Supposedenotes a set of samples with dimension .

Suppose denotes a multi-view dataset with views, and is a collection of data samples with feature dimension in -th view, where

Construct gragh:

,

where denotes the nearest instance set to .

Construct G:

,

GCN:

------- mth encoder layer

-------1th

where is the th view’s relevance coeffient matrix with added self-connection. is the identity matrix, and is the trainable parameter of the mth aotu-encoder layer and denotes the activation function.

is the th view’s final output that undergoes M-layer auto coding.

Weighted fusion:

where and are the consensus representation and view-specific representation generated by the encoder network corresponding to the th view of the th sample, respectively.

Reconstruct :

,m=1,2...M-1

KL:

