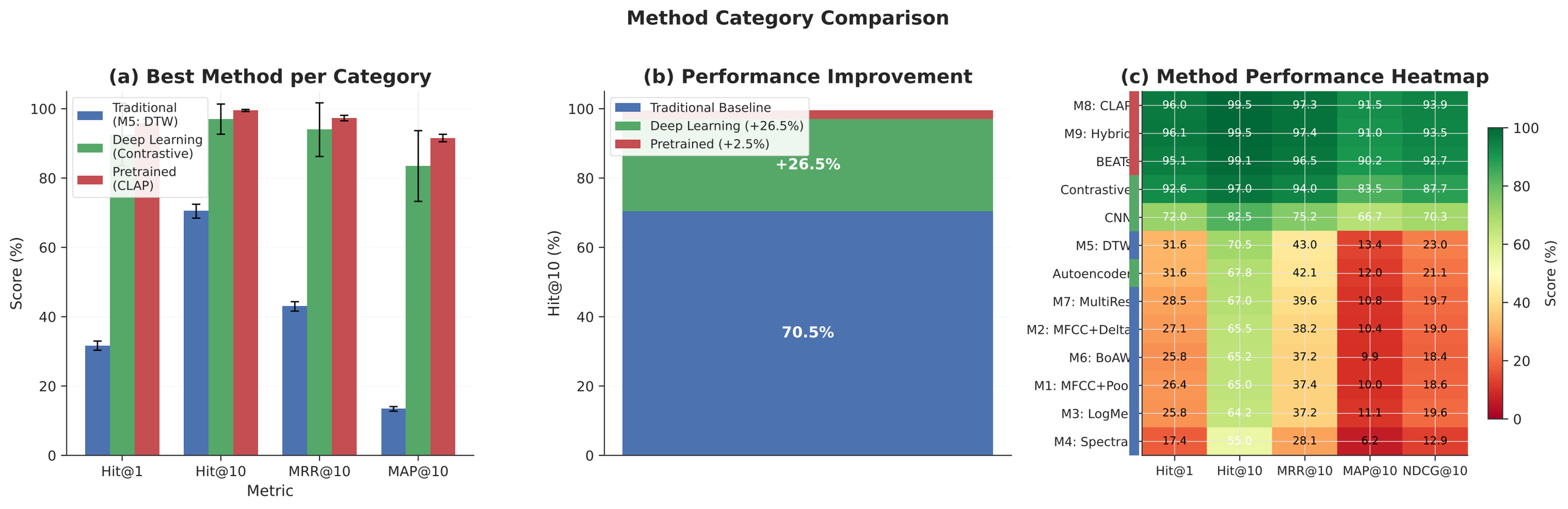
Method Category Comparison
(a) Best Method per Category
Traditional (M5: DTW)
Deep Learning (Contrastive)
Pretrained (CLAP)
Score (%)
Hit@1
Hit@10
MRR@10
MAP@10
Metric
(b) Performance Improvement
Traditional Baseline
Deep Learning (+26.5%)
Pretrained (+2.5%)
+26.5%
70.5%
Hit@10 (%)
(c) Method Performance Heatmap
M8: CLAP 96.0 99.5 97.3 91.5 93.9
M9: Hybrid 96.1 99.5 97.4 91.0 93.5
BEATs 95.1 99.1 96.5 90.2 92.7
Contrastive 92.6 97.0 94.0 83.5 87.7
CNN 72.0 82.5 75.2 66.7 70.3
M5: DTW 31.6 70.5 43.0 13.4 23.0
Autoencoder 31.6 67.8 42.1 12.0 21.1
M7: MultiRes 28.5 67.0 39.6 10.8 19.7
M2: MFCC+Delta 27.1 65.5 38.2 10.4 19.0
M6: BoAW 25.8 65.2 37.2 9.9 18.4
M1: MFCC+Poo 26.4 65.0 37.4 10.0 18.6
M3: LogMe 25.8 64.2 37.2 11.1 19.6
M4: Spectra 17.4 55.0 28.1 6.2 12.9
Hit@1
Hit@10
MRR@10
MAP@10
NDCG@10
Score (%)