

Supplementary Material

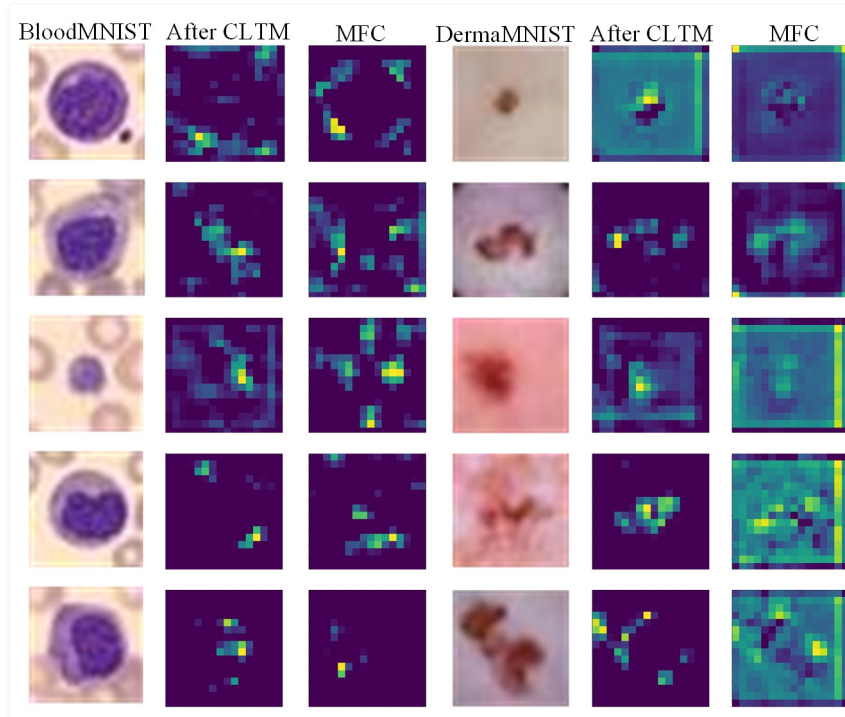


Fig (s1). Visual comparison of hotmaps of data after MFC and CLTM.

In Fig (s1), we can clearly see that after processing by CLTM, the thermal positions of hotmaps are more concentrated in the feature points of the original image, indicating that CLTM obtains more causal associations in the training process, and uses such associations to enhance the attention of features.