

# Jun Yu

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/8101642/publications.pdf>

Version: 2024-02-01

81  
papers

6,714  
citations

185998

28  
h-index

95083

68  
g-index

91  
all docs

91  
docs citations

91  
times ranked

4142  
citing authors

#	ARTICLE	IF	CITATIONS
1	Click Prediction for Web Image Reranking Using Multimodal Sparse Coding. IEEE Transactions on Image Processing, 2014, 23, 2019-2032.	6.0	491
2	Multimodal Deep Autoencoder for Human Pose Recovery. IEEE Transactions on Image Processing, 2015, 24, 5659-5670.	6.0	488
3	Deep Modular Co-Attention Networks for Visual Question Answering. , 2019, , .		430
4	Multi-modal Factorized Bilinear Pooling with Co-attention Learning for Visual Question Answering. , 2017, , .		388
5	Hierarchical Deep Click Feature Prediction for Fine-Grained Image Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44, 563-578.	9.7	359
6	Learning to Rank Using User Clicks and Visual Features for Image Retrieval. IEEE Transactions on Cybernetics, 2015, 45, 767-779.	6.2	344
7	Beyond Bilinear: Generalized Multimodal Factorized High-Order Pooling for Visual Question Answering. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29, 5947-5959.	7.2	304
8	iPrivacy: Image Privacy Protection by Identifying Sensitive Objects via Deep Multi-Task Learning. IEEE Transactions on Information Forensics and Security, 2017, 12, 1005-1016.	4.5	260
9	Multimodal Transformer With Multi-View Visual Representation for Image Captioning. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30, 4467-4480.	5.6	258
10	Semisupervised Multiview Distance Metric Learning for Cartoon Synthesis. IEEE Transactions on Image Processing, 2012, 21, 4636-4648.	6.0	254
11	High-Order Distance-Based Multiview Stochastic Learning in Image Classification. IEEE Transactions on Cybernetics, 2014, 44, 2431-2442.	6.2	246
12	Deep Multimodal Distance Metric Learning Using Click Constraints for Image Ranking. IEEE Transactions on Cybernetics, 2017, 47, 4014-4024.	6.2	246
13	Spatial Pyramid-Enhanced NetVLAD With Weighted Triplet Loss for Place Recognition. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31, 661-674.	7.2	229
14	Multimodal Face-Pose Estimation With Multitask Manifold Deep Learning. IEEE Transactions on Industrial Informatics, 2019, 15, 3952-3961.	7.2	225
15	Local Deep-Feature Alignment for Unsupervised Dimension Reduction. IEEE Transactions on Image Processing, 2018, 27, 2420-2432.	6.0	182
16	Leveraging Content Sensitiveness and User Trustworthiness to Recommend Fine-Grained Privacy Settings for Social Image Sharing. IEEE Transactions on Information Forensics and Security, 2018, 13, 1317-1332.	4.5	166
17	Exploiting Click Constraints and Multi-view Features for Image Re-ranking. IEEE Transactions on Multimedia, 2014, 16, 159-168.	5.2	155
18	Multitask Autoencoder Model for Recovering Human Poses. IEEE Transactions on Industrial Electronics, 2018, 65, 5060-5068.	5.2	136

#	ARTICLE	IF	CITATIONS
19	DeepSim: Deep similarity for image quality assessment. Neurocomputing, 2017, 257, 104-114.	3.5	129
20	Image clustering based on sparse patch alignment framework. Pattern Recognition, 2014, 47, 3512-3519.	5.1	95
21	Blind image quality prediction by exploiting multi-level deep representations. Pattern Recognition, 2018, 81, 432-442.	5.1	95
22	ActivityNet-QA: A Dataset for Understanding Complex Web Videos via Question Answering. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33, 9127-9134.	3.6	88
23	Rethinking Diversified and Discriminative Proposal Generation for Visual Grounding. , 2018, , .		63
24	Biologically inspired image quality assessment. Signal Processing, 2016, 124, 210-219.	2.1	61
25	On Exploring Undetermined Relationships for Visual Relationship Detection. , 2019, , .		57
26	Toward Realistic Face Photoâ€“Sketch Synthesis via Composition-Aided GANs. IEEE Transactions on Cybernetics, 2021, 51, 4350-4362.	6.2	50
27	Semisupervised image classification by mutual learning of multiple selfâ€“supervised models. International Journal of Intelligent Systems, 2022, 37, 3117-3141.	3.3	46
28	SPRNet: Single-Pixel Reconstruction for One-Stage Instance Segmentation. IEEE Transactions on Cybernetics, 2021, 51, 1731-1742.	6.2	43
29	Multi-interaction Network with Object Relation for Video Question Answering. , 2019, , .		42
30	Asymmetric Supervised Consistent and Specific Hashing for Cross-Modal Retrieval. IEEE Transactions on Image Processing, 2021, 30, 986-1000.	6.0	42
31	Deep Multimodal Neural Architecture Search. , 2020, , .		42
32	Integrating multi-level deep learning and concept ontology for large-scale visual recognition. Pattern Recognition, 2018, 78, 198-214.	5.1	39
33	Zero-Shot Learning via Robust Latent Representation and Manifold Regularization. IEEE Transactions on Image Processing, 2019, 28, 1824-1836.	6.0	35
34	Scalable Zero-Shot Learning via Binary Visual-Semantic Embeddings. IEEE Transactions on Image Processing, 2019, 28, 3662-3674.	6.0	30
35	Long-Term Video Question Answering via Multimodal Hierarchical Memory Attentive Networks. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31, 931-944.	5.6	30
36	Open-Ended Long-form Video Question Answering via Adaptive Hierarchical Reinforced Networks. , 2018, , .		30

#	ARTICLE	IF	CITATIONS
37	Compositional Attention Networks With Two-Stream Fusion for Video Question Answering. IEEE Transactions on Image Processing, 2020, 29, 1204-1218.	6.0	29
38	Long-Form Video Question Answering via Dynamic Hierarchical Reinforced Networks. IEEE Transactions on Image Processing, 2019, 28, 5939-5952.	6.0	28
39	Constrained Discriminative Projection Learning for Image Classification. IEEE Transactions on Image Processing, 2020, 29, 186-198.	6.0	26
40	Face biometric quality assessment via light CNN. Pattern Recognition Letters, 2018, 107, 25-32.	2.6	25
41	Human pose recovery by supervised spectral embedding. Neurocomputing, 2015, 166, 301-308.	3.5	24
42	Multi-task Compositional Network for Visual Relationship Detection. International Journal of Computer Vision, 2020, 128, 2146-2165.	10.9	22
43	Diversified dictionaries for multi-instance learning. Pattern Recognition, 2017, 64, 407-416.	5.1	18
44	User-Click-Data-Based Fine-Grained Image Recognition via Weakly Supervised Metric Learning. ACM Transactions on Multimedia Computing, Communications and Applications, 2018, 14, 1-23.	3.0	17
45	Fine-grained image recognition via weakly supervised click data guided bilinear CNN model. , 2017, , .		16
46	Unsupervised image segmentation via Stacked Denoising Auto-encoder and hierarchical patch indexing. Signal Processing, 2018, 143, 346-353.	2.1	16
47	$l_{2,1}$ Norm regularized fisher criterion for optimal feature selection. Neurocomputing, 2015, 166, 455-463.	3.5	15
48	Image Recognition by Predicted User Click Feature With Multidomain Multitask Transfer Deep Network. IEEE Transactions on Image Processing, 2019, 28, 6047-6062.	6.0	15
49	Effective De-identification Generative Adversarial Network for Face Anonymization. , 2021, , .		15
50	Adapting Stochastic Block Models to Power-Law Degree Distributions. IEEE Transactions on Cybernetics, 2019, 49, 626-637.	6.2	14
51	Complementary, Heterogeneous and Adversarial Networks for Image-to-Image Translation. IEEE Transactions on Image Processing, 2021, 30, 3487-3498.	6.0	14
52	Semantic embedding for indoor scene recognition by weighted hypergraph learning. Signal Processing, 2015, 112, 129-136.	2.1	13
53	Data-driven facial animation via semi-supervised local patch alignment. Pattern Recognition, 2016, 57, 1-20.	5.1	13
54	Representation learning of image composition for aesthetic prediction. Computer Vision and Image Understanding, 2020, 199, 103024.	3.0	13

#	ARTICLE	IF	CITATIONS
55	Toward Multi-Modal Conditioned Fashion Image Translation. IEEE Transactions on Multimedia, 2021, 23, 2361-2371.	5.2	13
56	Deep embedding of concept ontology for hierarchical fashion recognition. Neurocomputing, 2021, 425, 191-206.	3.5	13
57	Fine-grained Image Classification via Multi-scale Selective Hierarchical Biquadratic Pooling. ACM Transactions on Multimedia Computing, Communications and Applications, 2022, 18, 1-23.	3.0	12
58	Generalized Multi-View Collaborative Subspace Clustering. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32, 3561-3574.	5.6	11
59	Effective 3-D Shape Retrieval by Integrating Traditional Descriptors and Pointwise Convolution. IEEE Transactions on Multimedia, 2019, 21, 3164-3177.	5.2	10
60	Style-adaptive photo aesthetic rating via convolutional neural networks and multi-task learning. Neurocomputing, 2020, 395, 247-254.	3.5	10
61	Fine-grained image classification with factorized deep user click feature. Information Processing and Management, 2020, 57, 102186.	5.4	9
62	Multiview Consensus Structure Discovery. IEEE Transactions on Cybernetics, 2022, 52, 3469-3482.	6.2	9
63	Distributed feedback network for single-image deraining. Information Sciences, 2021, 572, 611-626.	4.0	9
64	TVENet: Temporal variance embedding network for fine-grained action representation. Pattern Recognition, 2020, 103, 107267.	5.1	8
65	Coupled Knowledge Transfer for Visual Data Recognition. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31, 1776-1789.	5.6	7
66	Deep Neural Network Boosted Large Scale Image Recognition Using User Click Data. , 2016, , .		6
67	An Individual-Difference-Aware Model for Cross-Person Gaze Estimation. IEEE Transactions on Image Processing, 2022, 31, 3322-3333.	6.0	6
68	Click data guided query modeling with click propagation and sparse coding. Multimedia Tools and Applications, 2018, 77, 22145-22158.	2.6	5
69	Deep Point Convolutional Approach for 3D Model Retrieval. , 2018, , .		5
70	TaoHighlight: Commodity-Aware Multi-Modal Video Highlight Detection in E-Commerce. IEEE Transactions on Multimedia, 2022, 24, 2606-2616.	5.2	4
71	Improving Stochastic Block Models by Incorporating Power-Law Degree Characteristic. , 2017, , .		4
72	Deformable Point Cloud Recognition Using Intrinsic Function and Deep Learning. Lecture Notes in Computer Science, 2018, , 89-101.	1.0	3

#	ARTICLE	IF	CITATIONS
73	Intra- and Inter-modal Multilinear Pooling with Multitask Learning for Video Grounding. Neural Processing Letters, 2020, 52, 1863-1879.	2.0	3
74	Incremental focal loss GANs. Information Processing and Management, 2020, 57, 102192.	5.4	3
75	Three-dimensional image-based human pose recovery with hypergraph regularized autoencoders. Multimedia Tools and Applications, 2017, 76, 10919-10937.	2.6	2
76	End-to-end visual grounding via region proposal networks and bilinear pooling. IET Computer Vision, 2019, 13, 131-138.	1.3	1
77	Relationship graph learning network for visual relationship detection. , 2021, , .		1
78	Query Modeling for Click Data Based Image Recognition Using Graph Based Propagation and Sparse Coding. Communications in Computer and Information Science, 2018, , 191-199.	0.4	1
79	Position constrained network for 3D human pose estimation. Multimedia Systems, 2023, 29, 459-468.	3.0	1
80	Deep click feature based query merging for robust image recognition. , 2018, , .		0
81	Contrastive learning of graph encoder for accelerating pedestrian trajectory prediction training. IET Image Processing, 0, , .	1.4	0