

Architecture	Architecture details	Params. [$\times 10^6$]	Av. GAP	Time (Hours)
ROI FC AV Fusion	FC Layers: 12K-12K-12K, DO: 0.3 (see Sec. ??)	502 (see Sec. ??)	81.84 81.8	60 48
Gated Resnet 8K	3 Resnet Layers: 8K-8K-8K-Gated, DO 0.4	238	82.89	48
Gated Resnet 10K	2 (2-conv) Resnet Layers: 10K-10K-Gated, DO 0.3	238	82.8	48
Gated Resnet 10K	2 (2-conv) Resnet Layers: 10K-10K-Gated, DO 0.4	238	82.8	48
Gated FC 10K	3 FC Layers: 10K-10K-10K-Gated, DO: 0.3	358	82.49	48
Gated FC 10K	3 FC Layers: 10K-10K-10K-Gated, DO: 0.4	358	82.77	48
Gated FC 12K	3 FC Layers: 12K-12K-12K-Gated, DO: 0.3	502	82.41	60
Gated FC 12K	3 FC Layers: 12K-12K-12K-Gated, DO: 0.4	502	82.72	60
LSTM	2 layers, 1024 cells	31	81.88	160
GRU	2 layers, 1200 cells	34	82.15	199
Gated BOW	4096 Clusters	39	81.97	348
Non-Gated BOW	8000 Clusters	17	81.88	178
Gated netVLAD	256 Clusters	308	82.89	304
Gated netRVLAD	256 Clusters	307	82.95	140
Gated netFV	128 Clusters	308	82.37	131