Table 3: Accuracy of action recognition techniques (Numbers are true recognition accuracy given in percentages. * Datasets in which the mean average precision is reported). The column *Type* indicates whether a method is purely Deep-net based(D), Representation Based(R) or Fused Solution(F).

Reported Paper	Method	Type	Dataset						
			HMDB51	UCF101	UCF50	UCF-Sports*	Holly-wood2*	Olympic Sports*	Sports1M
Wang et al. (2011)	Dense Traj (Traj + HOG+HOF+MBH)	R				88.2	58.3		
Kliper-Gross et al. (2012)	Motion Interchange Patterns	R	29.2		68.5				
Sadanand and Corso (2012)	General Video Wise Group Wise	R	26.9		 76.4 57.9				
Oneata et al. (2013)	MBH + SIFT + Sqrt + L2 Normalization	R	54.8		90		63.3	82.1	
Wang and Schmid (2013)	Without Human Detector With Human Detector	R	55.9 57.2		90.5 91.2		63 64.3	90.2 91.1	
Jain et al. (2013)	Traj + HoG + HoF + MBH + DCS on w-flow	R	52.1				62.5		
Peng et al. (2014b)	Stacked FVs + FV	R	66.8						
Peng et al. (2014a)	Hybrid-BoW	R	61.1	87.9	92.3				
Kantorov and Laptev (2014)	MPEG-Flow : VLAD encodings of	R	46.3						
Gaidon et al. (2014)	SDT tree ATEP	R	41.3				54.4	85.5	
Simonyan and Zisserman (2014)	Two-stream (CNN-M-2048)	D	59.4	88.0					
Karpathy et al. (2014)	Transfer Learning on Sports 1M Clip Hit @ 1 - Slow Fusion Video Hit @ 1 - Slow Fusion	D		65.4					 41.9 60.9
Sun et al. (2015)	Factorized Spatio Temporal Conv. Nets	D	59.1	88.1					
Wang et al. (2015b)	Two-Stream (ClarifaiNet) Two-Stream (GoogLeNet) Two-Stream (VGG-16)	D		88.0 89. 3 91.4					
Wang et al. (2015a)	TDD + Wang and Schmid (2013) TDD (Only)	F F	65.9 63.2	91.5 90.3					
Ng et al. (2015)	Conv Pooling Hit@1 (Best) LSTM Hit@1 (Best) Conv Pooling (Image + Opt Flow) LSTM (Image + Opt Flow)	D		 88.2 88.6					72.4 73.1
Fernando et al. (2015)	Rank Pooling	R	63.7				73.7		
Donahue et al. (2015)	LRCN- Weighted Avnerage of RBG + Flow	R		82.9					
Wu et al. (2015)	Adaptive Multi-Stream Fusion	D		92.6					
Jiang et al. (2015)	TrajShape+TrajMF TrajShape+TrajMF+ Wang and Schmid (2013)	R	48.4 57.3	78.5 87.2			55.2 65.4	80.6 91	
Lan et al. (2015)	Multi-Skip Feat. Stacking	R	65.1	89.1	94.4		68.0	91.4	
Hoai and Zisserman (2015)	Proposed SSD + RCS	R	62.2				72.7		
Tran et al. (2015)	C3D on SVM C3D + Wang and Schmid (2013) on SVM	D F		85.2 90.4					
Misra et al. (2016)	ImageNet pretrain + tuple verification HMDB + UCF101 Labels Only	D	29.9 30.6						
Wang et al. (2016)	Proposed Only (RBG + Opt Flow Networks)	D	62	92.4					
Fernando and Gould (2016)	End to End Rank-pooling	D				87	40.6		
Fernando et al. (2016)	Hierarchical Rank-pooling (CNN Features) Hierarchical RP on CNN+ Fernando et al. (2015)	D F	47.5 65.0	78.8 90.7			56.8 74.1		
Li et al. (2016)	VLAD^3 VLAD^3 + Wang and Schmid (2013)	F F		84.7 92.2				90.8 96.6	
Varol et al. (2016)	LTC$_{flow+RGB}$ LTC$_{flow+RGB}$ + Wang and Schmid (2013)	D F	64.8 67.2	91.7 92.7					
Feichtenhofer et al. (2016)	Two Stream Fusion (VGG-16) Two Stream Fusion (VGG-16) + Wang and Schmid (2013)	D F	65.4 69.2	92.5 93.5					
de Souza et al. (2016)	Hybrid fusion of Wang and Schmid (2013) + Deep-nets	F	70.4	92.5			72.6		