

---

# Hyperspherical Prototype Networks

Supplementary Materials

---

Anonymous Author(s)

Affiliation

Address

email

## 1 Evaluating hyperspherical prototypes on DenseNet-121

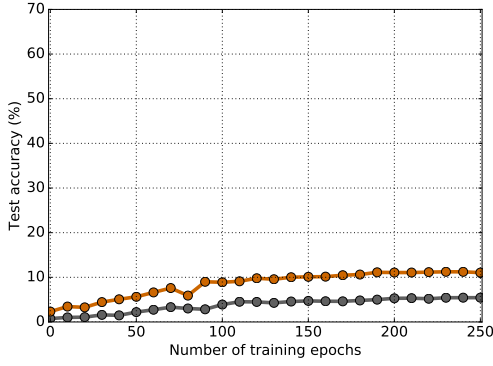
2 Table 1 provides an overview of the evaluation of hyperspherical prototypes with a DenseNet-121  
3 network architecture. The setup is identical to the setup of the first experiment of the main paper,  
4 where ResNet-32 was used as architecture. We reach the same conclusions as for the experiment  
5 using ResNet-32; our hyperspherical prototypes enable a large reduction in output dimensionality  
6 while maintaining performance, outperforming the baseline prototypes on both datasets.

<i>Dimensions</i>	<b>CIFAR-100</b>			<b>ImageNet-200</b>		
	10	25	100	25	50	200
One-hot	-	-	72.2	-	-	60.4
Word2vec	36.6	61.0	71.9	35.9	46.0	56.6
This paper	<b>60.5</b>	<b>69.1</b>	<b>73.4</b>	<b>56.0</b>	<b>60.6</b>	<b>62.6</b>

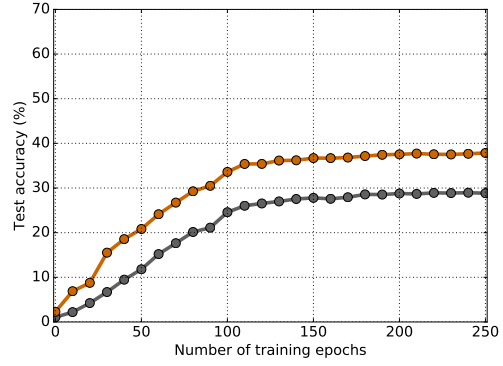
Table 1: Accuracy (%) of our hyperspherical prototypes compared to baseline prototypes using DenseNet-121. Akin to the experiments on ResNet-32, hyperspherical prototypes can handle any output dimensionality, unlike one-hot encodings, and obtain the best scores across datasets.

## 7 2 Accuracy per epoch for prototypes with privileged information

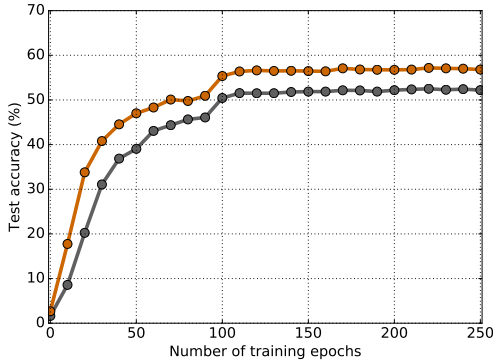
8 In Figure 1, we show the classification accuracy on CIFAR-100 using ResNet-32 as a function of the  
9 number of training epochs. We compare our approach both with and without the use of privileged  
10 information about class semantics, using 3, 5, 10, and 25 output dimensions. The Figure shows that  
11 the use of privileged information not only leads to better results, but also faster convergence. We  
12 attribute this to the semantic structure of the output space, which leads to a smoother optimization.  
13 For larger output spaces, the importance of privileged information is less vital, since there is more  
14 space for class prototypes to be separated from every other class, regardless of semantic structure.



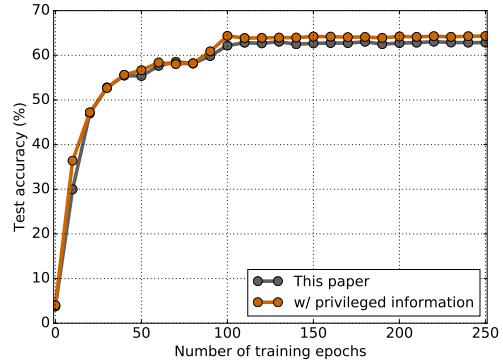
(a) 3 dimensions.



(b) 5 dimensions.



(c) 10 dimensions.



(d) 25 dimensions.

Figure 1: Classification accuracy (%) on CIFAR-100 using ResNet-32 for our approach with and without privileged information. We show the performance interpolated based on scores for every 10 epochs. Incorporating privileged information results in faster convergence and better overall results, especially in small output spaces.