

Li Fei-Fei

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/12097401/publications.pdf>

Version: 2024-02-01

95
papers

83,966
citations

87843

38
h-index

138417

58
g-index

97
all docs

97
docs citations

97
times ranked

47071
citing authors

#	ARTICLE	IF	CITATIONS
1	ImageNet: A large-scale hierarchical image database. , 2009, , .		29,755
2	ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision, 2015, 115, 211-252.	10.9	24,215
3	Large-Scale Video Classification with Convolutional Neural Networks. , 2014, , .		4,114
4	Perceptual Losses for Real-Time Style Transfer and Super-Resolution. Lecture Notes in Computer Science, 2016, , 694-711.	1.0	4,037
5	Deep visual-semantic alignments for generating image descriptions. , 2015, , .		2,539
6	Visual Genome: Connecting Language and Vision Using Crowdsourced Dense Image Annotations. International Journal of Computer Vision, 2017, 123, 32-73.	10.9	2,422
7	One-shot learning of object categories. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28, 594-611.	9.7	1,947
8	Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories. Computer Vision and Image Understanding, 2007, 106, 59-70.	3.0	1,537
9	Unsupervised Learning of Human Action Categories Using Spatial-Temporal Words. International Journal of Computer Vision, 2008, 79, 299-318.	10.9	1,227
10	Learning Generative Visual Models from Few Training Examples: An Incremental Bayesian Approach Tested on 101 Object Categories. , 0, , .		846
11	CLEVR: A Diagnostic Dataset for Compositional Language and Elementary Visual Reasoning. , 2017, , .		742
12	Image retrieval using scene graphs. , 2015, , .		569
13	Auto-DeepLab: Hierarchical Neural Architecture Search for Semantic Image Segmentation. , 2019, , .		547
14	Deep Visual-Semantic Alignments for Generating Image Descriptions. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39, 664-676.	9.7	542
15	What, where and who? Classifying events by scene and object recognition. , 2007, , .		537
16	Visual Relationship Detection with Language Priors. Lecture Notes in Computer Science, 2016, , 852-869.	1.0	444
17	Modeling mutual context of object and human pose in human-object interaction activities. , 2010, , .		410
18	Modeling Temporal Structure of Decomposable Motion Segments for Activity Classification. Lecture Notes in Computer Science, 2010, , 392-405.	1.0	340

#	ARTICLE	IF	CITATIONS
19	Natural Scene Categories Revealed in Distributed Patterns of Activity in the Human Brain. Journal of Neuroscience, 2009, 29, 10573-10581.	1.7	314
20	ImageNet: A large-scale hierarchical image database. , 2009, , .		313
21	What do we perceive in a glance of a real-world scene?. Journal of Vision, 2007, 7, 10.	0.1	312
22	What's the Point: Semantic Segmentation with Point Supervision. Lecture Notes in Computer Science, 2016, , 549-565.	1.0	288
23	Neural mechanisms of rapid natural scene categorization in human visual cortex. Nature, 2009, 460, 94-97.	13.7	271
24	Towards total scene understanding: Classification, annotation and segmentation in an automatic framework. , 2009, , .		247
25	Learning latent temporal structure for complex event detection. , 2012, , .		239
26	Inferring and Executing Programs for Visual Reasoning. , 2017, , .		239
27	Grouplet: A structured image representation for recognizing human and object interactions. , 2010, , .		221
28	What Does Classifying More Than 10,000 Image Categories Tell Us?. Lecture Notes in Computer Science, 2010, , 71-84.	1.0	220
29	Every Moment Counts: Dense Detailed Labeling of Actions in Complex Videos. International Journal of Computer Vision, 2018, 126, 375-389.	10.9	192
30	Simple line drawings suffice for functional MRI decoding of natural scene categories. Proceedings of the National Academy of Sciences of the United States of America, 2011, 108, 9661-9666.	3.3	189
31	Generating Semantically Precise Scene Graphs from Textual Descriptions for Improved Image Retrieval. , 2015, , .		183
32	Recognizing Human-Object Interactions in Still Images by Modeling the Mutual Context of Objects and Human Poses. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34, 1691-1703.	9.7	170
33	Tool Detection and Operative Skill Assessment in Surgical Videos Using Region-Based Convolutional Neural Networks. , 2018, , .		169
34	The Unreasonable Effectiveness of Noisy Data for Fine-Grained Recognition. Lecture Notes in Computer Science, 2016, , 301-320.	1.0	150
35	Action Genome: Actions As Compositions of Spatio-Temporal Scene Graphs. , 2020, , .		150
36	OPTIMOL: Automatic Online Picture Collection via Incremental Model Learning. International Journal of Computer Vision, 2010, 88, 147-168.	10.9	144

#	ARTICLE	IF	CITATIONS
37	Illuminating the dark spaces of healthcare with ambient intelligence. <i>Nature</i> , 2020, 585, 193-202.	13.7	139
38	Differential connectivity within the Parahippocampal Place Area. <i>NeuroImage</i> , 2013, 75, 228-237.	2.1	137
39	Distinct contributions of functional and deep neural network features to representational similarity of scenes in human brain and behavior. <i>ELife</i> , 2018, 7, .	2.8	132
40	Spatial-Temporal correlatons for unsupervised action classification. , 2008, , .		125
41	Towards fairer datasets. , 2020, , .		123
42	OPTIMOL: automatic Online Picture collecTION via Incremental MOdel Learning. , 2007, , .		113
43	Object Bank: An Object-Level Image Representation for High-Level Visual Recognition. <i>International Journal of Computer Vision</i> , 2014, 107, 20-39.	10.9	111
44	Two Distinct Scene-Processing Networks Connecting Vision and Memory. <i>ENeuro</i> , 2016, 3, ENEURO.0178-16.2016.	0.9	111
45	Reasoning about Object Affordances in a Knowledge Base Representation. <i>Lecture Notes in Computer Science</i> , 2014, , 408-424.	1.0	101
46	Why does natural scene categorization require little attention? Exploring attentional requirements for natural and synthetic stimuli. <i>Visual Cognition</i> , 2005, 12, 893-924.	0.9	94
47	Scalable multi-label annotation. , 2014, , .		91
48	Bedside Computer Vision “ Moving Artificial Intelligence from Driver Assistance to Patient Safety. <i>New England Journal of Medicine</i> , 2018, 378, 1271-1273.	13.9	91
49	Fine-Grained Recognition in the Wild: A Multi-task Domain Adaptation Approach. , 2017, , .		87
50	Learning task-oriented grasping for tool manipulation from simulated self-supervision. <i>International Journal of Robotics Research</i> , 2020, 39, 202-216.	5.8	87
51	Connectionist Temporal Modeling for Weakly Supervised Action Labeling. <i>Lecture Notes in Computer Science</i> , 2016, , 137-153.	1.0	86
52	Love Thy Neighbors: Image Annotation by Exploiting Image Metadata. , 2015, , .		76
53	Learning semantic relationships for better action retrieval in images. , 2015, , .		74
54	A computer vision system for deep learning-based detection of patient mobilization activities in the ICU. <i>Npj Digital Medicine</i> , 2019, 2, 11.	5.7	73

#	ARTICLE	IF	CITATIONS
55	Using Dependent Regions for Object Categorization in a Generative Framework. , 0, , .		72
56	Towards Scalable Dataset Construction: An Active Learning Approach. Lecture Notes in Computer Science, 2008, , 86-98.	1.0	69
57	Parcellating connectivity in spatial maps. PeerJ, 2015, 3, e784.	0.9	66
58	Visual scenes are categorized by function.. Journal of Experimental Psychology: General, 2016, 145, 82-94.	1.5	60
59	Visual categorization is automatic and obligatory: Evidence from Stroop-like paradigm. Journal of Vision, 2014, 14, 14-14.	0.1	59
60	Learning Object Categories From Internet Image Searches. Proceedings of the IEEE, 2010, 98, 1453-1466.	16.4	57
61	What you see is what you expect: rapid scene understanding benefits from prior experience. Attention, Perception, and Psychophysics, 2015, 77, 1239-1251.	0.7	56
62	Online Developmental Science to Foster Innovation, Access, and Impact. Trends in Cognitive Sciences, 2020, 24, 675-678.	4.0	53
63	Detecting Avocados to Zucchini: What Have We Done, and Where Are We Going?. , 2013, , .		48
64	Ethical issues in using ambient intelligence in health-care settings. The Lancet Digital Health, 2021, 3, e115-e123.	5.9	48
65	Automated abnormality detection in lower extremity radiographs using deep learning. Nature Machine Intelligence, 2019, 1, 578-583.	8.3	47
66	Embracing Error to Enable Rapid Crowdsourcing. , 2016, , .		44
67	Basic Level Category Structure Emerges Gradually across Human Ventral Visual Cortex. Journal of Cognitive Neuroscience, 2015, 27, 1427-1446.	1.1	42
68	Humanâ€™Object Interactions Are More than the Sum of Their Parts. Cerebral Cortex, 2017, 27, bhw077.	1.6	41
69	Quantifying Parkinsonâ€™s disease motor severity under uncertainty using MDS-UPDRS videos. Medical Image Analysis, 2021, 73, 102179.	7.0	37
70	A multi-view probabilistic model for 3D object classes. , 2009, , .		36
71	Situational Fusion of Visual Representation for Visual Navigation. , 2019, , .		36
72	Assessing the accuracy of automatic speech recognition for psychotherapy. Npj Digital Medicine, 2020, 3, 82.	5.7	35

#	ARTICLE	IF	CITATIONS
73	RubiksNet: Learnable 3D-Shift for Efficient Video Action Recognition. Lecture Notes in Computer Science, 2020, , 505-521.	1.0	33
74	Typicality sharpens category representations in object-selective cortex. NeuroImage, 2016, 134, 170-179.	2.1	32
75	Discovering Object Functionality. , 2013, , .		31
76	Automatic detection of hand hygiene using computer vision technology. Journal of the American Medical Informatics Association: JAMIA, 2020, 27, 1316-1320.	2.2	31
77	Voxel-level functional connectivity using spatial regularization. NeuroImage, 2012, 63, 1099-1106.	2.1	30
78	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
79	Good Exemplars of Natural Scene Categories Elicit Clearer Patterns than Bad Exemplars but Not Greater BOLD Activity. PLoS ONE, 2013, 8, e58594.	1.1	29
80	Binding is a local problem for natural objects and scenes. Vision Research, 2005, 45, 3133-3144.	0.7	26
81	Learning Task-Oriented Grasping for Tool Manipulation from Simulated Self-Supervision. , 0, , .		26
82	Extracting Moving People from Internet Videos. Lecture Notes in Computer Science, 2008, , 527-540.	1.0	25
83	Pinpointing the peripheral bias in neural scene-processing networks during natural viewing. Journal of Vision, 2016, 16, 9.	0.1	22
84	Evidence for similar patterns of neural activity elicited by picture- and word-based representations of natural scenes. NeuroImage, 2017, 155, 422-436.	2.1	21
85	Deep Affordance Foresight: Planning Through What Can Be Done in the Future. , 2021, , .		19
86	Task-set switching with natural scenes: Measuring the cost of deploying top-down attention. Journal of Vision, 2007, 7, 9.	0.1	17
87	Categorization influences detection: A perceptual advantage for representative exemplars of natural scene categories. Journal of Vision, 2017, 17, 21.	0.1	12
88	Scene Graph Prediction With Limited Labels. , 2019, 2019, 2580-2590.		12
89	Audio-Visual Speaker Localization Using Graphical Models. , 2006, , .		6
90	Neural Event Semantics for Grounded Language Understanding. Transactions of the Association for Computational Linguistics, 2021, 9, 875-890.	3.2	3

#	ARTICLE	IF	CITATIONS
91	A multi-view probabilistic model for 3D object classes. , 2009, , .		3
92	Thoracic Disease Identification and Localization with Limited Supervision. Advances in Computer Vision and Pattern Recognition, 2019, , 139-161.	0.9	3
93	Visual Intelligence through Human Interaction. Human-computer Interaction Series, 2021, , 257-314.	0.4	3
94	Towards total scene understanding: Classification, annotation and segmentation in an automatic framework. , 2009, , .		2
95	Mining discriminative adjectives and prepositions for natural scene recognition. , 2009, , .		1