	0.6780 + 0.0606	3.0000 + 0.0000	3.0000 + 0.0000
DivideMix	03	100	0.4822 + 0.0382	0.4817 + 0.0503	0.3390 + 0.0350	0.0838 + 0.0194	0.5231 + 0.0152	0.5486 + 0.0637	1.0526 + 0.0360	0.0744 + 0.0216	0.6449 + 0.0490	0.9998 + 0.0001	0.9998 + 0.0001
DivideMix	03	100	0.5661 + 0.0128	0.5554 + 0.0184	0.6394 + 0.0378	0.1610 + 0.0324	0.5134 + 0.0200	0.5596 + 0.0159	1.3845 + 0.0732	0.2463 + 0.0096	0.4290 + 0.0300	0.9999 + 0.0000	0.9999 + 0.0000
DivideMix	03	100	0.2765 + 0.0026	0.3333 + 0.0000	0.9630 + 0.0397	0.2639 + 0.0117	0.2765 + 0.0026	0.3333 + 0.0000	0.9661 + 0.0426	0.2615 + 0.0115	N/A	0.9993 + 0.0004	0.9993 + 0.0004
DivideMix	03	100	0.6661 + 0.0368	0.6400 + 0.0461	0.3386 + 0.0408	0.0485 + 0.0170	0.6922 + 0.0110	0.6813 + 0.0229	0.7379 + 0.0114	0.0766 + 0.0065	0.7694 + 0.0527	10.0000 + 0.0000	10.0000 + 0.0000
DivideMix	03	100	0.5033 + 0.0128	0.5554 + 0.0184	0.6394 + 0.0378	0.1610 + 0.0324	0.5134 + 0.0200	0.5596 + 0.0159	1.3845 + 0.0732	0.2463 + 0.0096	0.4290 + 0.0300	0.9999 + 0.0000	0.9999 + 0.0000
Fixmatch	01	020	0.5725 + 0.0371	0.6186 + 0.0227	0.4433 + 0.0692	0.1030 + 0.0294	0.5538 + 0.0109	0.6099 + 0.0190	0.9751 + 0.1206	0.2076 + 0.0095	0.4792 + 0.0305	0.2000 + 0.0000	0.2000 + 0.0000
Fixmatch	01	050	0.5817 + 0.0275	0.6071 + 0.0418	0.4223 + 0.0332	0.0702 + 0.0171	0.5433 + 0.0357	0.6455 + 0.0446	0.9925 + 0.0540	0.1596 + 0.0448	0.6577 + 0.0710	0.4999 + 0.0000	0.4999 + 0.0000
Fixmatch	01	100	0.6107 + 0.0195	0.6184 + 0.0223	0.4889 + 0.0194	0.0872 + 0.0140	0.6195 + 0.0129	0.6501 + 0.0148	1.0299 + 0.0129	0.0694 + 0.0094	0.6994 + 0.0181	1.0000 + 0.0000	1.0000 + 0.0000
Fixmatch + S2C2	01	01	0.5231 + 0.0224	0.5500 + 0.0253	0.5661 + 0.0430	0.1354 + 0.0301	0.5263 + 0.0380	0.5545 + 0.0268	1.2617 + 0.0712	0.2284 + 0.0127	0.4020 + 0.0285	0.0999 + 0.0000	0.0999 + 0.0000
Fixmatch + S2C2	01	020	0.5635 + 0.0250	0.6186 + 0.0365	0.4949 + 0.0827	0.0690 + 0.0236	0.5310 + 0.0296	0.6118 + 0.0143	0.8289 + 0.1531	0.2099 + 0.0024	0.5575 + 0.0495	0.2000 + 0.0000	0.2000 + 0.0000
Fixmatch + S2C2	01	050	0.6194 + 0.0152	0.6375 + 0.0189	0.4718 + 0.0249	0.0856 + 0.0129	0.6143 + 0.0152	0.6582 + 0.0139	1.0113 + 0.0113	0.0665 + 0.0059	0.7000 + 0.0000	1.0000 + 0.0000	1.0000 + 0.0000
Fixmatch + S2C2	01	100	0.4917 + 0.0820	0.5911 + 0.0765	0.5406 + 0.1225	0.1025 + 0.0203	0.4811 + 0.0200	0.5119 + 0.0000	0.7114 + 0.0000	0.1133 + 0.0000	0.4680 + 0.1847	1.0000 + 0.0000	1.0000 + 0.0000
Hot	01	01	0.5120 + 0.0535	0.5245 + 0.0724	0.5404 + 0.0576	0.1166 + 0.0211	0.4820 + 0.0505	0.4942 + 0.0491	2.0372 + 0.1984	0.2489 + 0.0211	0.3190 + 0.0322	0.0999 + 0.0000	0.0999 + 0.0000
Hot	01	020	0.5841 + 0.0632	0.5943 + 0.0757	0.5424 + 0.0245	0.0472 + 0.0195	0.5323 + 0.0240	0.5323 + 0.0240	1.0332 + 0.0195	0.1688 + 0.0112	0.5552 + 0.0664	1.0000 + 0.0000	1.0000 + 0.0000
Hot	01	050	0.5824 + 0.0119	0.5869 + 0.0261	0.4576 + 0.0212	0.0824 + 0.0057	0.5068 + 0.0333	0.5205 + 0.0135	2.2093 + 0.1483	0.2558 + 0.0125	0.3943 + 0.0354	0.4999 + 0.0000	0.4999 + 0.0000
Hot	01	100	0.6100 + 0.0430	0.6399 + 0.0096	0.6154 + 0.1843	0.1325 + 0.0458	0.5066 + 0.0247	0.5152 + 0.0268	2.3437 + 0.1291	0.2440 + 0.0022	0.5300 + 0.0837	1.0000 + 0.0000	1.0000 + 0.0000
Hot	02	050	0.5154 + 0.0549	0.5532 + 0.0379	0.6307 + 0.1192	0.1498 + 0.0436	0.5004 + 0.0395	0.4917 + 0.0336	2.4035 + 0.1994	0.2505 + 0.0138	0.3807 + 0.0598	0.9998 + 0.0000	0.9998 + 0.0000
Hot	02	100	0.6435 + 0.0057	0.6434 + 0.0064	0.6434 + 0.0064	0.6434 + 0.0064	0.6434 + 0.0064	0.6434 + 0.0064	2.4035 + 0.1994	0.2505 + 0.0138	0.3807 + 0.0598	0.9998 + 0.0000	0.9998 + 0.0000
Hot	02	100	0.5245 + 0.0245	0.5181 + 0.0707	0.4949 + 0.0366	0.1120 + 0.0403	0.5212 + 0.0450	0.5199 + 0.0469	2.0447 + 0.2054	0.2277 + 0.0111	0.3088 + 0.0239	0.9998 + 0.0001	0.9998 + 0.0001
Hot	05	100	0.6774 + 0.0160	0.6852 + 0.0179	0.5154 + 0.0719	0.1011 + 0.0119	0.5417 + 0.0470	0.5531 + 0.0419	2.2718 + 0.2064	0.2253 + 0.0160	0.5877 + 0.0784	5.0000 + 0.0000	5.0000 + 0.0000
Hot	10	01	0.5406 + 0.0232	0.5406 + 0.0232	0.5406 + 0.0232	0.5406 + 0.0232	0.5406 + 0.0232	0.5406 + 0.0232	2.4035 + 0.1994	0.2505 + 0.0138	0.3807 + 0.0598	0.9998 + 0.0000	0.9998 + 0.0000
Hot	10	01	0.6773 + 0.0325	0.6834 + 0.0221	0.4741 + 0.0269	0.0901 + 0.0034	0.5233 + 0.0482	0.5421 + 0.0411	2.2400 + 0.2797	0.2308 + 0.0163	0.5928 + 0.0891	10.0000 + 0.0000	10.0000 + 0.0000
Hot	10	01	0.5045 + 0.0380	0.5650 + 0.0436	0.5479 + 0.1371	0.1246 + 0.0593	0.5162 + 0.0386	0.5799 + 0.0608	0.6365 + 0.1354	0.1488 + 0.0511	0.4596 + 0.1317	0.0999 + 0.0000	0.0999 + 0.0000
Hot	10	01	0.5841 + 0.0632	0.5943 + 0.0757	0.5424 + 0.0245	0.0472 + 0.0195	0.5323 + 0.0240	0.5323 + 0.0240	1.0332 + 0.0195	0.1688 + 0.0112	0.5552 + 0.0664	1.0000 + 0.0000	1.0000 + 0.0000
Hot	10	050	0.6126 + 0.0010	0.6712 + 0.0244	0.7271 + 0.0431	0.0708 + 0.0244	0.5985 + 0.0209	0.6690 + 0.0602	4.081 + 0.0414	0.0634 + 0.0156	0.6434 + 0.0233	0.4999 + 0.0000	0.4999 + 0.0000
Hot	10	100	0.5800 + 0.0353	0.6397 + 0.0327	0.3826 + 0.0164	0.0655 + 0.0113	0.5907 + 0.0170	0.6682 + 0.0064	4.0002 + 0.0322	0.0222 + 0.0054	0.6442 + 0.0452	1.0000 + 0.0000	1.0000 + 0.0000
Hot	10	100	0.6575 + 0.0590	0.5945 + 0.0641	0.4578 + 0.1151	0.0815 + 0.0420	0.5865 + 0.0349	0.6271 + 0.0436	5.4320 + 0.0760	0.1094 + 0.0259	0.6292 + 0.0644	0.0999 + 0.0000	0.0999 + 0.0000
Hot	10	100	0.6075 + 0.0152	0.6389 + 0.0094	0.5406 + 0.0232	0.0856 + 0.0129	0.6143 + 0.0152	0.6582 + 0.0139	1.0113 + 0.0113	0.0665 + 0.0059	0.7000 + 0.0000	1.0000 + 0.0000	1.0000 + 0.0000
Hot	10	100	0.5977 + 0.0413	0.6624 + 0.0371	0.3789 + 0.0259	0.0583 + 0.0285	0.6105 + 0.0204	0.6832 + 0.0055	4.4387 + 0.0314	0.0774 + 0.0227	0.6225 + 0.0019	0.4999 + 0.0000	0.4999 + 0.0000
Hot	10	100	0.5985 + 0.0889	0.6648 + 0.0306	0.3738 + 0.0259	0.0488 + 0.0190	0.5968 + 0.0147	0.6821 + 0.0049	4.083 + 0.0176	0.0460 + 0.0222	0.7140 + 0.0713	1.0000 + 0.0000	1.0000 + 0.0000
Hot	10	100	0.6735 + 0.0189	0.6943 + 0.0245	0.4223 + 0.0173	0.0560 + 0.0132	0.4530 + 0.0250	0.4618 + 0.0173	4.5272 + 0.0246	0.0495 + 0.0162	0.7387 + 0.0959	0.9999 + 0.0000	0.9999 + 0.0000
Hot	10	100	0.5208 + 0.0441	0.5407 + 0.0393	0.4808 + 0.0715	0.0498 + 0.0298	0.4976 + 0.0217	0.5033 + 0.0191	4.0704 + 0.0619	0.0501 + 0.0122	0.4031 + 0.0630		



Method	<i>m</i>	<i>n</i>	<i>F1</i>	<i>ACC</i>	<i>KL</i>	<i>ECE</i>	$\hat{F}_1$	<i>ACC</i>	<i>KL</i>	<i>ECE</i>	$\kappa$	<i>b</i>	$\hat{b}$	
BYOL	01	0.10	0.6635 ± 0.0649	0.6697 ± 0.0671	0.7899 ± 0.1436	0.0330 ± 0.0022	0.6072 ± 0.0638	0.6012 ± 0.0662	1.0759 ± 0.1352	0.1051 ± 0.0221	0.5512 ± 0.0128	0.1000 ± 0.0000	0.1000 ± 0.0000	
BYOL	02	0.20	0.6295 ± 0.0493	0.7035 ± 0.0509	0.7056 ± 0.1014	0.0215 ± 0.0027	0.6482 ± 0.0552	0.6477 ± 0.0588	0.9227 ± 0.1221	0.0729 ± 0.0201	0.6446 ± 0.0242	0.2000 ± 0.0000	0.2000 ± 0.0000	
BYOL	03	0.50	0.7254 ± 0.0511	0.7371 ± 0.0519	0.6367 ± 0.1294	0.0376 ± 0.0168	0.6605 ± 0.0574	0.6557 ± 0.0581	0.8635 ± 0.1673	0.0385 ± 0.0229	0.6299 ± 0.0250	0.5000 ± 0.0000	0.5000 ± 0.0000	
BYOL	04	1.00	0.7869 ± 0.0224	0.8002 ± 0.0215	0.6510 ± 0.0527	0.0241 ± 0.0044	0.7165 ± 0.0299	0.7176 ± 0.0281	0.6772 ± 0.0350	0.1791 ± 0.0149	0.7191 ± 0.0149	1.0000 ± 0.0000	1.0000 ± 0.0000	
DivideMix	01	0.10	0.7192 ± 0.0098	0.7653 ± 0.1316	0.6350 ± 0.0417	0.1127 ± 0.1515	0.7579 ± 0.0173	0.3050 ± 0.0272	0.0609 ± 0.0099	0.8223 ± 0.0075	0.1000 ± 0.0000	0.1000 ± 0.0000	0.1000 ± 0.0000	
DivideMix	02	0.20	0.8235 ± 0.0385	0.8538 ± 0.0363	0.4832 ± 0.0292	0.0940 ± 0.0202	0.8207 ± 0.0339	0.8446 ± 0.0293	0.2156 ± 0.0124	0.0367 ± 0.0055	0.8349 ± 0.0196	0.2000 ± 0.0000	0.2000 ± 0.0000	
DivideMix	03	0.50	0.8936 ± 0.0112	0.9141 ± 0.0095	0.5470 ± 0.0214	0.0686 ± 0.0107	0.8987 ± 0.0096	0.9115 ± 0.0086	0.1201 ± 0.0111	0.0157 ± 0.0035	0.9158 ± 0.0052	0.5000 ± 0.0000	0.5000 ± 0.0000	
DivideMix	04	1.00	0.9016 ± 0.0068	0.9183 ± 0.0066	0.3371 ± 0.0118	0.0558 ± 0.0106	0.9107 ± 0.0047	0.9221 ± 0.0061	0.1015 ± 0.0025	0.0105 ± 0.0016	0.9464 ± 0.0056	1.0000 ± 0.0000	1.0000 ± 0.0000	
DivideMix	05	0.20	0.8836 ± 0.0119	0.9096 ± 0.0097	0.4384 ± 0.0295	0.0485 ± 0.0056	0.8940 ± 0.0036	0.9119 ± 0.0036	0.5785 ± 0.0322	0.0620 ± 0.0013	0.9164 ± 0.0026	0.9999 ± 0.0000	0.9999 ± 0.0000	
DivideMix	06	1.00	0.9013 ± 0.0062	0.9196 ± 0.0052	0.3419 ± 0.0058	0.0554 ± 0.0039	0.9194 ± 0.0042	0.9301 ± 0.0052	0.1064 ± 0.0116	0.0088 ± 0.0030	0.9539 ± 0.0035	3.0000 ± 0.0000	3.0000 ± 0.0000	
DivideMix	07	0.20	0.8599 ± 0.0060	0.8843 ± 0.0105	0.4762 ± 0.0156	0.0488 ± 0.0130	0.8513 ± 0.0093	0.8721 ± 0.0136	0.6793 ± 0.0132	0.0805 ± 0.0060	0.8673 ± 0.0103	0.9999 ± 0.0001	0.9999 ± 0.0001	
DivideMix	08	1.00	0.9022 ± 0.0008	0.9210 ± 0.0055	0.3367 ± 0.0071	0.0463 ± 0.0029	0.9187 ± 0.0029	0.9294 ± 0.0030	0.1071 ± 0.0041	0.0088 ± 0.0005	0.9526 ± 0.0034	5.0000 ± 0.0000	5.0000 ± 0.0000	
DivideMix	09	1.00	0.8881 ± 0.0061	0.9146 ± 0.0365	0.8462 ± 0.2578	0.1076 ± 0.0506	0.6854 ± 0.0260	0.7394 ± 0.0394	1.0537 ± 0.2942	0.1379 ± 0.0507	0.7598 ± 0.0503	0.9999 ± 0.0001	0.9999 ± 0.0001	
DivideMix	10	1.00	0.8986 ± 0.0036	0.9179 ± 0.0051	0.3445 ± 0.0113	0.0473 ± 0.0119	0.9185 ± 0.0012	0.9292 ± 0.0015	0.1157 ± 0.0128	0.0091 ± 0.0019	0.9549 ± 0.0048	10.0000 ± 0.0000	10.0000 ± 0.0000	
DivideMix	11	0.10	0.8036 ± 0.0249	0.7941 ± 0.0349	0.4480 ± 0.0354	0.0596 ± 0.0147	0.7680 ± 0.0247	0.7472 ± 0.0276	0.5243 ± 0.0324	0.0209 ± 0.0022	0.7902 ± 0.0332	0.1000 ± 0.0000	0.1000 ± 0.0000	
DivideMix	12	0.20	0.8657 ± 0.0068	0.8638 ± 0.0115	0.3156 ± 0.0149	0.0318 ± 0.0045	0.8360 ± 0.0148	0.8232 ± 0.0162	0.3881 ± 0.0259	0.0174 ± 0.0031	0.8607 ± 0.0054	0.2000 ± 0.0000	0.2000 ± 0.0000	
DivideMix	13	0.50	0.8838 ± 0.0109	0.9035 ± 0.0100	0.2523 ± 0.0183	0.0083 ± 0.0015	0.8929 ± 0.0064	0.8919 ± 0.0075	0.2988 ± 0.0193	0.0232 ± 0.0047	0.9176 ± 0.0017	0.5000 ± 0.0000	0.5000 ± 0.0000	
DivideMix	14	1.00	0.9121 ± 0.0068	0.9179 ± 0.0055	0.2429 ± 0.0251	0.0146 ± 0.0041	0.9084 ± 0.0044	0.9068 ± 0.0046	0.2853 ± 0.0223	0.0288 ± 0.0065	0.9420 ± 0.0042	1.0000 ± 0.0000	1.0000 ± 0.0000	
DivideMix	15	0.20	0.8974 ± 0.0067	0.9096 ± 0.0076	0.2609 ± 0.0198	0.0136 ± 0.0020	0.8919 ± 0.0076	0.8937 ± 0.0077	0.3051 ± 0.0334	0.0303 ± 0.0087	0.9192 ± 0.0028	0.9999 ± 0.0000	0.9999 ± 0.0000	
DivideMix	16	1.00	0.9074 ± 0.0046	0.9136 ± 0.0040	0.2677 ± 0.0178	0.0239 ± 0.0021	0.9146 ± 0.0025	0.9133 ± 0.0059	0.2853 ± 0.0086	0.0325 ± 0.0024	0.9519 ± 0.0022	3.0000 ± 0.0000	3.0000 ± 0.0000	
DivideMix	17	0.20	0.8686 ± 0.0063	0.8645 ± 0.0098	0.3174 ± 0.0096	0.0223 ± 0.0067	0.8386 ± 0.0047	0.8247 ± 0.0180	0.3909 ± 0.0121	0.0239 ± 0.0062	0.8643 ± 0.0039	0.9999 ± 0.0001	0.9999 ± 0.0001	
DivideMix	18	1.00	0.9112 ± 0.0075	0.9180 ± 0.0076	0.2541 ± 0.0223	0.0189 ± 0.0051	0.9115 ± 0.0053	0.9095 ± 0.0071	0.2839 ± 0.0184	0.0296 ± 0.0073	0.9491 ± 0.0022	5.0000 ± 0.0000	5.0000 ± 0.0000	
DivideMix	19	1.00	0.8155 ± 0.0214	0.8105 ± 0.0255	0.4241 ± 0.0305	0.0507 ± 0.0052	0.7802 ± 0.0171	0.7634 ± 0.0185	0.4972 ± 0.0289	0.0185 ± 0.0024	0.7922 ± 0.0085	0.9999 ± 0.0001	0.9999 ± 0.0001	
DivideMix	20	1.00	0.9115 ± 0.0097	0.9176 ± 0.0091	0.2675 ± 0.0283	0.0219 ± 0.0076	0.9122 ± 0.0067	0.9120 ± 0.0072	0.2932 ± 0.0185	0.0314 ± 0.0046	0.9516 ± 0.0040	10.0000 ± 0.0000	10.0000 ± 0.0000	
Fixmatch	01	0.10	0.8721 ± 0.0117	0.8887 ± 0.0107	0.3755 ± 0.0236	0.0440 ± 0.0056	0.8622 ± 0.0000	0.8573 ± 0.0000	0.5699 ± 0.0000	0.0793 ± 0.0000	0.8493 ± 0.0069	0.1000 ± 0.0000	0.1000 ± 0.0000	
Fixmatch	02	0.20	0.8842 ± 0.0097	0.8972 ± 0.0081	0.3225 ± 0.0170	0.0320 ± 0.0058	N/A	N/A	N/A	N/A	0.8653 ± 0.0006	0.2000 ± 0.0000	0.2000 ± 0.0000	
Fixmatch	03	0.50	0.8838 ± 0.0109	0.9090 ± 0.0079	0.2738 ± 0.0263	0.0171 ± 0.0075	N/A	N/A	N/A	N/A	0.8927 ± 0.0041	0.5000 ± 0.0000	0.5000 ± 0.0000	
Fixmatch	04	1.00	0.8860 ± 0.0085	0.9179 ± 0.0072	0.2664 ± 0.0165	0.0126 ± 0.0024	N/A	N/A	N/A	N/A	0.9397 ± 0.0046	1.0000 ± 0.0000	1.0000 ± 0.0000	
Fixmatch + S2C2	01	0.10	0.8404 ± 0.0059	0.8518 ± 0.0053	0.4158 ± 0.0312	0.0436 ± 0.0082	N/A	N/A	N/A	N/A	0.7722 ± 0.0088	0.1000 ± 0.0000	0.1000 ± 0.0000	
Fixmatch + S2C2	02	0.20	0.8697 ± 0.0105	0.8837 ± 0.0091	0.3373 ± 0.0284	0.0311 ± 0.0076	N/A	N/A	N/A	N/A	0.8274 ± 0.0058	0.2000 ± 0.0000	0.2000 ± 0.0000	
Fixmatch + S2C2	03	0.50	0.8882 ± 0.0068	0.9069 ± 0.0068	0.2786 ± 0.0220	0.0163 ± 0.0044	N/A	N/A	N/A	N/A	0.9044 ± 0.0043	0.5000 ± 0.0000	0.5000 ± 0.0000	
Fixmatch + S2C2	04	1.00	0.8865 ± 0.0119	0.9165 ± 0.0089	0.2663 ± 0.0203	0.0163 ± 0.0044	N/A	N/A	N/A	N/A	0.9295 ± 0.0110	1.0000 ± 0.0000	1.0000 ± 0.0000	
Het	01	0.10	0.7180 ± 0.0097	0.7376 ± 0.0086	0.6760 ± 0.1040	0.0567 ± 0.0042	0.6351 ± 0.0030	0.6359 ± 0.0018	2.0201 ± 0.0932	0.2411 ± 0.0057	0.7075 ± 0.0113	0.1000 ± 0.0000	0.1000 ± 0.0000	
Het	02	0.20	0.8035 ± 0.0040	0.8105 ± 0.0040	0.4032 ± 0.0240	0.0268 ± 0.0082	0.7120 ± 0.0052	0.7052 ± 0.0056	1.7863 ± 0.0369	0.2207 ± 0.0061	0.6664 ± 0.0029	0.2000 ± 0.0000	0.2000 ± 0.0000	
Het	03	0.50	0.8624 ± 0.0095	0.8773 ± 0.0106	0.3596 ± 0.0183	0.0351 ± 0.0082	0.8282 ± 0.0099	0.8221 ± 0.0116	1.4552 ± 0.0598	0.1555 ± 0.0077	0.7571 ± 0.0059	0.5000 ± 0.0000	0.5000 ± 0.0000	
Het	04	1.00	0.8912 ± 0.0110	0.9025 ± 0.0109	0.3435 ± 0.0295	0.0356 ± 0.0080	0.8825 ± 0.0113	0.8255 ± 0.0175	1.2066 ± 0.0609	0.1260 ± 0.0046	0.8562 ± 0.0212	1.0000 ± 0.0000	1.0000 ± 0.0000	
Het	05	0.20	0.8636 ± 0.0063	0.8810 ± 0.0065	0.3776 ± 0.0319	0.0315 ± 0.0086	0.7353 ± 0.0077	0.7368 ± 0.0059	1.4082 ± 0.0530	0.1487 ± 0.0078	0.7571 ± 0.0071	0.9999 ± 0.0000	0.9999 ± 0.0000	
Het	06	0.20	0.8976 ± 0.0081	0.9106 ± 0.0071	0.4009 ± 0.0260	0.0370 ± 0.0104	0.8404 ± 0.0105	0.8404 ± 0.0105	1.1432 ± 0.0459	0.0935 ± 0.0051	0.8234 ± 0.0093	0.9999 ± 0.0000	0.9999 ± 0.0000	
Het	07	0.20	0.8227 ± 0.0093	0.8352 ± 0.0099	0.4860 ± 0.0364	0.0389 ± 0.0113	0.7833 ± 0.0114	0.7302 ± 0.0114	0.7302 ± 0.0114	1.1670 ± 0.0926	0.1823 ± 0.0029	0.7051 ± 0.0037	0.9999 ± 0.0001	0.9999 ± 0.0001
Het	08	0.50	0.9012 ± 0.0059	0.9123 ± 0.0084	0.3512 ± 0.0195	0.0431 ± 0.0042	0.8482 ± 0.0039	0.8430 ± 0.0077	1.126 ± 0.0535	0.114 ± 0.0044	0.8814 ± 0.0163	0.5000 ± 0.0000	0.5000 ± 0.0000	
Het	09	1.00	0.9192 ± 0.0149	0.9285 ± 0.0149	0.2857 ± 0.0245	0.0367 ± 0.0135	0.6585 ± 0.0145	0.6585 ± 0.0145	1.0161 ± 0.0460	0.0621 ± 0.0029	0.9161 ± 0.0041	1.0000 ± 0.0000	1.0000 ± 0.0000	
Het	10	1.00	0.8998 ± 0.0029	0.9146 ± 0.0076	0.3625 ± 0.0370	0.0463 ± 0.0073	0.8422 ± 0.0033	0.8441 ± 0.0077	1.1334 ± 0.0074	0.1156 ± 0.0025	0.8875 ± 0.0168	10.0000 ± 0.0000	10.0000 ± 0.0000	
Mean	01	0.10	0.8991 ± 0.0102	0.8541 ± 0.0101	0.6609 ± 0.0384	0.0909 ± 0.0444	0.8027 ± 0.0142	0.8402 ± 0.0114	0.8312 ± 0.1321	0.1167 ± 0.0127	0.7823 ± 0.0231	0.1000 ± 0.0000	0.1000 ± 0.0000	
Mean	02	0.20	0.8430 ± 0.0114	0.8796 ± 0.0077	0.5109 ± 0.0263	0.0377 ± 0.0101	0.8424 ± 0.0053	0.7800 ± 0.0307	0.9958 ± 0.0551	0.0811 ± 0.0049	0.9191 ± 0.0052	0.2000 ± 0.0000	0.2000 ± 0.0000	
Mean	03	0.50	0.8654 ± 0.0101	0.8997 ± 0.0112	0.3671 ± 0.0285	0.0506 ± 0.0075	0.8789 ± 0.0050	0.8929 ± 0.0051	1.4353 ± 0.0169	0.0597 ± 0.0029	0.8825 ± 0.0026	0.5000 ± 0.0000	0.5000 ± 0.0000	
Mean	04	1.00	0.8740 ± 0.0051	0.9091 ± 0.0058	0.3150 ± 0.0133	0.0391 ± 0.0036	0.8836 ± 0.0043	0.9092 ± 0.0033	1.3259 ± 0.0080	0.0346 ± 0.0048	0.9284 ± 0.0086	1.0000 ± 0.0000	1.0000 ± 0.0000	
Mean+S2C2	01	0.10	0.8142 ± 0.0116	0.8383 ± 0.0124	0.5084 ± 0.0581	0.0702 ± 0.0116	0.8501 ± 0.0146	0.8194 ± 0.0180	0.9140 ± 0.2070	0.178 ± 0.0203	0.7681 ± 0.0022	0.1000 ± 0.0000	0.1000 ± 0.0000	
Mean+S2C2	02	0.20	0.8689 ± 0.0094	0.8861 ± 0.0071	0.4009 ± 0.0260	0.0370 ± 0.0104	0.8501 ± 0.0146	0.8194 ± 0.0180	1.1432 ± 0.0459	0.0935 ± 0.0051	0.8234 ± 0.0093	0.2000 ± 0.0000	0.2000 ± 0.0000	
Mean+S2C2	03	0.50	0.8802 ± 0.0096	0.8980 ± 0.0097	0.3406 ± 0.0240	0.0369 ± 0.0053	0.8754 ± 0.0041	0.8872 ± 0.0065	1.5203 ± 0.0238	0.0658 ± 0.0069	0.8482 ± 0.0054	0.5000 ± 0.0000	0.5000 ± 0.0000	
Mean+S2C2	04	1.00	0.8890 ± 0.0121	0.9115 ± 0.0098	0.2996 ± 0.0239	0.0249 ± 0.0090	0.8563 ± 0.0043	0.9013 ± 0.0059	0.5939 ± 0.0112	0.0379 ± 0.0010	0.8988 ± 0.0159	1.0000 ± 0.0000	1.0000 ± 0.0000	
MOCoV1	01	0.10	0.8975 ± 0.0068	0.9161 ± 0.0076	0.1996 ± 0.0187	0.0136 ± 0.0011	0.8714 ± 0.0042	0.8714 ± 0.0042	0.3909 ± 0.0030	0.0261 ± 0.0002	0.9000 ± 0.0000	0.1000 ± 0.0000	0.1000 ± 0.0000	
MOCoV2	01	0.20	0.7628 ± 0.0060	0.7582 ± 0.0061	0.5973 ± 0.1102	0.0933 ± 0.1919	0.7085 ± 0.0045	0.6911 ± 0.0067	0.7691 ± 0.0147	0.2882 ± 0.0048	0.7430 ± 0.0064	0.2000 ± 0.0000	0.2000 ± 0.0000	
MOCoV2	02	0.50	0.7862 ± 0.0101	0.7838 ± 0.0022	0.5293 ± 0.0400	0.0932 ± 0.0102	0.7356 ± 0.0055	0.7213 ± 0.0048	0.6842 ± 0.01					

Method	m	n	F1	ACC	KL	ECE	F0.5	ACC	ACC	ECE	κ	η	β
BYOL	01	01	0.3876 ± 0.0252	0.4428 ± 0.0640	1.7667 ± 0.0465	0.3935 ± 0.1115	0.4123 ± 0.0444	0.4744 ± 0.0533	3.1910 ± 2.1229	0.3828 ± 0.1778	-0.0210 ± 0.0297	0.0968 ± 0.0000	0.0968 ± 0.0000
BYOL	01	02	0.4082 ± 0.0271	0.4755 ± 0.0172	2.2524 ± 0.0823	0.3587 ± 0.0396	0.4666 ± 0.0640	0.4859 ± 0.0532	4.7936 ± 0.3150	0.4619 ± 0.0756	-0.1304 ± 0.0897	0.1976 ± 0.0000	0.1976 ± 0.0000
BYOL	01	05	0.5523 ± 0.0681	0.5769 ± 0.0662	1.7833 ± 0.1397	0.2832 ± 0.0334	0.5125 ± 0.0567	0.5138 ± 0.0549	3.7137 ± 0.4669	0.4149 ± 0.0621	-0.0159 ± 0.0216	0.5000 ± 0.0000	0.5000 ± 0.0000
BYOL	01	10	0.4576 ± 0.0506	0.5045 ± 0.0188	1.0714 ± 0.1787	0.3317 ± 0.0358	0.5303 ± 0.0391	0.5308 ± 0.0384	3.4709 ± 0.4725	0.4088 ± 0.0486	0.1854 ± 0.0563	1.0000 ± 0.0000	1.0000 ± 0.0000
DivideMix	01	01	0.5114 ± 0.1521	0.5618 ± 0.1065	0.2375 ± 0.0925	0.1301 ± 0.0352	0.5039 ± 0.0755	0.5504 ± 0.0361	0.1516 ± 0.0477	0.1659 ± 0.1432	-0.0319 ± 0.0270	0.0968 ± 0.0000	0.0968 ± 0.0000
DivideMix	01	02	0.3941 ± 0.0028	0.5000 ± 0.0000	0.2405 ± 0.1325	0.1245 ± 0.0824	0.2921 ± 0.0409	0.4958 ± 0.0059	0.2050 ± 0.0691	0.1153 ± 0.0683	-0.0020 ± 0.0029	0.1976 ± 0.0000	0.1976 ± 0.0000
DivideMix	01	05	0.5035 ± 0.1534	0.5682 ± 0.0964	0.4484 ± 0.3751	0.1189 ± 0.0392	0.5178 ± 0.1735	0.5852 ± 0.1205	-0.1945 ± 0.1307	0.0000 ± 0.0000	N/A	0.5000 ± 0.0000	0.5000 ± 0.0000
DivideMix	01	10	0.4245 ± 0.0417	0.5152 ± 0.0214	0.4584 ± 0.2640	0.1866 ± 0.1017	0.4507 ± 0.0787	0.5303 ± 0.0429	-0.2258 ± 0.0882	0.0000 ± 0.0000	N/A	1.0000 ± 0.0000	1.0000 ± 0.0000
DivideMix	02	05	0.2589 ± 0.0042	0.5000 ± 0.0000	0.9107 ± 0.3204	0.2598 ± 0.0579	0.2589 ± 0.0042	0.5000 ± 0.0000	0.5951 ± 0.2229	0.4554 ± 0.0666	N/A	1.0000 ± 0.0000	1.0000 ± 0.0000
DivideMix	05	02	0.4528 ± 0.0858	0.5317 ± 0.0449	0.9881 ± 0.6769	0.1670 ± 0.0375	0.5588 ± 0.2556	0.6307 ± 0.1848	-0.2171 ± 0.1388	0.0000 ± 0.0000	N/A	3.0000 ± 0.0000	3.0000 ± 0.0000
DivideMix	05	02	0.3941 ± 0.0028	0.5000 ± 0.0000	0.2650 ± 0.0715	0.1712 ± 0.0354	0.3941 ± 0.0028	0.5000 ± 0.0000	0.2491 ± 0.0625	0.1413 ± 0.0662	N/A	0.9879 ± 0.0000	0.9879 ± 0.0000
DivideMix	05	10	0.4941 ± 0.1401	0.5606 ± 0.0857	0.1737 ± 0.0453	0.1043 ± 0.0217	0.5380 ± 0.2021	0.6080 ± 0.1527	-0.3307 ± 0.0403	0.0000 ± 0.0000	0.0257 ± 0.0363	5.0000 ± 0.0000	5.0000 ± 0.0000
DivideMix	10	01	0.4507 ± 0.0077	0.5303 ± 0.0429	0.3316 ± 0.1985	0.1621 ± 0.0820	0.4911 ± 0.1357	0.5610 ± 0.0862	0.2239 ± 0.1260	0.1288 ± 0.0665	-0.0482 ± 0.0959	0.9877 ± 0.0000	0.9877 ± 0.0000
DivideMix	10	01	0.5106 ± 0.1304	0.5685 ± 0.0807	1.3351 ± 0.8443	0.1922 ± 0.0591	0.7144 ± 0.2281	0.7499 ± 0.1768	-0.0955 ± 0.2390	0.0000 ± 0.0000	0.2913 ± 0.1017	10.0000 ± 0.0000	10.0000 ± 0.0000
ELR+	01	01	0.3507 ± 0.0028	0.5000 ± 0.0000	1.0651 ± 0.4053	0.3998 ± 0.1489	0.3507 ± 0.0028	0.5000 ± 0.0000	1.1410 ± 0.4974	0.4067 ± 0.1507	N/A	0.9688 ± 0.0000	0.9688 ± 0.0000
ELR+	01	02	0.3941 ± 0.0028	0.5000 ± 0.0000	0.9736 ± 0.3767	0.3182 ± 0.0882	0.3941 ± 0.0028	0.5000 ± 0.0000	0.8746 ± 0.1536	0.5150 ± 0.0140	N/A	0.1976 ± 0.0000	0.1976 ± 0.0000
ELR+	01	05	0.3941 ± 0.0028	0.5000 ± 0.0000	0.9555 ± 0.8235	0.2609 ± 0.0533	0.4099 ± 0.0211	0.5076 ± 0.0107	0.7676 ± 0.4255	0.2692 ± 0.0842	N/A	0.5000 ± 0.0000	0.5000 ± 0.0000
ELR+	01	10	0.3941 ± 0.0028	0.5000 ± 0.0000	1.4424 ± 0.8302	0.2979 ± 0.0700	0.3941 ± 0.0028	0.5000 ± 0.0000	0.9467 ± 0.1838	0.3210 ± 0.0200	N/A	1.0000 ± 0.0000	1.0000 ± 0.0000
ELR+	02	05	0.2589 ± 0.0042	0.5000 ± 0.0000	1.8229 ± 1.1816	0.5854 ± 0.0449	0.2589 ± 0.0042	0.5000 ± 0.0000	1.4355 ± 0.2269	0.6068 ± 0.0225	N/A	1.0000 ± 0.0000	1.0000 ± 0.0000
ELR+	05	02	0.3941 ± 0.0028	0.5000 ± 0.0000	1.5959 ± 1.0502	0.2957 ± 0.0642	0.3941 ± 0.0028	0.5000 ± 0.0000	0.8352 ± 0.1046	0.3111 ± 0.0147	N/A	3.0000 ± 0.0000	3.0000 ± 0.0000
ELR+	05	10	0.3901 ± 0.0075	0.4917 ± 0.0118	1.7464 ± 1.6319	0.2416 ± 0.1352	0.3921 ± 0.0040	0.4958 ± 0.0059	0.8547 ± 0.5370	0.2443 ± 0.1225	N/A	0.9879 ± 0.0000	0.9879 ± 0.0000
ELR+	10	01	0.3941 ± 0.0028	0.5000 ± 0.0000	2.8352 ± 1.1547	0.4237 ± 0.0986	0.3941 ± 0.0028	0.5000 ± 0.0000	1.1284 ± 0.0847	0.3277 ± 0.0100	N/A	5.0000 ± 0.0000	5.0000 ± 0.0000
ELR+	10	01	0.3507 ± 0.0028	0.5000 ± 0.0000	1.8608 ± 1.4022	0.3974 ± 0.0954	0.3507 ± 0.0028	0.5000 ± 0.0000	0.9649 ± 0.4867	0.3624 ± 0.0565	N/A	0.9877 ± 0.0000	0.9877 ± 0.0000
ELR+	10	01	0.3941 ± 0.0028	0.5000 ± 0.0000	1.2563 ± 0.5650	0.3058 ± 0.0241	0.3941 ± 0.0028	0.5000 ± 0.0000	0.8273 ± 0.1527	0.3079 ± 0.0087	N/A	10.0000 ± 0.0000	10.0000 ± 0.0000
Fixmatch	01	01	0.5354 ± 0.0844	0.705 ± 0.0883	2.8881 ± 1.2537	0.4150 ± 0.0973	0.5029 ± 0.0646	0.5440 ± 0.0436	4.4583 ± 0.7279	0.4871 ± 0.0630	-0.0314 ± 0.3157	0.0968 ± 0.0000	0.0968 ± 0.0000
Fixmatch	01	02	0.5373 ± 0.0843	0.5561 ± 0.0267	1.6716 ± 0.2950	0.2227 ± 0.0765	0.5407 ± 0.0734	0.5543 ± 0.0261	3.5638 ± 1.3333	0.4528 ± 0.0395	-0.0166 ± 0.2348	0.1976 ± 0.0000	0.1976 ± 0.0000
Fixmatch	01	05	0.5440 ± 0.0929	0.5924 ± 0.0833	1.6682 ± 0.6100	0.2867 ± 0.0437	0.5129 ± 0.0430	0.5150 ± 0.0517	1.7527 ± 0.2873	0.3896 ± 0.0566	0.0528 ± 0.0471	0.5000 ± 0.0000	0.5000 ± 0.0000
Fixmatch	01	10	0.7304 ± 0.0077	0.7157 ± 0.0026	0.9890 ± 0.3003	0.1652 ± 0.0654	0.6308 ± 0.0296	0.6283 ± 0.0262	1.0441 ± 0.2233	0.2296 ± 0.0203	0.3836 ± 0.1161	1.0000 ± 0.0000	1.0000 ± 0.0000
Fixmatch + S2C2	01	01	0.4160 ± 0.1365	0.508 ± 0.0435	2.1917 ± 1.1148	0.4060 ± 0.1656	0.5724 ± 0.0000	0.5880 ± 0.0000	2.1986 ± 0.0000	0.2896 ± 0.0000	-0.0166 ± 0.0510	0.9688 ± 0.0000	0.9688 ± 0.0000
Fixmatch + S2C2	01	02	0.4845 ± 0.1497	0.5000 ± 0.0000	1.4764 ± 0.5344	0.2944 ± 0.0968	0.4845 ± 0.1497	0.5000 ± 0.0000	1.9891 ± 0.0000	0.3348 ± 0.0000	-0.0166 ± 0.0510	0.1976 ± 0.0000	0.1976 ± 0.0000
Fixmatch + S2C2	01	05	0.570 ± 0.0400	0.6163 ± 0.0270	1.4494 ± 0.4580	0.2932 ± 0.0426	N/A	N/A	N/A	N/A	0.2026 ± 0.0628	0.5000 ± 0.0000	0.5000 ± 0.0000
Fixmatch + S2C2	01	10	0.6140 ± 0.0975	0.6288 ± 0.0683	1.4053 ± 0.4179	0.2537 ± 0.0380	0.6253 ± 0.0193	0.6274 ± 0.0124	1.3231 ± 0.1553	0.2839 ± 0.0095	0.4801 ± 0.0990	1.0000 ± 0.0000	1.0000 ± 0.0000
Mean	01	02	0.2815 ± 0.0263	0.4660 ± 0.0327	0.9560 ± 0.3770	0.2457 ± 0.0764	0.5447 ± 0.0674	0.5672 ± 0.0319	3.2532 ± 1.3520	0.4380 ± 0.1536	0.0968 ± 0.0000	0.0968 ± 0.0000	0.0968 ± 0.0000
Mean	01	02	0.4115 ± 0.1343	0.5036 ± 0.0743	0.8013 ± 0.4976	0.3428 ± 0.0600	0.4699 ± 0.0587	0.5382 ± 0.0302	1.4664 ± 1.0702	0.3295 ± 0.0966	-0.0795 ± 0.1883	1.0000 ± 0.0000	1.0000 ± 0.0000
Mean	01	05	0.5122 ± 0.1723	0.5931 ± 0.0780	0.2652 ± 0.0883	0.1390 ± 0.0932	0.5156 ± 0.1836	0.5929 ± 0.0756	0.1950 ± 0.0318	0.1978 ± 0.0950	0.0605 ± 0.0856	0.5000 ± 0.0000	0.5000 ± 0.0000
Mean	01	10	0.4538 ± 0.0586	0.5451 ± 0.0402	0.8548 ± 0.5012	0.2545 ± 0.0468	0.4538 ± 0.0586	0.5451 ± 0.0402	0.8548 ± 0.5012	0.2545 ± 0.0468	0.1978 ± 0.0950	0.9688 ± 0.0000	0.9688 ± 0.0000
Mean+S2C2	01	01	0.5162 ± 0.0589	0.5605 ± 0.0460	2.1551 ± 1.1033	0.3731 ± 0.0967	0.5757 ± 0.0618	0.6017 ± 0.0355	2.2855 ± 0.8693	0.3563 ± 0.0988	0.0688 ± 0.0935	0.9688 ± 0.0000	0.9688 ± 0.0000
Mean+S2C2	01	02	0.5652 ± 0.1237	0.6013 ± 0.0648	2.1546 ± 0.8013	0.2368 ± 0.1589	0.5876 ± 0.0311	0.6099 ± 0.0313	2.6191 ± 0.4734	0.3526 ± 0.0355	0.0667 ± 0.0287	0.1976 ± 0.0000	0.1976 ± 0.0000
Mean+S2C2	01	05	0.5602 ± 0.1217	0.6307 ± 0.0972	1.4299 ± 0.9798	0.2610 ± 0.0551	0.6451 ± 0.0548	0.6609 ± 0.0682	0.8951 ± 0.5244	0.2171 ± 0.0597	0.1051 ± 0.0775	0.5000 ± 0.0000	0.5000 ± 0.0000
Mean+S2C2	01	10	0.5997 ± 0.0959	0.6255 ± 0.0715	1.4179 ± 0.9798	0.2230 ± 0.0664	0.6451 ± 0.0548	0.6609 ± 0.0682	0.8951 ± 0.5244	0.2171 ± 0.0597	0.1051 ± 0.0775	0.5000 ± 0.0000	0.5000 ± 0.0000
MOCOv2	01	01	0.5094 ± 0.0575	0.5285 ± 0.0424	1.0560 ± 0.6299	0.2875 ± 0.1028	0.5346 ± 0.0367	0.5470 ± 0.0361	2.9361 ± 1.9313	0.3182 ± 0.1338	-0.0968 ± 0.0348	0.9688 ± 0.0000	0.9688 ± 0.0000
MOCOv2	01	02	0.4479 ± 0.1155	0.5182 ± 0.0858	1.9671 ± 0.9367	0.3009 ± 0.0889	0.4848 ± 0.0542	0.5118 ± 0.0529	2.9719 ± 0.5079	0.3755 ± 0.0646	0.0689 ± 0.1664	0.1976 ± 0.0000	0.1976 ± 0.0000
MOCOv2	01	05	0.4287 ± 0.0258	0.5064 ± 0.0124	0.9143 ± 0.6770	0.2744 ± 0.0324	0.5322 ± 0.0154	0.5349 ± 0.0166	0.8613 ± 0.5261	0.2907 ± 0.0815	0.0635 ± 0.0799	0.5000 ± 0.0000	0.5000 ± 0.0000
MOCOv2	01	10	0.5127 ± 0.0351	0.5613 ± 0.0209	2.932 ± 0.1087	0.1580 ± 0.0323	0.6508 ± 0.0506	0.6477 ± 0.0555	0.2716 ± 0.1221	0.1394 ± 0.0666	0.2677 ± 0.0897	1.0000 ± 0.0000	1.0000 ± 0.0000
Baseline	01	01	0.3507 ± 0.0028	0.5000 ± 0.0000	4.9204 ± 1.0767	0.4462 ± 0.1408	0.6785 ± 0.0757	0.6799 ± 0.0551	3.8083 ± 0.1782	0.2957 ± 0.0548	0.1140 ± 0.0889	0.9688 ± 0.0000	0.9688 ± 0.0000
Baseline	01	02	0.4295 ± 0.1321	0.4783 ± 0.0815	1.1238 ± 1.4298	0.2984 ± 0.1437	0.4992 ± 0.0526	0.7005 ± 0.0284	3.7132 ± 0.1157	0.2742 ± 0.0475	0.3316 ± 0.1088	0.9688 ± 0.0000	0.9688 ± 0.0000
Baseline	01	05	0.6419 ± 0.0559	0.6761 ± 0.0217	1.1234 ± 0.3645	0.2443 ± 0.0654	0.7207 ± 0.0639	0.7297 ± 0.0622	3.7159 ± 0.3444	0.2634 ± 0.0650	0.1937 ± 0.0117	0.5000 ± 0.0000	0.5000 ± 0.0000
Baseline	01	10	0.6711 ± 0.0445	0.6662 ± 0.0355	1.7292 ± 0.4832	0.2142 ± 0.0204	0.7505 ± 0.1017	0.7581 ± 0.0017	3.4827 ± 0.1190	0.2312 ± 0.0201	1.0000 ± 0.0000	1.0000 ± 0.0000	1.0000 ± 0.0000
Baseline	01	02	0.4994 ± 0.0902	0.5157 ± 0.0892	0.6419 ± 0.2084	0.1413 ± 0.0384	0.4687 ± 0.0191	0.4687 ± 0.0191	0.8414 ± 0.1683	0.391 ± 0.0567	0.747 ± 0.0751	0.1976 ± 0.0000	0.1976 ± 0.0000
Baseline	03	01	0.5825 ± 0.1029	0.6086 ± 0.0620	0.3873 ± 0.1927	0.1630 ± 0.0696	0.7549 ± 0.0861	0.7617 ± 0.0770	1.1571 ± 0.2343	0.1093 ± 0.0217	1.0000 ± 0.0000	3.0000 ± 0.0000	3.0000 ± 0.0000
Baseline	05	02	0.4733 ± 0.0789	0.5560 ± 0.0632	0.5558 ± 0.2512	0.2459 ± 0.0627	0.7883 ± 0.0088	0.8010 ± 0.0132	0.4824 ± 0.0574	0.0387 ± 0.0274	0.3368 ± 0.1070	0.9879 ± 0.0000	0.9879 ± 0.0000
Baseline	10	01	0.6036 ± 0.0930	0.6219 ± 0.0643	0.5230 ± 0.2831	0.1761 ± 0.0841	0.8287 ± 0.0549	0.8286 ± 0.0493	0.5882 ± 0.0463	0.0753 ± 0.0145	1.0000 ± 0.0000	5.0000 ± 0.0000	5.0000 ± 0.0000
Baseline	10	01	0.4573 ± 0.0952	0.5310 ± 0.0117	1.6119 ± 0.6793	0.3597 ± 0.0924	0.4573 ± 0.0952	0.5310 ± 0.0117	1.6119 ± 0.6793	0.3597 ± 0.0924	0.5310 ± 0.0117	1.0000 ± 0.0000	1.0000 ± 0.0000
Baseline	10	01	0.5141 ± 0.0222	0.5617 ± 0.0116	0.3069 ±								









