essential for the existence of the non-collapsed solution. neutral

Tian, Chen, and Ganguli (2021) showed that a regularizer is

In this paper, we make a first attempt towards the second question, by studying a family of algorithms named DirectSet, in which the DirectPred

Although we tried to train split network with the same training data we used, we failed to reproduce their results and used the model trained by the authors [32].