Table 2: Datasets for action recognition

Dataset	Source	No. of Videos	Video Duration	Training Protocol	No. of Classes	Videos/Class	Example Classes
KTH (Schuldt et al., 2004)	Recoded videos on both outdoors and indoors	600	4s	Training and Testing are divided on subjects	6	-	Walk, Jog, Run
Weizmann (Blank et al., 2005)	Outdoor video recordings on still backgrounds	90		Leave out one cross validation	10	-	Walk, Jump, Jumping Jack, Skip
UCF-Sports (Rodriguez et al., 2008)	Television sports broadcasts(eg. BBC, ESPN) (780x480)	150	6.39s	Classification accuracy on provided train test splits by Tian Yan, Discriminative figure-centric models for	10	6 - 22	Diving, Golf-swing, Kicking
Hollywood2 (Marszalek et al., 2009)	Clips from 69 Hollywood movies (33 training and 36 testing) annotated based on movie script	1707		mAP of each class(884 Test videos and 823 training videos obtained from separate training and testing movies)	12	20 - 140	Answer-phone, Eat, Handshake
Olympic Sports (Niebles et al., 2010)	Youtube Video Sequences			mAP of each class on provided train-test splits	16	50	High-jump, Long-jump, Tripple-jump
HMDB-51 (Kuehne et al., 2013)	Youtube, Movies	7000	2 - 3s	Classification accuracy of 30 test clips with training on 70 clips (3 splits are provided)	51	Over 101	Brush-hair, Kick, Kiss
UCF-50 (Reddy and Shah, 2013)	Youtube Video Sequences	-	-	Leave out one cross validation	50	-	Rowing, Fencing, Punch
UCF-101 (Soomro et al., 2012)	Youtube Video Sequences	13320	2 - 5 s	Classification accuracy on 3 train and test spits	101	Over 100	Diving, Skiing, Apply Eye Makeup
Sports 1-M (Karpathy et al., 2014)	Youtube Sports Videos annotated automatically from YouTube topics	1133158		70% of as training while testing and validation sets are respectively 20% and 10%.	487	1000 - 3000	Cricket, disc golf, gliding