Juan Carlos Niebles

List of Publications by Year
in descending order

67 papers	6,799 citations	430442 18 h-index	642321 23 g-index
68 all docs	68 docs citations	68 times ranked	3608 citing authors

#	Article	IF	Citations
1	Unsupervised Learning of Human Action Categories Using Spatial-Temporal Words. International Journal of Computer Vision, 2008, 79, 299-318.	10.9	1,227
2	ActivityNet: A large-scale video benchmark for human activity understanding. , 2015, , .		1,148
3	Dense-Captioning Events in Videos. , 2017, , .		595
4	Modeling Temporal Structure of Decomposable Motion Segments for Activity Classification. Lecture Notes in Computer Science, 2010, , 392-405.	1.0	340
5	A Hierarchical Model of Shape and Appearance for Human Action Classification. , 2007, , .		316
6	SST: Single-Stream Temporal Action Proposals. , 2017, , .		286
7	Vision-based action recognition of earthmoving equipment using spatio-temporal features and support vector machine classifiers. Advanced Engineering Informatics, 2013, 27, 652-663.	4.0	216
8	DAPs: Deep Action Proposals for Action Understanding. Lecture Notes in Computer Science, 2016, , 768-784.	1.0	197
9	Automated 2D detection of construction equipment and workers from site video streams using histograms of oriented gradients and colors. Automation in Construction, 2013, 32, 24-37.	4.8	183
10	Fast Temporal Activity Proposals for Efficient Detection of Human Actions in Untrimmed Videos. , 2016, , .		178
11	Peeking Into the Future: Predicting Future Person Activities and Locations in Videos. , 2019, , .		136
12	Spatial-Temporal correlatons for unsupervised action classification. , 2008, , .		125
13	End-to-End, Single-Stream Temporal Action Detection in Untrimmed Videos. , 2017, , .		101
14	Spatiotemporal Relationship Reasoning for Pedestrian Intent Prediction. IEEE Robotics and Automation Letters, 2020, 5, 3485-3492.	3.3	96
15	Vision-based workface assessment using depth images for activity analysis of interior construction operations. Automation in Construction, 2014, 48, 74-87.	4.8	91
16	Connectionist Temporal Modeling for Weakly Supervised Action Labeling. Lecture Notes in Computer Science, 2016, , 137-153.	1.0	86
17	D3TW: Discriminative Differentiable Dynamic Time Warping for Weakly Supervised Action Alignment and Segmentation. , 2019, , .		84
18	Action-Agnostic Human Pose Forecasting. , 2019, , .		77

#	Article	IF	Citations
19	What Makes a Video a Video: Analyzing Temporal Information in Video Understanding Models and Datasets. , 2018, , .		71
20	Imitation Learning for Human Pose Prediction. , 2019, , .		66
21	On the relationship between visual attributes and convolutional networks. , 2015, , .		62
22	Graph Distillation for Action Detection with Privileged Modalities. Lecture Notes in Computer Science, 2018, , 174-192.	1.0	54
23	Discriminative Hierarchical Modeling of Spatio-temporally Composable Human Activities. , 2014, , .		51
24	Peeking Into the Future: Predicting Future Person Activities and Locations in Videos. , 2019, , .		50
25	Automated Vision-Based Recognition of Construction Worker Actions for Building Interior Construction Operations Using RGBD Cameras. , 2012, , .		46
26	Finding "It": Weakly-Supervised Reference-Aware Visual Grounding in Instructional Videos. , 2018, , .		46
27	Sparse composition of body poses and atomic actions for human activity recognition in RGB-D videos. Image and Vision Computing, 2017, 59, 63-75.	2.7	45
28	Neural Task Graphs: Generalizing to Unseen Tasks From a Single Video Demonstration. , 2019, , .		42
29	Quantifying Parkinsonâ€™s disease motor severity under uncertainty using MDS-UPDRS videos. Medical Image Analysis, 2021, 73, 102179.	7.0	37
30	Interpretable Visual Question Answering by Visual Grounding From Attention Supervision Mining. , 2019, , .		36
31	Visual Forecasting by Imitating Dynamics in Natural Sequences. , 2017, , .		35
32	Socially and Contextually Aware Human Motion and Pose Forecasting. IEEE Robotics and Automation Letters, 2020, 5, 6033-6040.	3.3	35
33	Representation Learning with Statistical Independence to Mitigate Bias. , 2021, 2021, 2512-2522.		35
34	Temporal Modular Networks for Retrieving Complex Compositional Activities in Videos. Lecture Notes in Computer Science, 2018, , 569-586.	1.0	34
35	Title Generation for User Generated Videos. Lecture Notes in Computer Science, 2016, , 609-625.	1.0	33
36	A Hierarchical Pose-Based Approach to Complex Action Understanding Using Dictionaries of Actionlets and Motion Poselets. , 2016, , .		32

#	Article	IF	Citations
37	Agent-Centric Risk Assessment: Accident Anticipation and Risky Region Localization. , 2017, , .		31
38	Vision-Based Estimation of MDS-UPDRS Gait Scores for Assessing Parkinson's Disease Motor Severity. Lecture Notes in Computer Science, 2020, 12263, 637-647.	1.0	30
39	Segmenting the Future. IEEE Robotics and Automation Letters, 2020, 5, 4202-4209.	3.3	26
40	Disentangling Human Dynamics for Pedestrian Locomotion Forecasting with Noisy Supervision. , 2020, , .		26
41	Efficient extraction of human motion volumes by tracking. , 2010, , .		25
42	Robust Manhattan Frame estimation from a single RGB-D image. , 2015, , .		25
43	Learning Temporal Action Proposals With Fewer Labels. , 2019, , .		25
44	Extracting Moving People from Internet Videos. Lecture Notes in Computer Science, 2008, , 527-540.	1.0	25
45	Home Action Genome: Cooperative Compositional Action Understanding. , 2021, , .		25
46	Spatio-temporal Human-Object Interactions for Action Recognition in Videos. , 2013, , .		24
47	Unsupervised Visual-Linguistic Reference Resolution in Instructional Videos. , 2017, , .		23
48	Learning Privacy-preserving Optics for Human Pose Estimation. , 2021, , .		22
49	Automated Visual Recognition of Construction Equipment Actions Using Spatio-Temporal Features and Multiple Binary Support Vector Machines. , 2012, , .		20
50	Real-Time and Automated Recognition and 2D Tracking of Construction Workers and Equipment from Site Video Streams. , 2012, , .		20
51	A Deep Learning Based Behavioral Approach to Indoor Autonomous Navigation. , 2018, , .		20
52	Procedure Planning in Instructional Videos. Lecture Notes in Computer Science, 2020, , 334-350.	1.0	19
53	Collecting and Annotating Human Activities in Web Videos. , 2014, , .		18
54	TRiPOD: Human Trajectory and Pose Dynamics Forecasting in the Wild. , 2021, , .		18

#	Article	IF	Citations
55	End-to-End Joint Semantic Segmentation of Actors and Actions in Video. Lecture Notes in Computer Science, 2018, , 734-749.	1.0	17
56	Continuous Relaxation of Symbolic Planner for One-Shot Imitation Learning. , 2019, , .		15
57	Automated Worker Activity Analysis in Indoor Environments for Direct-Work Rate Improvement from Long Sequences of RGB-D Images. , 2014, , .		12
58	Explaining VQA predictions using visual grounding and a knowledge base. Image and Vision Computing, 2020, 101, 103968.	2.7	12
59	Metadata Normalization. , 2021, 2021, 10912-10922.		11
60	Behavioral Indoor Navigation With Natural Language Directions. , 2018, , .		4
61	Liquid Pouring Monitoring via Rich Sensory Inputs. Lecture Notes in Computer Science, 2018, , 352-369.	1.0	4
62	Risky Region Localization with Point Supervision. , 2017, , .		3
63	Vision-Based Construction Activity Analysis in Long Video Sequences via Hidden Markov Models: Experiments on Earthmoving Operations. , 2018, , .		3
64	Camera Motion and Surrounding Scene Appearance as Context for Action Recognition. Lecture Notes in Computer Science, 2015, , 583-597.	1.0	2
65	Mining discriminative adjectives and prepositions for natural scene recognition. , 2009, , .		1
66	Mining discriminative adjectives and prepositions for natural scene recognition. , 2009, , .		0
67	A shoreline-estimation system using remote radar sensing and image-processing techniques. DYNA (Colombia), 2017, 84, 151-159.	0.2	0