

Supplementary Material: Self-supervised Detransformation Autoencoder for Representation Learning in Open Set Recognition

Jingyun Jia^[0000–0003–0865–049X] and Philip K. Chan^[0000–0002–3878–4205]

Florida Institute of Technology, Melbourne FL 32901, USA
jiaj2018@my.fit.edu pkc@fit.edu

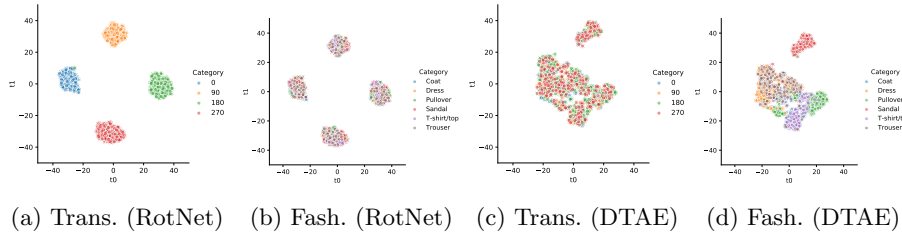


Fig. 1: The t-SNE plots of the learned representations for the Fashion-MNIST training set in the pre-training stage: (a): the representations on transformation (Trans.) classes learned by RotNet; (b): the representations on fashion (Fash.) classes learned by RotNet; similarly for DTAE in (c) and (d).

To understand what kind of information is in the learned representations after pre-training, we analyze how the representations are associated with the transformation classes and fashion (target) classes. We color the representations of the Fashion-MNIST training set based on transformation and fashion classes separately in Figure 1. If the representations have much information for the transformation classes, representations in the same transformation class (color) will be clustered together; otherwise, they will be dispersed. Similarly, we can use the same technique to understand the amount of fashion information in the learned representations. By comparing Figures 1a and 1b, we observe that RotNet learns more transformation than fashion information in the representations. However, from Figures 1c and 1d, we find that DTAE learns more fashion than transformation information in the representations. Therefore, RotNet is more biased towards the transformation information, while DTAE is more biased towards the fashion information, which is more desirable in our case. Moreover, from Figures 1a and 1c, we observe that the representations learned by DTAE contain much less transformation information than those learned by RotNet. However, from Figures 1b and 1d, we find that the representations learned by DTAE contain more fashion (target) information than those learned by Rot-

Net. In summary, DTAE is not only more biased towards fashion (target) than transformation information, and it also learns more fashion information and less transformation information than RotNet.