); (b) Different combinations of scale levels – in particular, "level1" corresponds to the global CNN representation and "level1+level2+level3" corresponds to the proposed MOP-CNN method. (c) Published numbers for state-of-the-art methods.

	method	feature dimension	accuracy
(a)	Avg (Multi-Scale)	4,096	39.62
	Avg (Concatenation)	12,288	47.50
	Max (Multi-Scale)	4,096	43.51
	Max (Concatenation)	12,288	48.50
	VLAD (Multi-Scale)	4,096	47.32
(b)	level1	4,096	39.57
	level2	4,096	45.34
	level3	4,096	40.21
	level1 + level2	8,192	49.91
	level1 + level3	8,192	49.52
	level2 + level3	8,192	49.66
	level1 + level2 + level3 (MOP-CNN)	12,288	**51.98**
(c)	Xiao et al.	–	38.00
	DeCAF	4,096	40.94
	FV (SIFT + Local Color Statistic)	256,000	47.20