Supplementary material – Hubs and Hyperspheres: Reducing Hubness and Improving Transductive Few-shot Learning with Hyperspherical Embeddings

Daniel J. Trosten*,† Rwiddhi Chakraborty*,† Sigurd Løkse,† Kristoffer Knutsen Wickstrøm,†
Robert Jenssen†‡§¶ Michael C. Kampffmeyer†‡
Department of Physics and Technology, UiT The Arctic University of Norway
firstname[.middle initial].lastname@uit.no

1. Introduction

Here we provide proofs for our theoretical results on the hyperspherical uniform and hubness; the decomposition of the KL divergence; and the minima of our methods' loss functions. We also give additional details on the implementation and hyperparameters for noHub and noHub-S– and include the complete tables of 1-shot and 5-shot results for all classifiers, datasets and feature extractors. Finally, we briefly reflect on potential negative societal impacts of our work.

2. Hyperspherical Uniform Eliminates Hubness

Definition 1 (Uniform PDF on the hypersphere.). *The uniform probability density function (PDF) on the unit hypersphere $\mathbb{S}_d = \{\boldsymbol{x} \in \mathbb{R}^d \mid ||\boldsymbol{x}|| = 1\} \subset \mathbb{R}^d$ is*

$$u_{\mathbb{S}_d}(\boldsymbol{x}) = A_d^{-1} \delta(||\boldsymbol{x}|| - 1) \tag{1}$$

where $A_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}$ is the surface area of $\mathbb{S}_d$, and $\delta(\cdot)$ is the Dirac delta distribution.

Lemma 1 (Trisection of hypersphere). *The trisection of the hypersphere along coordinate i is given by the three-tuple of disjoint sets $(\mathbb{S}_d^{i,+}, \mathbb{S}_d^{i,-}, \mathbb{S}_d^{i,0})$ where*

$$\mathbb{S}_d^{i,+} = \{\boldsymbol{x} = [x^1, \dots, x^d]^\top \in \mathbb{S}_d \mid x^i > 0\} \tag{2}$$

$$\mathbb{S}_d^{i,-} = \{\boldsymbol{x} = [x^1, \dots, x^d]^\top \in \mathbb{S}_d \mid x^i < 0\} \tag{3}$$

$$\mathbb{S}_d^{i,0} = \{\boldsymbol{x} = [x^1, \dots, x^d]^\top \in \mathbb{S}_d \mid x^i = 0\} \tag{4}$$

and

$$\mathbb{S}_d^{i,+} \cup \mathbb{S}_d^{i,-} \cup \mathbb{S}_d^{i,0} = \mathbb{S}_d \tag{5}$$

Then we have

$$\mathbb{S}_d^{i,+} = -\mathbb{S}_d^{i,-} = \{-\boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{S}_d^{i,-}\} \tag{6}$$

Proof. Let $\boldsymbol{x} \in \mathbb{S}_d^{i,+}$, then

$$||(-x)|| = ||x|| = 1, \tag{7}$$

and

$$-(x^i) < 0. \tag{8}$$

Hence $\boldsymbol{x} \in -\mathbb{S}_d^{i,-}$, and $\mathbb{S}_d^{i,+} \subseteq -\mathbb{S}_d^{i,-}$.

Similarly, let $-\boldsymbol{x} \in -\mathbb{S}_d^{i,-}$, then

$$|| - (-x)|| = ||x|| = 1, \tag{9}$$

and

$$-(-x^i) = x^i > 0. \tag{10}$$

Hence $\boldsymbol{x} \in \mathbb{S}_d^{i,+}$, and $-\mathbb{S}_d^{i,-} \subseteq \mathbb{S}_d^{i,+}$.

It then follows that $\mathbb{S}_d^{i,+} = -\mathbb{S}_d^{i,-}$. □

Proposition 1. *Suppose $\boldsymbol{X}$ has PDF $u_{\mathbb{S}_d}(\boldsymbol{x})$. Then*

$$\mathbb{E}(\boldsymbol{X}) = 0 \tag{11}$$

Proof. The expectation $\mathbb{E}(\boldsymbol{X})$ is given by

$$\mathbb{E}(\boldsymbol{X}) = \int_{\mathbb{R}^d} \boldsymbol{x} u_{\mathbb{S}_d}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \tag{12}$$

Since $u_{\mathbb{S}_d}$ is non-zero only on the hypersphere $\mathbb{S}_d$, the integral can be rewritten as a surface integral over $\mathbb{S}_d$

$$\mathbb{E}(\boldsymbol{X}) = \int_{\mathbb{S}_d} \boldsymbol{x} A_d^{-1} \mathrm{d}S. \tag{13}$$

*Equal contributions.
†UiT Machine Learning group (machine-learning.uit.no) and Visual Intelligence Centre (visual-intelligence.no).
‡Norwegian Computing Center.
§Department of Computer Science, University of Copenhagen.
¶Pioneer Centre for AI (aicentre.dk).

Decomposing the integral over the trisection of $\mathbb{S}_d$ along coordinate i gives

$$\int_{\mathbb{S}_d} \boldsymbol{x} A_d^{-1} \mathrm{d}S = \tag{14}$$

$$A_d^{-1} \left(\int_{\mathbb{S}_d^{i,+}} \boldsymbol{x} \mathrm{d}S + \int_{\mathbb{S}_d^{i,-}} \boldsymbol{x} \mathrm{d}S + \int_{\mathbb{S}_d^{i,0}} \boldsymbol{x} \mathrm{d}S \right). \tag{15}$$

By Lemma 1 we have

$$\mathbb{S}_d^{i,+} = -\mathbb{S}_d^{i,-} \Rightarrow \int_{\mathbb{S}_d^{i,+}} \boldsymbol{x} \mathrm{d}S = - \int_{\mathbb{S}_d^{i,-}} \boldsymbol{x} \mathrm{d}S. \tag{16}$$

Furthermore, since the set $\mathbb{S}_d^{i,0}$ has zero width along coordinate i, $\int_{\mathbb{S}_d^{i,0}} \boldsymbol{x} \mathrm{d}S = 0$. Hence

$$\mathbb{E}(\boldsymbol{X}) = \tag{17}$$

$$A_d^{-1} \left(\int_{\mathbb{S}_d^{i,+}} \boldsymbol{x} \mathrm{d}S - \int_{\mathbb{S}_d^{i,+}} \boldsymbol{x} \mathrm{d}S + \int_{\mathbb{S}_d^{i,0}} \boldsymbol{x} \mathrm{d}S \right) = 0 \tag{18}$$

□

Proposition 2. *Let $\Pi_{\boldsymbol{p}}$ be the tangent plane of $\mathbb{S}_d$ at an arbitrary point $\boldsymbol{p} \in \mathbb{S}_d$. Then, for any direction $\boldsymbol{\theta}^*$ in $\Pi_{\boldsymbol{p}}$ the directional derivative of $u_{\mathbb{S}_d}$ along $\boldsymbol{\theta}^*$ is*

$$\nabla_{\boldsymbol{\theta}^*} u_{\mathbb{S}_d} = 0 \tag{19}$$

Proof. $u_{\mathbb{S}_d}(\boldsymbol{x})$ can be written in polar coordinates as

$$u_{\mathbb{S}_d}(\boldsymbol{x}(r, \boldsymbol{\theta})) = u_{\mathbb{S}_d}^{\text{Polar}}(r, \boldsymbol{\theta}) = A_d^{-1} \delta(r-1) \tag{20}$$

The gradient of $u_{\mathbb{S}_d}^{\text{Polar}}(r, \boldsymbol{\theta})$ is then

$$\nabla_{(r,\boldsymbol{\theta})} u_{\mathbb{S}_d}^{\text{Polar}}(r, \boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial}{\partial r} u_{\mathbb{S}_d}^{\text{Polar}}(r, \boldsymbol{\theta}) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{21}$$

For an arbitrary point $\boldsymbol{p} \in \mathbb{S}_d$, an arbitrary unit vector (direction), $\boldsymbol{\theta}^*$, in the tangent plane $\Pi_{\boldsymbol{p}}$ is given by

$$\boldsymbol{\theta}^* = \begin{bmatrix} 0 \\ \theta_1^* \\ \vdots \\ \theta_{d-1}^* \end{bmatrix} \tag{22}$$

The directional derivative of $u_{\mathbb{S}_d}(\boldsymbol{x})$ along $\boldsymbol{\theta}^*$ is then

$$\nabla_{\boldsymbol{\theta}^*} u_{\mathbb{S}_d}(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial}{\partial r} u_{\mathbb{S}_d}^{\text{Polar}}(r, \boldsymbol{\theta}) \\ 0 \\ \vdots \\ 0 \end{bmatrix}^\top \cdot \begin{bmatrix} 0 \\ \theta_1^* \\ \vdots \\ \theta_{d-1}^* \end{bmatrix} = 0 \tag{23}$$

□

3. Method

Computing κ_i. Following [5], we compute κ_i using a binary search such that

$$|\log_2(P) - H(P_i)| \leq 0.1 \cdot \log_2(P) \tag{24}$$

where P is a hyperparameter, and $H(P_i)$ is the Shannon entropy of similarities for representation i

$$H(P_i) = \sum_{j=1}^{n} p_{i|j} \log_2(p_{i|j}). \tag{25}$$

Decomposition of $KL(P||Q)$. Recall that

$$p_{ij} = \frac{p_{i|j} + p_{j|i}}{2}, \quad p_{i|j} = \frac{\exp(\kappa_i \boldsymbol{x}_i^\top \boldsymbol{x}_j)}{\sum_{l,m} \exp(\kappa_i \boldsymbol{x}_l^\top \boldsymbol{x}_m)} \tag{26}$$

and

$$q_{ij} = \frac{\exp(\kappa \boldsymbol{z}_i^\top \boldsymbol{z}_j)}{\sum_{l,m} \exp(\kappa \boldsymbol{z}_l^\top \boldsymbol{z}_m)}. \tag{27}$$

Since p_{ij} is constant w.r.t. q_{ij}, we have

$$\underset{\boldsymbol{z}_1,\dots,\boldsymbol{z}_n \in \mathbb{S}_d}{\arg\min} \; KL(P||Q) = \underset{\boldsymbol{z}_1,\dots,\boldsymbol{z}_n \in \mathbb{S}_d}{\arg\min} \sum_{i,j} p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{28}$$

$$= \underset{\boldsymbol{z}_1,\dots,\boldsymbol{z}_n \in \mathbb{S}_d}{\arg\min} \underbrace{\sum_{i,j} p_{ij} \log p_{ij}}_{\text{constant}} - \sum_{i,j} p_{ij} \log q_{ij} \tag{29}$$

$$= \underset{\boldsymbol{z}_1,\dots,\boldsymbol{z}_n \in \mathbb{S}_d}{\arg\min} \underbrace{- \sum_{i,j} p_{ij} \log q_{ij}}_{=:\tilde{\mathcal{L}}} \tag{30}$$

Minimizing $KL(P||Q)$ over $\boldsymbol{z}_1, \dots, \boldsymbol{z}_n \in \mathbb{S}_d$ is therefore equivalent to minimizing $\tilde{\mathcal{L}}$.

Decomposing $\tilde{\mathcal{L}}$ gives

$$\tilde{\mathcal{L}} = - \sum_{i,j} p_{ij} \kappa \boldsymbol{z}_i^\top \boldsymbol{z}_j + \tag{31}$$

$$\sum_{i,j} \left(p_{ij} \log \sum_{l,m} \exp(\kappa \boldsymbol{z}_l^\top \boldsymbol{z}_m) \right) \tag{32}$$

$$= - \sum_{i,j} p_{ij} \kappa \boldsymbol{z}_i^\top \boldsymbol{z}_j \tag{33}$$

$$+ \left(\log \sum_{l,m} \exp(\kappa \boldsymbol{z}_l^\top \boldsymbol{z}_m) \right) \cdot \underbrace{\left(\sum_{i,j} p_{ij} \right)}_{=1} \tag{34}$$

$$= \underbrace{- \sum_{i,j} p_{ij} \kappa \boldsymbol{z}_i^\top \boldsymbol{z}_j}_{=: \mathcal{L}_{\text{LSP}}} + \underbrace{\log \sum_{l,m} \exp(\kappa \boldsymbol{z}_l^\top \boldsymbol{z}_m)}_{=: \mathcal{L}_{\text{Unif}}} \tag{35}$$

Thus, we have shown that

$$\underset{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_n\in\mathbb{S}_d}{\arg\min}\; KL(P||Q) = \underset{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_n\in\mathbb{S}_d}{\arg\min}\; \mathcal{L}_{\text{LSP}} + \mathcal{L}_{\text{Unif}}. \quad (36)$$

4. Theoretical Results

Proposition 3. *Let $W_{ij} = \frac{1}{2}\kappa p_{ij}$, where $\sum_{i,j} p_{ij} = 1$, and let $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n \in \mathbb{S}_d$. Then we have*

$$\mathcal{L}_{\text{LSP}} = \sum_{i,j} \|\boldsymbol{z}_i - \boldsymbol{z}_j\|^2 W_{ij} - \kappa. \quad (37)$$

Proof. We have

$$\begin{aligned}
\mathcal{L}_{\text{LSP}} &= -\kappa \sum_{i,j} p_{ij} \boldsymbol{z}_i^\top \boldsymbol{z}_j && (38)\\
&= -2\sum_{i,j} \frac{1}{2}\kappa p_{ij} \boldsymbol{z}_i^\top \boldsymbol{z}_j + \sum_{i,j} 2\frac{1}{2}\kappa p_{ij} - \kappa && (39)\\
&(\textstyle\sum_{i,j} p_{ij} = 1)\\
&= -2\sum_{i,j} \boldsymbol{z}_i^\top \boldsymbol{z}_j W_{ij} + \sum_{i,j} (\|\boldsymbol{z}_i\| + \|\boldsymbol{z}_j\|) W_{ij} - \kappa && (40)\\
&(\|\boldsymbol{z}_i\| = \|\boldsymbol{z}_j\| = 1)\\
&= \sum_{i,j} (\|\boldsymbol{z}_i\| - 2\boldsymbol{z}_i^\top \boldsymbol{z}_j + \|\boldsymbol{z}_j\|) W_{ij} - \kappa && (41)\\
&= \sum_{i,j} |\boldsymbol{z}_i - \boldsymbol{z}_j\|^2 W_{ij} - \kappa. && (42)
\end{aligned}$$

□

Proposition 4 (Minimizing $\mathcal{L}_{\text{Unif}}$ maximizes entropy)**.** *Let $H_2(\cdot)$ be the 2-order Rényi entropy, estimated with a kernel density estimator using a Gaussian kernel. Then*

$$\underset{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_n\in\mathbb{S}_d}{\arg\min}\; \mathcal{L}_{\text{Unif}} = \underset{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_n\in\mathbb{S}_d}{\arg\max}\; H_2(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n). \quad (43)$$

Proof. Using a Gaussian kernel, the 2-order Rényi entropy can be estimated as [4, Eq. (2.13)]

$$H_2(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n) = -\log\left(\frac{1}{n^2}\sum_{l,m} \exp(-\frac{1}{2}\kappa\|\boldsymbol{z}_l - \boldsymbol{z}_m\|^2)\right) \quad (44)$$

Thus, we have

$$\begin{aligned}
&\underset{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_n\in\mathbb{S}_d}{\arg\max}\; H_2(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n) && (45)\\
&= \underset{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_n\in\mathbb{S}_d}{\arg\max}\; -\log\left(\frac{1}{n^2}\sum_{l,m} \exp(-\frac{1}{2}\kappa\|\boldsymbol{z}_l - \boldsymbol{z}_m\|^2)\right) && (46)\\
&= \underset{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_n\in\mathbb{S}_d}{\arg\min}\; \log\left(\sum_{l,m} \exp(-\frac{1}{2}\kappa\|\boldsymbol{z}_l - \boldsymbol{z}_m\|^2)\right) && (47)\\
&= \underset{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_n\in\mathbb{S}_d}{\arg\min}\; \log\left(\sum_{l,m} \exp(-\frac{1}{2}\kappa(\|\boldsymbol{z}_l\|^2\right. && (48)\\
&\left. - 2\boldsymbol{z}_l^\top \boldsymbol{z}_m + \|\boldsymbol{z}_m\|^2)\right) && (49)\\
&= \underset{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_n\in\mathbb{S}_d}{\arg\min}\; \log\left(\sum_{l,m} \exp(-\kappa(1 - \boldsymbol{z}_l^\top \boldsymbol{z}_m))\right) && (50)\\
&(\|\boldsymbol{z}_l\| = \|\boldsymbol{z}_m\| = 1)\\
&= \underset{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_n\in\mathbb{S}_d}{\arg\min}\; \log\left(\exp(-\kappa)\sum_{l,m} \exp(\kappa\boldsymbol{z}_l^\top \boldsymbol{z}_m)\right) && (51)\\
&= \underset{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_n\in\mathbb{S}_d}{\arg\min}\; \log\sum_{l,m} \exp(\kappa\boldsymbol{z}_l^\top \boldsymbol{z}_m) && (52)\\
&= \underset{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_n\in\mathbb{S}_d}{\arg\min}\; \mathcal{L}_{\text{Unif}}. && (53)
\end{aligned}$$

□

Definition 2 (Normalized counting measure)**.** *The normalized counting measure associated with a set B on A is*

$$\nu_B(A) = \frac{|B \cap A|}{|B|} \quad (54)$$

Definition 3 (Normalized surface area measure on $\mathbb{S}_d$)**.** *The normalized surface area measure on the hypersphere $\mathbb{S}_d \subset \mathbb{R}^d$, of a subset $S' \subset \mathbb{S}_d$ is*

$$\sigma_d(S') = \frac{\int_{S'} \mathrm{d}S}{\int_{\mathbb{S}_d} \mathrm{d}S} = A_d^{-1}\int_{S'} \mathrm{d}S \quad (55)$$

where A_d is defined as in Eq. (1), and $\int \mathrm{d}S$ denotes the surface integral on $\mathbb{S}_d$.

Definition 4 (Weak* convergence of measures [8])**.** *A sequence of Borel measures $\{\mu_n\}_{n=1}^{\infty}$ in $\mathbb{R}^d$ converges weak* to a Borel measure μ, if for all continuous functions $f : \mathbb{R}^d \to \mathbb{R}$,*

$$\lim_{n\to\infty} \int f(x)\mathrm{d}\mu_n(x) = \int f(x)\mathrm{d}\mu(x) \quad (56)$$

Proposition 5 (Minimizer of $\mathcal{L}_{\text{Unif}}$). *For each $n > 0$, the n point minimizer of $\mathcal{L}_{\text{Unif}}$ is*

$$z_1^\star, \ldots, z_n^\star = \underset{z_1,\ldots,z_n \in \mathbb{S}_d}{\arg\min} \; \mathcal{L}_{\text{Unif}}. \tag{57}$$

Then $\nu_{\{z_1^\star,\ldots,z_n^\star\}}$ converge weak to σ_d as $n \to \infty$.*

Proof. We have

$$\underset{z_1,\ldots,z_n \in \mathbb{S}_d}{\arg\min} \; \mathcal{L}_{\text{Unif}} \tag{58}$$

$$= \underset{z_1,\ldots,z_n \in \mathbb{S}_d}{\arg\min} \; \log \sum_{l,m} \exp(\kappa z_l^\top z_m) \tag{59}$$

$$= \underset{z_1,\ldots,z_n \in \mathbb{S}_d}{\arg\min} \; \sum_{l,m} \exp(\kappa z_l^\top z_m) \tag{60}$$

(monotonicity of logarithm)

$$= \underset{z_1,\ldots,z_n \in \mathbb{S}_d}{\arg\min} \; \sum_{1 \leq l < m \leq n} \exp(\kappa z_l^\top z_m) \tag{61}$$

(symmetry of inner product)

$$= \underset{z_1,\ldots,z_n \in \mathbb{S}_d}{\arg\min} \; \sum_{1 \leq l < m \leq n} \underbrace{\exp(-\kappa ||z_l - z_m||_2^2)}_{=: \, G(z_l, z_m)} \tag{62}$$

(multiplication by positive constant)

$$= \underset{z_1,\ldots,z_n \in \mathbb{S}_d}{\arg\min} \; \sum_{1 \leq l < m \leq n} G(z_l, z_m) \tag{63}$$

The result then follows directly from [8, Proposition 2]. □

5. Experiments

5.1. Implementation details

This section covers the additional implementation details not provided in the main paper. These include the initialization of the embeddings in Algorithm 1, hyperparameters, additional transformations wherever required, the architectures used, and a note on accessing the code, datasets, and dataset splits.

Initialization and normalization. Instead of a random initialization of our embeddings Z_0, we follow a PCA based initialization, as in [5]. The weights are computed using the cached features from the base classes, the support and query features are then transformed using these weights. This procedure is also fast as we do not need to compute the PCA weights on every episode. To ensure that the resulting features lie on the hypersphere after each gradient update in noHub and noHub-S, we re-normalize the embeddings using L2 normalization.

Hyperparameters. noHub and noHub-S have the following hyperparameters.

- P – perplexity for computing the κ_i.

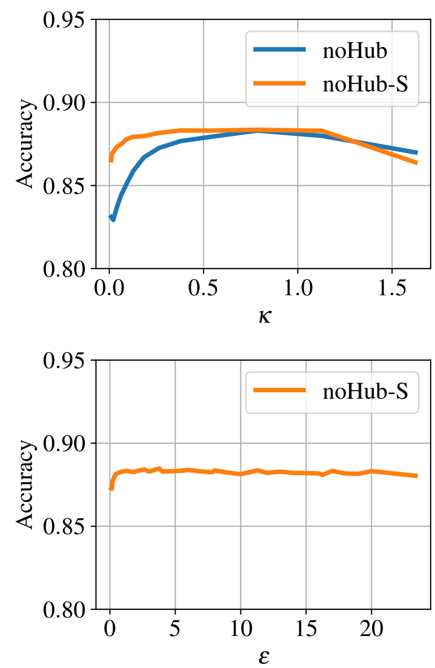

Figure 1. Accuracy for different values for κ and ε. Neither noHub nor noHub-S are particularly sensitive the the choice of these parameters.

- T – number of iterations.
- α – tradeoff parameter in the loss ($\mathcal{L}_{\text{noHub}} = \alpha \mathcal{L}_{\text{LSP}} + (1-\alpha)\mathcal{L}_{\text{Unif}}$).
- η – learning rate for the Adam optimizer.
- κ – concentration parameter for the embeddings.
- ε – exaggeration of similarities between supports from different classes.
- d – dimensionality of embeddings.

All hyperparameter values used in in noHub and noHub-S are given in Table 1

Code. The code for our experiments is available at: https://github.com/uitml/noHub

Data splits. Details to access the datasets used with the requisite splits (both are consistent with [6]) are available in the code repository.

Base feature extractors.

- **Resnet-18**: As in [1, 6], we use the weights from [6]. The model is trained using a cross-entropy loss on the base classes.
- **WideRes28-10**: Following [3, 9], we use the weights from [3]. The model is pre-trained using a combination of cross-entropy and rotation prediction [2], and then fine-tuned with Manifold Mixup [7].

Arch.	Param.	Method	mini 1-shot	mini 5-shot	tiered 1-shot	tiered 5-shot	CUB 1-shot	CUB 5-shot
ResNet18	P	noHub	45	45	45	45	45	45
		noHub-S	45	45	40	45	45	45
	T	noHub	50	50	50	50	50	50
		noHub-S	150	150	150	150	150	150
	α	noHub	0.2	0.2	0.2	0.2	0.2	0.2
		noHub-S	0.3	0.2	0.2	0.2	0.3	0.2
	η	noHub	0.1	0.1	0.1	0.1	0.1	0.1
		noHub-S	0.1	0.1	0.1	0.1	0.1	0.1
	κ	noHub	0.5	0.5	0.5	0.5	0.5	0.5
		noHub-S	0.5	0.5	0.5	0.5	0.5	0.5
	ε	noHub	–	–	–	–	–	–
		noHub-S	8	8	5	8	8	8
	d	noHub	400	400	400	400	400	400
		noHub-S	400	400	400	400	400	400
WideRes28-10	P	noHub	45	45	45	45	45	45
		noHub-S	45	45	40	35	45	30
	T	noHub	50	50	50	50	50	50
		noHub-S	150	150	150	150	150	150
	α	noHub	0.2	0.2	0.2	0.2	0.2	0.2
		noHub-S	0.3	0.2	0.2	0.1	0.3	0.1
	η	noHub	0.1	0.1	0.1	0.1	0.1	0.1
		noHub-S	0.1	0.1	0.1	0.1	0.1	0.1
	κ	noHub	0.5	0.5	0.5	0.5	0.5	0.5
		noHub-S	0.5	0.5	0.5	0.2	0.5	0.2
	ε	noHub	–	–	–	–	–	–
		noHub-S	8	8	5	12	8	8
	d	noHub	400	400	400	400	400	400
		noHub-S	400	400	400	400	400	400

Table 1. Hyperparameter values used in our experiments.

5.2. Results

FSL performance. The complete lists of accuracies and hubness metrics for all embeddings, classifiers, and feature extractors, are given in Tables 2, 3, 4, and 5. The exhaustive results in these tables form the basis of Table 1, Table 2 and Table 3 in the main text. The two proposed approaches consistently outperform prior embeddings across several classifiers, feature extractors and datasets.

Effect of the κ and ε hyperparameters. The plots in Figure 1 show accuracy on *tiered* 5-shot with SIAMESE for increasing κ and ε. Neither method is particularly sensitive to the choice of κ and ε, and noHub-S is less sensitive to variations in κ, than noHub. Choosing $\kappa \in [0.5, 1]$ and $\varepsilon \in [3, 20]$ will result in high classification accuracy

6. Potential Negative Societal Impacts

As is the case with most methodological research in machine learning, the methods developed in this work could be used in downstream applications with potential negative societal impacts. Real world machine learning-based systems that interact with humans, or the environment in general, should therefore be properly tested and equipped with adequate safety measures.

Since our work relies on a large number of labeled examples from the base classes, un-discovered biases from the base dataset could be transferred to the trained models. Furthermore, the small number of examples in the inference stage could make the query predictions biased towards the included support examples, and not accurately reflect the diversity of the novel classes.

Arch.	Clf.	Emb.	mini Acc	mini Skew	mini Hub. Occ.	tiered Acc	tiered Skew	tiered Hub. Occ.	CUB Acc	CUB Skew	CUB Hub. Occ.
ResNet18	ILPC	None	64.07 (0.28)	1.411 (0.01)	0.408 (0.001)	75.5 (0.28)	1.213 (0.009)	0.41 (0.001)	76.06 (0.27)	0.886 (0.006)	0.34 (0.001)
		L2	69.28 (0.27)	0.966 (0.007)	0.298 (0.001)	77.84 (0.28)	0.811 (0.007)	0.267 (0.001)	79.91 (0.26)	0.688 (0.006)	0.236 (0.001)
		CL2	71.48 (0.27)	0.661 (0.005)	0.229 (0.001)	79.8 (0.27)	0.679 (0.006)	0.249 (0.001)	80.97 (0.26)	0.553 (0.005)	0.203 (0.001)
		ZN	71.48 (0.27)	0.677 (0.006)	0.227 (0.001)	79.95 (0.27)	0.694 (0.006)	0.263 (0.001)	81.49 (0.25)	0.57 (0.005)	0.217 (0.001)
		ReRep	65.49 (0.28)	3.688 (0.007)	0.559 (0.001)	76.75 (0.28)	3.61 (0.01)	0.55 (0.001)	77.73 (0.26)	3.563 (0.007)	0.512 (0.001)
		EASE	71.79 (0.28)	0.515 (0.005)	0.157 (0.001)	80.2 (0.27)	0.48 (0.005)	0.158 (0.001)	81.88 (0.25)	0.463 (0.004)	0.153 (0.001)
		TCPR	71.77 (0.28)	0.647 (0.005)	0.223 (0.001)	80.01 (0.28)	0.652 (0.006)	0.249 (0.001)	81.75 (0.25)	0.534 (0.004)	0.203 (0.001)
		noHub	73.18 (0.28)	0.308 (0.005)	**0.094 (0.001)**	80.76 (0.28)	0.296 (0.004)	**0.101 (0.001)**	82.74 (0.26)	0.32 (0.004)	**0.112 (0.001)**
		noHub-S	**74.02 (0.28)**	**0.276 (0.004)**	0.13 (0.001)	**81.34 (0.27)**	**0.281 (0.004)**	0.127 (0.001)	**83.92 (0.25)**	**0.296 (0.003)**	0.163 (0.001)
	LaplacianShot	None	68.92 (0.23)	1.341 (0.009)	0.408 (0.001)	76.43 (0.25)	1.214 (0.009)	0.41 (0.001)	79.17 (0.23)	0.887 (0.006)	0.34 (0.001)
		L2	69.3 (0.23)	0.945 (0.007)	0.302 (0.001)	77.2 (0.25)	0.808 (0.007)	0.265 (0.001)	79.65 (0.23)	0.682 (0.006)	0.236 (0.001)
		CL2	70.68 (0.23)	0.661 (0.005)	0.231 (0.001)	77.98 (0.24)	0.689 (0.006)	0.248 (0.001)	79.99 (0.22)	0.547 (0.005)	0.201 (0.001)
		ZN	70.51 (0.23)	0.688 (0.006)	0.233 (0.001)	77.51 (0.24)	0.697 (0.006)	0.264 (0.001)	79.86 (0.22)	0.564 (0.005)	0.217 (0.001)
		ReRep	72.75 (0.24)	3.653 (0.007)	0.548 (0.001)	78.95 (0.25)	3.605 (0.011)	0.549 (0.001)	82.38 (0.22)	3.565 (0.007)	0.512 (0.001)
		EASE	72.19 (0.23)	0.526 (0.005)	0.161 (0.001)	79.34 (0.24)	0.481 (0.005)	0.158 (0.001)	81.5 (0.22)	0.459 (0.004)	0.152 (0.001)
		TCPR	71.79 (0.23)	0.654 (0.005)	0.228 (0.001)	78.41 (0.24)	0.651 (0.005)	0.249 (0.001)	80.86 (0.22)	0.537 (0.004)	0.203 (0.001)
		noHub	73.63 (0.25)	0.305 (0.005)	**0.094 (0.001)**	**80.84 (0.25)**	0.3 (0.005)	**0.101 (0.001)**	83.23 (0.22)	0.318 (0.004)	**0.112 (0.001)**
		noHub-S	**73.79 (0.25)**	**0.276 (0.004)**	0.13 (0.001)	80.83 (0.25)	**0.275 (0.004)**	0.125 (0.001)	**83.47 (0.22)**	**0.299 (0.003)**	0.164 (0.001)
	ObliqueManifold	None	68.89 (0.23)	1.412 (0.01)	0.407 (0.001)	77.07 (0.25)	1.21 (0.009)	0.409 (0.001)	79.4 (0.22)	0.887 (0.006)	0.341 (0.001)
		L2	68.92 (0.23)	0.964 (0.007)	0.299 (0.001)	77.17 (0.25)	0.806 (0.007)	0.266 (0.001)	79.32 (0.22)	0.691 (0.006)	0.237 (0.001)
		CL2	70.86 (0.24)	0.66 (0.005)	0.228 (0.001)	78.92 (0.25)	0.68 (0.006)	0.249 (0.001)	80.29 (0.23)	0.547 (0.005)	0.202 (0.001)
		ZN	71.25 (0.24)	0.679 (0.006)	0.227 (0.001)	79.54 (0.25)	0.697 (0.006)	0.263 (0.001)	81.38 (0.23)	0.562 (0.005)	0.216 (0.001)
		ReRep	73.3 (0.25)	3.682 (0.007)	0.559 (0.001)	80.26 (0.26)	3.608 (0.01)	0.551 (0.001)	**83.84 (0.23)**	3.559 (0.008)	0.513 (0.001)
		EASE	68.4 (0.24)	0.516 (0.005)	0.156 (0.001)	77.33 (0.25)	0.477 (0.004)	0.158 (0.001)	79.03 (0.24)	0.461 (0.004)	0.152 (0.001)
		TCPR	70.74 (0.24)	0.646 (0.005)	0.223 (0.001)	78.92 (0.25)	0.649 (0.005)	0.249 (0.001)	80.18 (0.23)	0.537 (0.004)	0.204 (0.001)
		noHub	72.55 (0.26)	0.309 (0.005)	**0.095 (0.001)**	79.97 (0.26)	0.302 (0.005)	**0.102 (0.001)**	82.21 (0.24)	0.319 (0.004)	**0.112 (0.001)**
		noHub-S	**74.24 (0.26)**	**0.274 (0.004)**	0.13 (0.001)	**80.84 (0.26)**	**0.282 (0.004)**	0.127 (0.001)	83.67 (0.23)	**0.294 (0.003)**	0.162 (0.001)
	SIAMESE	None	20.0 (0.0)	1.345 (0.009)	0.407 (0.001)	20.0 (0.0)	1.222 (0.009)	0.41 (0.001)	20.0 (0.0)	0.885 (0.006)	0.339 (0.001)
		L2	73.77 (0.24)	0.949 (0.007)	0.301 (0.001)	80.46 (0.26)	0.811 (0.007)	0.265 (0.001)	83.1 (0.23)	0.691 (0.006)	0.237 (0.001)
		CL2	75.56 (0.26)	0.666 (0.005)	0.232 (0.001)	82.1 (0.26)	0.68 (0.006)	0.248 (0.001)	84.35 (0.24)	0.549 (0.005)	0.201 (0.001)
		ZN	20.0 (0.0)	0.686 (0.006)	0.232 (0.001)	20.0 (0.0)	0.69 (0.006)	0.262 (0.001)	20.0 (0.0)	0.565 (0.005)	0.217 (0.001)
		ReRep	20.0 (0.0)	3.653 (0.007)	0.549 (0.001)	20.0 (0.0)	3.616 (0.01)	0.549 (0.001)	20.0 (0.0)	3.559 (0.007)	0.512 (0.001)
		EASE	76.05 (0.27)	0.529 (0.005)	0.162 (0.001)	82.57 (0.27)	0.485 (0.005)	0.159 (0.001)	85.24 (0.24)	0.464 (0.004)	0.153 (0.001)
		TCPR	75.99 (0.26)	0.655 (0.005)	0.227 (0.001)	82.65 (0.26)	0.651 (0.005)	0.249 (0.001)	85.34 (0.23)	0.535 (0.004)	0.203 (0.001)
		noHub	76.65 (0.28)	0.308 (0.005)	**0.095 (0.001)**	82.94 (0.27)	0.303 (0.004)	**0.101 (0.001)**	**85.88 (0.24)**	0.322 (0.004)	**0.112 (0.001)**
		noHub-S	**76.68 (0.28)**	**0.275 (0.004)**	0.13 (0.001)	**83.09 (0.27)**	**0.281 (0.004)**	0.128 (0.001)	85.81 (0.24)	**0.295 (0.003)**	0.161 (0.001)
	SimpleShot	None	56.14 (0.2)	1.349 (0.009)	0.407 (0.001)	63.34 (0.23)	1.211 (0.009)	0.408 (0.001)	64.02 (0.21)	0.887 (0.006)	0.341 (0.001)
		L2	60.15 (0.2)	0.937 (0.007)	0.301 (0.001)	68.02 (0.23)	0.812 (0.007)	0.265 (0.001)	69.05 (0.21)	0.691 (0.006)	0.236 (0.001)
		CL2	63.1 (0.2)	0.667 (0.005)	0.233 (0.001)	69.76 (0.22)	0.679 (0.006)	0.249 (0.001)	70.16 (0.2)	0.549 (0.005)	0.201 (0.001)
		ZN	63.39 (0.2)	0.68 (0.005)	0.231 (0.001)	70.04 (0.22)	0.698 (0.006)	0.264 (0.001)	71.03 (0.2)	0.564 (0.005)	0.216 (0.001)
		ReRep	66.66 (0.22)	3.655 (0.007)	0.548 (0.001)	73.23 (0.23)	3.604 (0.01)	0.549 (0.001)	76.8 (0.21)	3.565 (0.007)	0.513 (0.001)
		EASE	64.0 (0.2)	0.521 (0.005)	0.16 (0.001)	71.0 (0.21)	0.479 (0.005)	0.158 (0.001)	72.38 (0.2)	0.466 (0.004)	0.153 (0.001)
		TCPR	63.33 (0.2)	0.651 (0.005)	0.228 (0.001)	69.82 (0.22)	0.65 (0.005)	0.25 (0.001)	70.75 (0.2)	0.532 (0.004)	0.204 (0.001)
		noHub	69.38 (0.22)	0.315 (0.005)	**0.095 (0.001)**	76.72 (0.23)	0.303 (0.004)	**0.102 (0.001)**	78.21 (0.21)	0.32 (0.004)	**0.112 (0.001)**
		noHub-S	**71.1 (0.22)**	**0.276 (0.004)**	0.13 (0.001)	**78.35 (0.23)**	**0.283 (0.004)**	0.127 (0.001)	**80.31 (0.21)**	**0.296 (0.003)**	0.162 (0.001)
	α-TIM	None	56.39 (0.2)	1.342 (0.009)	0.406 (0.001)	63.32 (0.23)	1.216 (0.009)	0.411 (0.001)	64.02 (0.22)	0.886 (0.006)	0.341 (0.001)
		L2	67.91 (0.23)	0.942 (0.007)	0.301 (0.001)	74.94 (0.24)	0.814 (0.007)	0.266 (0.001)	77.49 (0.23)	0.694 (0.006)	0.236 (0.001)
		CL2	65.68 (0.21)	0.665 (0.005)	0.232 (0.001)	73.23 (0.23)	0.681 (0.006)	0.248 (0.001)	73.79 (0.21)	0.552 (0.005)	0.202 (0.001)
		ZN	63.36 (0.2)	0.682 (0.005)	0.232 (0.001)	70.19 (0.22)	0.693 (0.006)	0.263 (0.001)	70.85 (0.21)	0.566 (0.005)	0.215 (0.001)
		ReRep	66.37 (0.22)	3.656 (0.007)	0.55 (0.001)	73.24 (0.24)	3.605 (0.011)	0.55 (0.001)	76.86 (0.21)	3.555 (0.007)	0.514 (0.001)
		EASE	65.32 (0.2)	0.526 (0.005)	0.163 (0.001)	71.88 (0.22)	0.477 (0.005)	0.158 (0.001)	73.03 (0.21)	0.459 (0.004)	0.151 (0.001)
		TCPR	66.19 (0.21)	0.65 (0.005)	0.227 (0.001)	73.24 (0.23)	0.649 (0.005)	0.25 (0.001)	74.07 (0.21)	0.532 (0.004)	0.203 (0.001)
		noHub	70.08 (0.23)	0.312 (0.005)	**0.094 (0.001)**	77.39 (0.24)	0.304 (0.004)	**0.101 (0.001)**	79.19 (0.22)	0.319 (0.004)	**0.112 (0.001)**
		noHub-S	**72.04 (0.23)**	**0.273 (0.004)**	0.13 (0.001)	**79.13 (0.24)**	**0.282 (0.004)**	0.126 (0.001)	**81.42 (0.22)**	**0.296 (0.003)**	0.161 (0.001)

Table 2. Resnet-18: 1-shot.

Arch.	Clf.	Emb.	mini Acc	mini Skew	mini Hub. Occ.	tiered Acc	tiered Skew	tiered Hub. Occ.	CUB Acc	CUB Skew	CUB Hub. Occ.
WideRes28-10	ILPC	None	71.27 (0.28)	1.595 (0.01)	0.46 (0.001)	75.01 (0.28)	1.807 (0.01)	0.494 (0.001)	89.75 (0.19)	1.072 (0.009)	0.367 (0.001)
		L2	76.41 (0.26)	0.773 (0.006)	0.295 (0.001)	78.25 (0.27)	0.731 (0.006)	0.274 (0.001)	90.27 (0.2)	0.473 (0.004)	0.228 (0.001)
		CL2	74.13 (0.27)	0.993 (0.009)	0.29 (0.001)	78.2 (0.27)	0.815 (0.006)	0.306 (0.001)	90.34 (0.2)	0.524 (0.004)	0.267 (0.001)
		ZN	77.76 (0.26)	0.728 (0.005)	0.287 (0.001)	79.42 (0.27)	0.776 (0.006)	0.302 (0.001)	90.21 (0.2)	0.516 (0.004)	0.263 (0.001)
		ReRep	62.51 (0.34)	3.56 (0.002)	0.704 (0.001)	60.66 (0.37)	3.55 (0.002)	0.776 (0.001)	87.44 (0.25)	3.033 (0.008)	0.472 (0.001)
		EASE	78.01 (0.26)	0.47 (0.004)	0.176 (0.001)	79.64 (0.27)	0.479 (0.004)	0.175 (0.001)	90.76 (0.19)	0.437 (0.003)	0.212 (0.001)
		TCPR	78.37 (0.26)	0.584 (0.005)	0.237 (0.001)	79.55 (0.28)	0.683 (0.006)	0.265 (0.001)	90.77 (0.19)	0.476 (0.004)	0.23 (0.001)
		noHub	78.84 (0.27)	0.293 (0.004)	**0.112 (0.001)**	80.75 (0.28)	0.3 (0.004)	**0.112 (0.001)**	90.91 (0.2)	0.189 (0.004)	**0.109 (0.001)**
		noHub-S	**79.77 (0.26)**	**0.262 (0.004)**	0.148 (0.001)	**81.24 (0.27)**	**0.278 (0.004)**	0.135 (0.001)	**91.28 (0.19)**	**0.16 (0.004)**	0.13 (0.001)
	LaplacianShot	None	72.56 (0.23)	1.599 (0.01)	0.459 (0.001)	75.58 (0.25)	1.795 (0.01)	0.495 (0.001)	88.71 (0.19)	1.071 (0.009)	0.369 (0.001)
		L2	75.18 (0.23)	0.777 (0.006)	0.296 (0.001)	77.03 (0.24)	0.732 (0.006)	0.274 (0.001)	89.73 (0.17)	0.474 (0.004)	0.229 (0.001)
		CL2	71.29 (0.24)	0.987 (0.009)	0.29 (0.001)	75.42 (0.25)	0.819 (0.006)	0.309 (0.001)	89.61 (0.18)	0.52 (0.004)	0.268 (0.001)
		ZN	75.18 (0.22)	0.724 (0.005)	0.286 (0.001)	77.0 (0.24)	0.768 (0.006)	0.301 (0.001)	89.22 (0.18)	0.517 (0.004)	0.263 (0.001)
		ReRep	75.25 (0.22)	3.562 (0.002)	0.704 (0.001)	77.12 (0.24)	3.548 (0.002)	0.776 (0.001)	88.98 (0.18)	3.024 (0.008)	0.47 (0.001)
		EASE	77.29 (0.22)	0.473 (0.004)	0.177 (0.001)	78.97 (0.24)	0.475 (0.004)	0.175 (0.001)	90.06 (0.17)	0.435 (0.003)	0.213 (0.001)
		TCPR	76.77 (0.22)	0.593 (0.005)	0.236 (0.001)	77.49 (0.24)	0.686 (0.006)	0.264 (0.001)	89.42 (0.17)	0.475 (0.004)	0.231 (0.001)
		noHub	**79.13 (0.23)**	0.29 (0.004)	**0.111 (0.001)**	80.5 (0.25)	0.302 (0.004)	**0.112 (0.001)**	**90.73 (0.18)**	0.19 (0.004)	**0.109 (0.001)**
		noHub-S	79.13 (0.23)	**0.259 (0.004)**	0.147 (0.001)	**80.59 (0.24)**	**0.277 (0.004)**	0.135 (0.001)	90.61 (0.17)	**0.164 (0.004)**	0.13 (0.001)
	ObliqueManifold	None	76.02 (0.22)	1.599 (0.01)	0.46 (0.001)	77.75 (0.25)	1.801 (0.01)	0.494 (0.001)	90.82 (0.18)	1.07 (0.009)	0.368 (0.001)
		L2	76.11 (0.22)	0.779 (0.006)	0.295 (0.001)	77.74 (0.25)	0.731 (0.006)	0.274 (0.001)	90.89 (0.18)	0.475 (0.004)	0.228 (0.001)
		CL2	74.43 (0.24)	0.985 (0.009)	0.289 (0.001)	77.98 (0.25)	0.816 (0.007)	0.307 (0.001)	90.6 (0.18)	0.523 (0.004)	0.267 (0.001)
		ZN	77.69 (0.23)	0.724 (0.005)	0.286 (0.001)	79.32 (0.24)	0.767 (0.006)	0.301 (0.001)	90.73 (0.18)	0.519 (0.004)	0.263 (0.001)
		ReRep	78.08 (0.23)	3.56 (0.002)	0.703 (0.001)	79.46 (0.25)	3.549 (0.002)	0.777 (0.001)	91.16 (0.18)	3.032 (0.008)	0.471 (0.001)
		EASE	74.77 (0.23)	0.472 (0.004)	0.178 (0.001)	77.07 (0.25)	0.473 (0.004)	0.174 (0.001)	89.2 (0.18)	0.439 (0.003)	0.212 (0.001)
		TCPR	77.39 (0.23)	0.587 (0.005)	0.236 (0.001)	78.75 (0.24)	0.687 (0.006)	0.265 (0.001)	89.93 (0.19)	0.474 (0.004)	0.23 (0.001)
		noHub	78.44 (0.24)	0.292 (0.004)	**0.112 (0.001)**	79.99 (0.26)	0.302 (0.004)	**0.113 (0.001)**	90.59 (0.19)	0.185 (0.004)	**0.108 (0.001)**
		noHub-S	**79.89 (0.24)**	**0.259 (0.004)**	0.148 (0.001)	**80.67 (0.26)**	**0.279 (0.004)**	0.137 (0.001)	**91.37 (0.18)**	**0.162 (0.004)**	0.13 (0.001)
	SIAMESE	None	45.69 (0.31)	1.594 (0.009)	0.459 (0.001)	75.29 (0.28)	1.801 (0.01)	0.495 (0.001)	61.36 (0.55)	1.074 (0.009)	0.37 (0.001)
		L2	80.2 (0.23)	0.776 (0.006)	0.296 (0.001)	80.89 (0.26)	0.735 (0.006)	0.275 (0.001)	91.98 (0.18)	0.476 (0.004)	0.23 (0.001)
		CL2	75.23 (0.27)	0.988 (0.009)	0.289 (0.001)	79.59 (0.27)	0.82 (0.006)	0.307 (0.001)	92.17 (0.18)	0.518 (0.004)	0.266 (0.001)
		ZN	20.0 (0.0)	0.726 (0.005)	0.286 (0.001)	20.0 (0.0)	0.775 (0.006)	0.302 (0.001)	20.0 (0.0)	0.517 (0.004)	0.264 (0.001)
		ReRep	36.69 (0.28)	3.561 (0.002)	0.705 (0.001)	67.41 (0.29)	3.55 (0.002)	0.776 (0.001)	57.62 (0.56)	3.027 (0.008)	0.472 (0.001)
		EASE	81.19 (0.25)	0.474 (0.004)	0.178 (0.001)	82.04 (0.26)	0.476 (0.004)	0.176 (0.001)	91.99 (0.19)	0.436 (0.003)	0.213 (0.001)
		TCPR	81.27 (0.24)	0.582 (0.005)	0.236 (0.001)	81.89 (0.26)	0.681 (0.006)	0.264 (0.001)	91.91 (0.18)	0.477 (0.004)	0.232 (0.001)
		noHub	81.97 (0.25)	0.291 (0.004)	**0.111 (0.001)**	82.8 (0.27)	0.298 (0.004)	**0.112 (0.001)**	92.53 (0.18)	0.189 (0.004)	**0.109 (0.001)**
		noHub-S	**82.0 (0.26)**	**0.258 (0.004)**	0.148 (0.001)	**82.85 (0.27)**	**0.278 (0.004)**	0.137 (0.001)	**92.63 (0.18)**	**0.159 (0.004)**	0.13 (0.001)
	SimpleShot	None	55.66 (0.21)	1.6 (0.01)	0.459 (0.001)	54.71 (0.22)	1.81 (0.01)	0.494 (0.001)	70.92 (0.23)	1.073 (0.009)	0.369 (0.001)
		L2	65.78 (0.2)	0.781 (0.006)	0.296 (0.001)	68.75 (0.22)	0.737 (0.006)	0.275 (0.001)	82.85 (0.19)	0.475 (0.004)	0.228 (0.001)
		CL2	64.33 (0.2)	0.981 (0.009)	0.288 (0.001)	67.66 (0.22)	0.817 (0.006)	0.307 (0.001)	82.8 (0.19)	0.52 (0.004)	0.267 (0.001)
		ZN	67.31 (0.2)	0.73 (0.005)	0.287 (0.001)	69.14 (0.22)	0.769 (0.006)	0.302 (0.001)	82.79 (0.19)	0.517 (0.004)	0.263 (0.001)
		ReRep	67.38 (0.2)	3.56 (0.002)	0.704 (0.001)	70.17 (0.22)	3.55 (0.002)	0.777 (0.001)	84.86 (0.19)	3.026 (0.008)	0.47 (0.001)
		EASE	68.62 (0.2)	0.47 (0.004)	0.177 (0.001)	70.26 (0.21)	0.477 (0.004)	0.175 (0.001)	84.14 (0.18)	0.437 (0.003)	0.213 (0.001)
		TCPR	68.45 (0.2)	0.589 (0.005)	0.236 (0.001)	68.68 (0.22)	0.685 (0.006)	0.264 (0.001)	82.28 (0.19)	0.477 (0.004)	0.231 (0.001)
		noHub	75.06 (0.21)	0.29 (0.004)	**0.111 (0.001)**	76.7 (0.23)	0.301 (0.004)	**0.111 (0.001)**	88.06 (0.18)	0.188 (0.004)	**0.108 (0.001)**
		noHub-S	**76.86 (0.21)**	**0.258 (0.004)**	0.148 (0.001)	**78.4 (0.23)**	**0.274 (0.004)**	0.135 (0.001)	**89.25 (0.18)**	**0.162 (0.004)**	0.13 (0.001)
	α-TIM	None	60.31 (0.2)	1.603 (0.01)	0.458 (0.001)	69.42 (0.25)	1.811 (0.01)	0.494 (0.001)	73.83 (0.21)	1.072 (0.009)	0.369 (0.001)
		L2	72.11 (0.22)	0.778 (0.006)	0.295 (0.001)	74.45 (0.23)	0.73 (0.006)	0.275 (0.001)	85.96 (0.19)	0.476 (0.004)	0.229 (0.001)
		CL2	68.5 (0.21)	0.988 (0.009)	0.29 (0.001)	72.17 (0.23)	0.811 (0.006)	0.306 (0.001)	85.6 (0.18)	0.522 (0.004)	0.267 (0.001)
		ZN	67.69 (0.2)	0.73 (0.005)	0.287 (0.001)	68.94 (0.22)	0.769 (0.006)	0.302 (0.001)	83.03 (0.19)	0.518 (0.004)	0.263 (0.001)
		ReRep	73.15 (0.23)	3.56 (0.002)	0.704 (0.001)	76.19 (0.25)	3.551 (0.002)	0.778 (0.001)	88.55 (0.18)	3.027 (0.008)	0.472 (0.001)
		EASE	69.83 (0.2)	0.468 (0.004)	0.176 (0.001)	71.54 (0.22)	0.481 (0.004)	0.175 (0.001)	84.9 (0.19)	0.436 (0.003)	0.213 (0.001)
		TCPR	71.6 (0.21)	0.586 (0.005)	0.237 (0.001)	72.71 (0.22)	0.689 (0.006)	0.264 (0.001)	84.99 (0.19)	0.479 (0.004)	0.231 (0.001)
		noHub	75.87 (0.22)	0.29 (0.004)	**0.111 (0.001)**	77.83 (0.23)	0.302 (0.004)	**0.112 (0.001)**	88.7 (0.17)	0.189 (0.004)	**0.108 (0.001)**
		noHub-S	**77.76 (0.22)**	**0.259 (0.004)**	0.147 (0.001)	**79.04 (0.24)**	**0.276 (0.004)**	0.136 (0.001)	**89.77 (0.17)**	**0.163 (0.003)**	0.13 (0.001)

Table 3. WideRes28-10: 1-shot.

Arch.	Clf.	Emb.	mini Acc	mini Skew	mini Hub. Occ.	tiered Acc	tiered Skew	tiered Hub. Occ.	CUB Acc	CUB Skew	CUB Hub. Occ.
ResNet18	ILPC	None	76.46 (0.18)	1.503 (0.01)	0.421 (0.001)	84.46 (0.18)	1.334 (0.008)	0.433 (0.001)	85.86 (0.14)	0.981 (0.005)	0.364 (0.001)
		L2	80.9 (0.16)	1.051 (0.007)	0.314 (0.001)	86.23 (0.17)	0.912 (0.006)	0.289 (0.001)	88.03 (0.13)	0.808 (0.005)	0.264 (0.001)
		CL2	81.64 (0.16)	0.778 (0.005)	0.262 (0.001)	86.88 (0.17)	0.823 (0.006)	0.281 (0.001)	88.44 (0.13)	0.695 (0.005)	0.235 (0.001)
		ZN	81.61 (0.16)	0.793 (0.005)	0.258 (0.001)	86.9 (0.17)	0.841 (0.006)	0.297 (0.001)	88.44 (0.12)	0.717 (0.004)	0.25 (0.001)
		ReRep	74.83 (0.19)	1.623 (0.003)	0.871 (0.001)	83.96 (0.19)	1.722 (0.004)	0.873 (0.001)	84.54 (0.15)	1.432 (0.003)	0.869 (0.001)
		EASE	81.75 (0.16)	0.618 (0.005)	0.182 (0.001)	86.84 (0.17)	0.593 (0.004)	0.181 (0.001)	88.85 (0.12)	0.606 (0.004)	0.186 (0.001)
		TCPR	81.76 (0.16)	0.766 (0.005)	0.254 (0.001)	86.78 (0.17)	0.801 (0.005)	0.284 (0.001)	88.69 (0.13)	0.683 (0.004)	0.237 (0.001)
		noHub	82.09 (0.16)	**0.295 (0.004)**	0.097 (0.001)	86.81 (0.17)	**0.289 (0.004)**	0.102 (0.001)	88.85 (0.13)	**0.333 (0.004)**	0.12 (0.001)
		noHub-S	**82.33 (0.16)**	0.488 (0.006)	**0.086 (0.001)**	**87.05 (0.17)**	0.475 (0.006)	**0.091 (0.001)**	**89.12 (0.13)**	0.438 (0.006)	**0.097 (0.001)**
	LaplacianShot	None	81.97 (0.15)	1.442 (0.009)	0.422 (0.001)	86.17 (0.16)	1.336 (0.008)	0.432 (0.001)	88.58 (0.12)	0.985 (0.005)	0.365 (0.001)
		L2	81.89 (0.14)	1.035 (0.007)	0.319 (0.001)	86.19 (0.16)	0.913 (0.006)	0.289 (0.001)	88.52 (0.11)	0.811 (0.005)	0.264 (0.001)
		CL2	81.93 (0.14)	0.786 (0.005)	0.265 (0.001)	86.16 (0.16)	0.82 (0.006)	0.282 (0.001)	88.46 (0.12)	0.7 (0.005)	0.235 (0.001)
		ZN	82.57 (0.14)	0.803 (0.005)	0.263 (0.001)	86.67 (0.16)	0.838 (0.006)	0.296 (0.001)	88.88 (0.11)	0.714 (0.004)	0.25 (0.001)
		ReRep	82.32 (0.14)	1.633 (0.003)	0.863 (0.001)	86.09 (0.16)	1.721 (0.004)	0.873 (0.001)	88.74 (0.12)	1.431 (0.002)	0.869 (0.001)
		EASE	82.57 (0.14)	0.627 (0.005)	0.186 (0.001)	86.82 (0.15)	0.596 (0.004)	0.182 (0.001)	88.94 (0.11)	0.608 (0.004)	0.185 (0.001)
		TCPR	82.24 (0.14)	0.781 (0.005)	0.259 (0.001)	86.27 (0.16)	0.797 (0.005)	0.284 (0.001)	88.63 (0.11)	0.687 (0.004)	0.236 (0.001)
		noHub	82.55 (0.15)	0.285 (0.004)	0.096 (0.001)	86.75 (0.16)	0.29 (0.004)	0.103 (0.001)	**89.08 (0.11)**	**0.329 (0.004)**	0.12 (0.001)
		noHub-S	**82.81 (0.14)**	**0.25 (0.005)**	**0.073 (0.001)**	**87.12 (0.16)**	**0.214 (0.005)**	**0.077 (0.001)**	88.99 (0.11)	0.438 (0.006)	**0.096 (0.001)**
	ObliqueManifold	None	83.53 (0.15)	1.497 (0.01)	0.421 (0.001)	87.85 (0.15)	1.334 (0.009)	0.433 (0.001)	90.28 (0.11)	0.987 (0.005)	0.364 (0.001)
		L2	83.66 (0.15)	1.051 (0.007)	0.314 (0.001)	87.83 (0.15)	0.922 (0.006)	0.289 (0.001)	90.21 (0.11)	0.81 (0.005)	0.263 (0.001)
		CL2	83.62 (0.15)	0.775 (0.005)	0.261 (0.001)	88.1 (0.15)	0.823 (0.006)	0.281 (0.001)	90.09 (0.11)	0.701 (0.005)	0.236 (0.001)
		ZN	**83.86 (0.15)**	0.795 (0.005)	0.258 (0.001)	**88.47 (0.15)**	0.835 (0.006)	0.296 (0.001)	**90.47 (0.11)**	0.716 (0.004)	0.251 (0.001)
		ReRep	82.44 (0.15)	1.62 (0.003)	0.871 (0.001)	86.85 (0.16)	1.725 (0.004)	0.872 (0.001)	89.83 (0.11)	1.431 (0.003)	0.869 (0.001)
		EASE	82.83 (0.15)	0.628 (0.005)	0.185 (0.001)	87.63 (0.16)	0.597 (0.005)	0.182 (0.001)	89.74 (0.12)	0.609 (0.004)	0.186 (0.001)
		TCPR	83.51 (0.15)	0.766 (0.005)	0.255 (0.001)	88.09 (0.15)	0.795 (0.005)	0.283 (0.001)	90.28 (0.11)	0.687 (0.004)	0.236 (0.001)
		noHub	83.28 (0.15)	**0.287 (0.004)**	0.096 (0.001)	87.58 (0.16)	**0.288 (0.004)**	0.102 (0.001)	89.89 (0.12)	**0.334 (0.004)**	0.121 (0.001)
		noHub-S	83.25 (0.16)	0.487 (0.006)	**0.086 (0.001)**	87.82 (0.16)	0.469 (0.006)	**0.091 (0.001)**	89.38 (0.17)	nan (nan)	**0.097 (0.001)**
	SIAMESE	None	20.0 (0.0)	1.441 (0.009)	0.421 (0.001)	20.0 (0.0)	1.339 (0.009)	0.433 (0.001)	20.0 (0.0)	0.984 (0.005)	0.364 (0.001)
		L2	83.14 (0.14)	1.035 (0.007)	0.319 (0.001)	87.04 (0.16)	0.912 (0.006)	0.288 (0.001)	89.48 (0.12)	0.808 (0.005)	0.264 (0.001)
		CL2	84.04 (0.15)	0.788 (0.005)	0.264 (0.001)	87.9 (0.16)	0.816 (0.006)	0.28 (0.001)	90.14 (0.12)	0.698 (0.005)	0.235 (0.001)
		ZN	20.0 (0.0)	0.8 (0.005)	0.263 (0.001)	20.0 (0.0)	0.84 (0.006)	0.296 (0.001)	20.0 (0.0)	0.713 (0.004)	0.251 (0.001)
		ReRep	20.0 (0.0)	1.633 (0.003)	0.863 (0.001)	20.0 (0.0)	1.724 (0.004)	0.872 (0.001)	20.0 (0.0)	1.428 (0.002)	0.869 (0.001)
		EASE	84.61 (0.15)	0.63 (0.005)	0.187 (0.001)	88.33 (0.16)	0.594 (0.004)	0.182 (0.001)	90.42 (0.12)	0.607 (0.004)	0.185 (0.001)
		TCPR	84.39 (0.15)	0.772 (0.005)	0.259 (0.001)	88.26 (0.16)	0.791 (0.005)	0.283 (0.001)	90.5 (0.11)	0.686 (0.004)	0.235 (0.001)
		noHub	84.05 (0.16)	0.292 (0.004)	0.096 (0.001)	87.87 (0.17)	**0.291 (0.004)**	0.103 (0.001)	90.34 (0.12)	**0.334 (0.004)**	0.12 (0.001)
		noHub-S	**84.67 (0.15)**	**0.247 (0.005)**	**0.074 (0.001)**	**88.43 (0.16)**	0.473 (0.006)	**0.092 (0.001)**	**90.52 (0.12)**	0.443 (0.006)	**0.097 (0.001)**
	SimpleShot	None	78.5 (0.14)	1.436 (0.009)	0.422 (0.001)	83.95 (0.16)	1.339 (0.008)	0.432 (0.001)	85.65 (0.12)	0.987 (0.005)	0.364 (0.001)
		L2	79.89 (0.14)	1.04 (0.007)	0.318 (0.001)	84.5 (0.16)	0.914 (0.006)	0.287 (0.001)	86.46 (0.12)	0.812 (0.005)	0.263 (0.001)
		CL2	80.0 (0.14)	0.786 (0.005)	0.264 (0.001)	84.66 (0.16)	0.821 (0.006)	0.28 (0.001)	86.3 (0.12)	0.698 (0.005)	0.236 (0.001)
		ZN	80.57 (0.14)	0.806 (0.005)	0.264 (0.001)	84.97 (0.16)	0.839 (0.006)	0.296 (0.001)	86.76 (0.12)	0.716 (0.005)	0.25 (0.001)
		ReRep	80.86 (0.14)	1.631 (0.003)	0.863 (0.001)	85.05 (0.16)	1.721 (0.004)	0.872 (0.001)	87.83 (0.12)	1.432 (0.002)	0.869 (0.001)
		EASE	80.13 (0.14)	0.624 (0.005)	0.186 (0.001)	84.74 (0.16)	0.598 (0.004)	0.183 (0.001)	86.76 (0.12)	0.607 (0.004)	0.186 (0.001)
		TCPR	80.15 (0.14)	0.78 (0.005)	0.259 (0.001)	84.86 (0.15)	0.796 (0.005)	0.283 (0.001)	86.8 (0.12)	0.687 (0.004)	0.235 (0.001)
		noHub	**82.13 (0.14)**	0.286 (0.004)	0.096 (0.001)	**86.31 (0.16)**	0.289 (0.004)	0.104 (0.001)	**88.46 (0.11)**	**0.329 (0.004)**	0.12 (0.001)
		noHub-S	81.22 (0.14)	**0.25 (0.005)**	**0.074 (0.001)**	86.22 (0.15)	**0.213 (0.005)**	**0.078 (0.001)**	87.6 (0.12)	0.433 (0.006)	**0.097 (0.001)**
	α-TIM	None	78.51 (0.15)	1.45 (0.009)	0.42 (0.001)	83.86 (0.16)	1.341 (0.009)	0.433 (0.001)	85.7 (0.12)	0.981 (0.005)	0.363 (0.001)
		L2	80.02 (0.16)	1.036 (0.007)	0.318 (0.001)	84.49 (0.18)	0.92 (0.006)	0.288 (0.001)	87.88 (0.13)	0.812 (0.005)	0.264 (0.001)
		CL2	80.46 (0.16)	0.784 (0.005)	0.264 (0.001)	84.86 (0.17)	0.82 (0.006)	0.281 (0.001)	87.53 (0.13)	0.701 (0.005)	0.235 (0.001)
		ZN	80.32 (0.14)	0.802 (0.005)	0.263 (0.001)	84.93 (0.16)	0.834 (0.006)	0.295 (0.001)	86.95 (0.12)	0.715 (0.004)	0.25 (0.001)
		ReRep	81.05 (0.14)	1.63 (0.003)	0.863 (0.001)	85.18 (0.16)	1.718 (0.004)	0.872 (0.001)	87.63 (0.12)	1.43 (0.002)	0.87 (0.001)
		EASE	79.13 (0.15)	0.632 (0.005)	0.188 (0.001)	84.04 (0.17)	0.596 (0.004)	0.181 (0.001)	86.7 (0.13)	0.607 (0.004)	0.186 (0.001)
		TCPR	80.52 (0.16)	0.776 (0.005)	0.259 (0.001)	85.01 (0.17)	0.796 (0.005)	0.283 (0.001)	87.81 (0.13)	0.681 (0.004)	0.234 (0.001)
		noHub	**81.39 (0.15)**	0.29 (0.004)	0.096 (0.001)	86.09 (0.16)	0.292 (0.004)	0.103 (0.001)	**88.16 (0.12)**	**0.336 (0.004)**	0.121 (0.001)
		noHub-S	81.37 (0.15)	**0.253 (0.005)**	**0.074 (0.001)**	**86.14 (0.16)**	**0.219 (0.005)**	**0.078 (0.001)**	87.97 (0.12)	0.437 (0.006)	**0.096 (0.001)**

Table 4. Resnet-18: 5-shot.

Arch.	Clf.	Emb.	mini Acc	mini Skew	mini Hub. Occ.	tiered Acc	tiered Skew	tiered Hub. Occ.	CUB Acc	CUB Skew	CUB Hub. Occ.
WideRes28-10	ILPC	None	81.93 (0.16)	1.717 (0.01)	0.473 (0.001)	84.34 (0.17)	1.927 (0.011)	0.509 (0.001)	93.18 (0.11)	1.164 (0.008)	0.396 (0.001)
		L2	85.74 (0.14)	0.888 (0.005)	0.322 (0.001)	86.26 (0.17)	0.859 (0.005)	0.306 (0.001)	93.77 (0.1)	0.636 (0.004)	0.266 (0.001)
		CL2	83.33 (0.16)	1.12 (0.009)	0.318 (0.001)	85.99 (0.17)	0.957 (0.006)	0.338 (0.001)	93.79 (0.1)	0.703 (0.004)	0.309 (0.001)
		ZN	85.96 (0.14)	0.858 (0.005)	0.32 (0.001)	86.77 (0.16)	0.909 (0.006)	0.335 (0.001)	93.73 (0.1)	0.696 (0.004)	0.305 (0.001)
		ReRep	72.11 (0.27)	1.601 (0.003)	0.819 (0.001)	71.68 (0.3)	1.616 (0.004)	0.845 (0.001)	91.52 (0.13)	1.301 (0.005)	0.548 (0.002)
		EASE	85.89 (0.14)	0.577 (0.004)	0.198 (0.001)	86.83 (0.17)	0.583 (0.004)	0.193 (0.001)	**93.87 (0.1)**	0.576 (0.004)	0.242 (0.001)
		TCPR	86.29 (0.14)	0.715 (0.004)	0.27 (0.001)	86.96 (0.17)	0.819 (0.005)	0.295 (0.001)	93.82 (0.1)	0.634 (0.004)	0.265 (0.001)
		noHub	86.07 (0.15)	**0.295 (0.004)**	0.115 (0.001)	86.75 (0.17)	**0.299 (0.004)**	**0.115 (0.001)**	93.72 (0.1)	**0.2 (0.004)**	**0.101 (0.001)**
		noHub-S	**86.41 (0.14)**	0.499 (0.006)	**0.104 (0.001)**	**87.31 (0.17)**	0.406 (0.005)	0.121 (0.001)	93.79 (0.1)	0.416 (0.005)	0.126 (0.001)
	LaplacianShot	None	85.23 (0.13)	1.711 (0.01)	0.474 (0.001)	86.14 (0.15)	1.921 (0.011)	0.509 (0.001)	92.61 (0.1)	1.164 (0.008)	0.395 (0.001)
		L2	85.9 (0.13)	0.892 (0.006)	0.321 (0.001)	86.47 (0.15)	0.867 (0.006)	0.304 (0.001)	93.17 (0.09)	0.635 (0.004)	0.267 (0.001)
		CL2	82.08 (0.15)	1.112 (0.009)	0.318 (0.001)	84.62 (0.16)	0.954 (0.006)	0.34 (0.001)	93.01 (0.1)	0.702 (0.004)	0.309 (0.001)
		ZN	85.97 (0.13)	0.86 (0.005)	0.319 (0.001)	86.67 (0.15)	0.912 (0.006)	0.335 (0.001)	93.3 (0.1)	0.698 (0.004)	0.305 (0.001)
		ReRep	84.34 (0.14)	1.599 (0.003)	0.819 (0.001)	85.61 (0.16)	1.615 (0.004)	0.845 (0.001)	92.2 (0.1)	1.304 (0.005)	0.549 (0.002)
		EASE	86.24 (0.13)	0.573 (0.004)	0.198 (0.001)	86.74 (0.15)	0.582 (0.004)	0.194 (0.001)	93.31 (0.09)	0.578 (0.004)	0.243 (0.001)
		TCPR	86.16 (0.13)	0.712 (0.004)	0.269 (0.001)	85.72 (0.16)	0.813 (0.005)	0.293 (0.001)	92.99 (0.1)	0.638 (0.004)	0.264 (0.001)
		noHub	**86.25 (0.13)**	**0.292 (0.004)**	0.115 (0.001)	**86.78 (0.16)**	**0.299 (0.004)**	**0.115 (0.001)**	**93.38 (0.09)**	**0.197 (0.004)**	**0.1 (0.001)**
		noHub-S	85.79 (0.13)	0.494 (0.006)	**0.103 (0.001)**	86.44 (0.16)	0.397 (0.005)	0.12 (0.001)	93.36 (0.1)	0.42 (0.005)	0.126 (0.001)
	ObliqueManifold	None	87.46 (0.13)	1.712 (0.01)	0.472 (0.001)	88.16 (0.15)	1.913 (0.01)	0.509 (0.001)	94.75 (0.09)	1.161 (0.008)	0.395 (0.001)
		L2	87.61 (0.13)	0.889 (0.005)	0.321 (0.001)	88.14 (0.15)	0.862 (0.006)	0.306 (0.001)	**94.8 (0.09)**	0.642 (0.004)	0.268 (0.001)
		CL2	86.03 (0.14)	1.112 (0.009)	0.317 (0.001)	87.64 (0.16)	0.949 (0.006)	0.338 (0.001)	94.67 (0.09)	0.703 (0.004)	0.31 (0.001)
		ZN	87.88 (0.13)	0.852 (0.005)	0.32 (0.001)	**88.43 (0.15)**	0.908 (0.006)	0.335 (0.001)	94.77 (0.08)	0.697 (0.004)	0.306 (0.001)
		ReRep	87.62 (0.12)	1.599 (0.003)	0.819 (0.001)	88.15 (0.15)	1.616 (0.004)	0.845 (0.001)	94.48 (0.09)	1.302 (0.005)	0.547 (0.002)
		EASE	86.75 (0.13)	0.573 (0.004)	0.198 (0.001)	87.78 (0.15)	0.583 (0.004)	0.193 (0.001)	94.16 (0.09)	0.57 (0.004)	0.24 (0.001)
		TCPR	**87.94 (0.12)**	0.718 (0.004)	0.271 (0.001)	88.15 (0.15)	0.816 (0.005)	0.294 (0.001)	94.47 (0.09)	0.635 (0.004)	0.265 (0.001)
		noHub	87.23 (0.13)	**0.297 (0.004)**	0.115 (0.001)	87.95 (0.16)	**0.296 (0.004)**	**0.114 (0.001)**	94.13 (0.09)	**0.197 (0.004)**	**0.1 (0.001)**
		noHub-S	87.13 (0.14)	0.495 (0.006)	**0.103 (0.001)**	87.84 (0.16)	0.399 (0.005)	0.12 (0.001)	94.06 (0.09)	0.421 (0.005)	0.126 (0.001)
	SIAMESE	None	58.82 (0.31)	1.722 (0.01)	0.473 (0.001)	82.56 (0.22)	1.93 (0.01)	0.511 (0.001)	82.22 (0.37)	1.154 (0.008)	0.396 (0.001)
		L2	87.11 (0.13)	0.894 (0.006)	0.321 (0.001)	87.34 (0.15)	0.861 (0.005)	0.305 (0.001)	94.15 (0.1)	0.638 (0.004)	0.266 (0.001)
		CL2	83.99 (0.16)	1.107 (0.009)	0.318 (0.001)	86.71 (0.16)	0.953 (0.006)	0.339 (0.001)	94.48 (0.09)	0.704 (0.004)	0.31 (0.001)
		ZN	20.0 (0.0)	0.856 (0.005)	0.319 (0.001)	20.0 (0.0)	0.913 (0.006)	0.334 (0.001)	20.0 (0.0)	0.702 (0.004)	0.305 (0.001)
		ReRep	36.41 (0.3)	1.597 (0.003)	0.818 (0.001)	76.49 (0.24)	1.613 (0.004)	0.846 (0.001)	60.36 (0.6)	1.299 (0.005)	0.547 (0.002)
		EASE	87.82 (0.13)	0.579 (0.004)	0.199 (0.001)	88.06 (0.16)	0.586 (0.004)	0.192 (0.001)	94.36 (0.09)	0.571 (0.004)	0.241 (0.001)
		TCPR	87.8 (0.13)	0.717 (0.004)	0.27 (0.001)	87.95 (0.16)	0.822 (0.005)	0.295 (0.001)	94.25 (0.1)	0.637 (0.004)	0.266 (0.001)
		noHub	87.78 (0.14)	**0.29 (0.004)**	0.114 (0.001)	87.99 (0.17)	**0.297 (0.004)**	**0.115 (0.001)**	94.56 (0.09)	**0.196 (0.004)**	**0.1 (0.001)**
		noHub-S	**88.03 (0.13)**	0.492 (0.006)	**0.103 (0.001)**	**88.31 (0.16)**	0.398 (0.005)	0.12 (0.001)	**94.69 (0.09)**	0.416 (0.005)	0.127 (0.001)
	SimpleShot	None	78.56 (0.14)	1.709 (0.01)	0.473 (0.001)	80.32 (0.16)	1.937 (0.01)	0.51 (0.001)	89.27 (0.11)	1.16 (0.008)	0.395 (0.001)
		L2	83.81 (0.13)	0.887 (0.005)	0.322 (0.001)	84.82 (0.15)	0.86 (0.006)	0.305 (0.001)	92.06 (0.1)	0.632 (0.004)	0.266 (0.001)
		CL2	81.05 (0.14)	1.12 (0.009)	0.318 (0.001)	83.82 (0.16)	0.956 (0.006)	0.337 (0.001)	92.19 (0.1)	0.701 (0.004)	0.31 (0.001)
		ZN	83.92 (0.13)	0.858 (0.005)	0.32 (0.001)	85.1 (0.15)	0.912 (0.006)	0.335 (0.001)	92.17 (0.1)	0.699 (0.004)	0.305 (0.001)
		ReRep	79.26 (0.16)	1.597 (0.003)	0.819 (0.001)	82.7 (0.16)	1.617 (0.004)	0.846 (0.001)	91.48 (0.11)	1.299 (0.005)	0.549 (0.002)
		EASE	83.65 (0.13)	0.579 (0.004)	0.199 (0.001)	84.47 (0.15)	0.585 (0.004)	0.193 (0.001)	92.01 (0.1)	0.572 (0.004)	0.241 (0.001)
		TCPR	83.77 (0.13)	0.717 (0.004)	0.27 (0.001)	84.81 (0.15)	0.815 (0.005)	0.294 (0.001)	91.84 (0.1)	0.634 (0.004)	0.264 (0.001)
		noHub	**85.73 (0.13)**	**0.294 (0.004)**	0.115 (0.001)	**86.58 (0.15)**	**0.298 (0.004)**	**0.115 (0.001)**	93.21 (0.09)	**0.195 (0.004)**	**0.1 (0.001)**
		noHub-S	84.39 (0.13)	0.494 (0.006)	**0.103 (0.001)**	86.38 (0.15)	0.407 (0.005)	0.12 (0.001)	**93.39 (0.09)**	0.421 (0.005)	0.127 (0.001)
	α-TIM	None	80.61 (0.15)	1.711 (0.01)	0.473 (0.001)	83.05 (0.18)	1.928 (0.01)	0.51 (0.001)	84.89 (0.29)	1.153 (0.008)	0.396 (0.001)
		L2	83.71 (0.16)	0.892 (0.005)	0.323 (0.001)	84.69 (0.18)	0.863 (0.005)	0.304 (0.001)	92.88 (0.1)	0.633 (0.004)	0.266 (0.001)
		CL2	82.35 (0.16)	1.111 (0.009)	0.318 (0.001)	84.06 (0.18)	0.949 (0.006)	0.339 (0.001)	92.81 (0.1)	0.7 (0.004)	0.31 (0.001)
		ZN	83.93 (0.13)	0.857 (0.005)	0.321 (0.001)	85.07 (0.15)	0.912 (0.006)	0.336 (0.001)	92.15 (0.1)	0.698 (0.004)	0.306 (0.001)
		ReRep	83.4 (0.14)	1.596 (0.003)	0.82 (0.001)	84.4 (0.16)	1.615 (0.004)	0.845 (0.001)	93.19 (0.09)	1.302 (0.005)	0.547 (0.002)
		EASE	82.72 (0.14)	0.576 (0.004)	0.2 (0.001)	83.86 (0.16)	0.583 (0.004)	0.193 (0.001)	92.31 (0.1)	0.572 (0.004)	0.242 (0.001)
		TCPR	84.21 (0.15)	0.718 (0.004)	0.27 (0.001)	84.63 (0.18)	0.814 (0.005)	0.293 (0.001)	92.44 (0.1)	0.635 (0.004)	0.265 (0.001)
		noHub	**85.56 (0.13)**	**0.293 (0.004)**	0.115 (0.001)	**86.37 (0.16)**	**0.3 (0.004)**	**0.115 (0.001)**	92.89 (0.1)	**0.193 (0.004)**	**0.099 (0.001)**
		noHub-S	83.96 (0.15)	0.496 (0.006)	**0.102 (0.001)**	86.01 (0.16)	0.395 (0.005)	0.12 (0.001)	**93.24 (0.1)**	0.422 (0.005)	0.126 (0.001)

Table 5. WideRes28-10: 5-shot.

References

[1] Malik Boudiaf, Ziko Imtiaz Masud, Jérôme Rony, José Dolz, Pablo Piantanida, and Ismail Ben Ayed. Transductive Information Maximization For Few-Shot Learning. In *NeurIPS*, 2020. 4

[2] Spyros Gidaris, Praveer Singh, and Nikos Komodakis. Unsupervised Representation Learning by Predicting Image Rotations. In *ICLR*, 2018. 4

[3] Puneet Mangla, Mayank Singh, Abhishek Sinha, Nupur Kumari, Vineeth N Balasubramanian, and Balaji Krishnamurthy. Charting the Right Manifold: Manifold Mixup for Few-shot Learning. In *WACV*, 2020. 4

[4] Jose C Principe. *Information theoretic learning: Renyi's entropy and kernel perspectives*. Springer Science & Business Media, 2010. 3

[5] Laurens van der Maaten and Geoffrey E. Hinton. Visualizing data using t-sne. *JMLR*, 2008. 2, 4

[6] Olivier Veilleux, Malik Boudiaf, Pablo Piantanida, and Ismail Ben Ayed. Realistic evaluation of transductive few-shot learning. In *NeurIPS*, 2021. 4

[7] Vikas Verma, Alex Lamb, Christopher Beckham, Amir Najafi, Ioannis Mitliagkas, David Lopez-Paz, and Yoshua Bengio. Manifold Mixup: Better Representations by Interpolating Hidden States. In *ICML*, 2019. 4

[8] Tongzhou Wang and Phillip Isola. Understanding Contrastive Representation Learning through Alignment and Uniformity on the Hypersphere. In *ICML*, 2020. 3, 4

[9] Hao Zhu and Piotr Koniusz. EASE: Unsupervised Discriminant Subspace Learning for Transductive Few-Shot Learning. In *CVPR*, 2022. 4