

Conventional metric learning methods usually assume that the training and test samples are captured in similar scenarios so that their distributions are assumed to be the same. This assumption does not hold in many real visual recognition applications, especially when samples are captured across different data sets. In this paper, we propose a new deep transfer metric learning (DTML) method to learn a set of hierarchical nonlinear transformations for cross-domain visual recognition by transferring discriminative knowledge from the labeled source domain to the unlabeled target domain. Specifically, our DTML learns a deep metric network by maximizing the inter-class variations and minimizing the intra-class variations, and minimizing the distribution divergence between the source domain and the target domain at the top layer of the network. To better exploit the discriminative information from the source domain, we further develop a deeply supervised transfer metric learning (DSTML) method by including an additional objective on DTML, where the output of both the hidden layers and the top layer are optimized jointly. To preserve the local manifold of input data points in the metric space, we present two new methods, DTML with autoencoder regularization and DSTML with autoencoder regularization. Experimental results on face verification, person re-identification, and handwritten digit recognition validate the effectiveness of the proposed methods.