

method	I2I	T2I	K2I
V-full	33.68	–	–
V	41.65	–	–
CCA (V+T)	54.90	64.07	95.60
CCA (V+K)	61.77	–	92.60
CCA (V+T+K)	62.56	78.88	97.20
CCA (V+C)	61.69	–	–
CCA (V+T+C)	62.44	75.92	97.80
Structural learning	57.77	–	–
Wsabie	57.15	–	–

) retrieval. The protocols

for I2I and T2I are described in Section ?? . For K2I, each of the 10 ground truth classes is used as a query once. The evaluation metric is average precision (**V-full** refers to the concatenated 38,512-dimensional visual features. In all the other approaches, **V** refers to the 4,500-dimensional PCA-reduced features, **T** to the 500-dimensional sparse SVD-reduced tag features, and **C** is computed based on 20 NC clusters. **Structural learning** refers to the method of and **Wsabie** refers to the method of . We have obtained standard deviations from five random database/query splits, and they are around 0.25