

# Heng Tao Shen

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/3831803/publications.pdf>

Version: 2024-02-01

219  
papers

14,313  
citations

44069

48  
h-index

37204

96  
g-index

224  
all docs

224  
docs citations

224  
times ranked

7354  
citing authors

#	ARTICLE	IF	CITATIONS
1	Supervised Discrete Hashing. , 2015, , .		838
2	Near-duplicate video retrieval. ACM Computing Surveys, 2013, 45, 1-23.	23.0	769
3	A Survey on Learning to Hash. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40, 769-790.	13.9	650
4	Adversarial Cross-Modal Retrieval. , 2017, , .		509
5	Video Captioning With Attention-Based LSTM and Semantic Consistency. IEEE Transactions on Multimedia, 2017, 19, 2045-2055.	7.2	491
6	Inter-media hashing for large-scale retrieval from heterogeneous data sources. , 2013, , .		442
7	Learning Discriminative Binary Codes for Large-scale Cross-modal Retrieval. IEEE Transactions on Image Processing, 2017, 26, 2494-2507.	9.8	345
8	Binary Multi-View Clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41, 1774-1782.	13.9	342
9	Unsupervised Deep Hashing with Similarity-Adaptive and Discrete Optimization. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40, 3034-3044.	13.9	294
10	Multiple feature hashing for real-time large scale near-duplicate video retrieval. , 2011, , .		256
11	Discovering popular routes from trajectories. , 2011, , .		256
12	Deep Fuzzy Hashing Network for Efficient Image Retrieval. IEEE Transactions on Fuzzy Systems, 2021, 29, 166-176.	9.8	236
13	Linear cross-modal hashing for efficient multimedia search. , 2013, , .		221
14	Effective Multiple Feature Hashing for Large-Scale Near-Duplicate Video Retrieval. IEEE Transactions on Multimedia, 2013, 15, 1997-2008.	7.2	213
15	Transfer Independently Together: A Generalized Framework for Domain Adaptation. IEEE Transactions on Cybernetics, 2019, 49, 2144-2155.	9.5	208
16	Locality Preserving Joint Transfer for Domain Adaptation. IEEE Transactions on Image Processing, 2019, 28, 6103-6115.	9.8	185
17	Describing Video With Attention-Based Bidirectional LSTM. IEEE Transactions on Cybernetics, 2019, 49, 2631-2641.	9.5	179
18	Video Captioning by Adversarial LSTM. IEEE Transactions on Image Processing, 2018, 27, 5600-5611.	9.8	174

#	ARTICLE	IF	CITATIONS
19	From Deterministic to Generative: Multimodal Stochastic RNNs for Video Captioning. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30, 3047-3058.	11.3	160
20	Heterogeneous Domain Adaptation Through Progressive Alignment. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30, 1381-1391.	11.3	154
21	Multitask Spectral Clustering by Exploring Intertask Correlation. IEEE Transactions on Cybernetics, 2015, 45, 1083-1094.	9.5	150
22	Maximum Density Divergence for Domain Adaptation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43, 3918-3930.	13.9	149
23	Cross-Modal Attention With Semantic Consistence for Image-Text Matching. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31, 5412-5425.	11.3	148
24	Hierarchical LSTMs with Adaptive Attention for Visual Captioning. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 42, 1-1.	13.9	143
25	Ternary Adversarial Networks With Self-Supervision for Zero-Shot Cross-Modal Retrieval. IEEE Transactions on Cybernetics, 2020, 50, 2400-2413.	9.5	142
26	Multi-source Skyline Query Processing in Road Networks. , 2007, , .		138
27	A Fast Optimization Method for General Binary Code Learning. IEEE Transactions on Image Processing, 2016, 25, 5610-5621.	9.8	138
28	Exploiting Subspace Relation in Semantic Labels for Cross-Modal Hashing. IEEE Transactions on Knowledge and Data Engineering, 2021, 33, 3351-3365.	5.7	133
29	Exploring Auxiliary Context: Discrete Semantic Transfer Hashing for Scalable Image Retrieval. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29, 5264-5276.	11.3	127
30	Self-taught dimensionality reduction on the high-dimensional small-sized data. Pattern Recognition, 2013, 46, 215-229.	8.1	126
31	Dimensionality reduction by Mixed Kernel Canonical Correlation Analysis. Pattern Recognition, 2012, 45, 3003-3016.	8.1	121
32	Hashing on Nonlinear Manifolds. IEEE Transactions on Image Processing, 2015, 24, 1839-1851.	9.8	120
33	Hierarchical LSTM with Adjusted Temporal Attention for Video Captioning. , 2017, , .		111
34	Zero-Shot Hashing via Transferring Supervised Knowledge. , 2016, , .		109
35	Collective Reconstructive Embeddings for Cross-Modal Hashing. IEEE Transactions on Image Processing, 2019, 28, 2770-2784.	9.8	109
36	Asymmetric Binary Coding for Image Search. IEEE Transactions on Multimedia, 2017, 19, 2022-2032.	7.2	108

#	ARTICLE	IF	CITATIONS
37	Optimized Graph Learning Using Partial Tags and Multiple Features for Image and Video Annotation. IEEE Transactions on Image Processing, 2016, 25, 4999-5011.	9.8	107
38	Sparse hashing for fast multimedia search. ACM Transactions on Information Systems, 2013, 31, 1-24.	4.9	105
39	Towards effective indexing for very large video sequence database. , 2005, , .		102
40	A Distance-Computation-Free Search Scheme for Binary Code Databases. IEEE Transactions on Multimedia, 2016, 18, 484-495.	7.2	101
41	Semi-Paired Discrete Hashing: Learning Latent Hash Codes for Semi-Paired Cross-View Retrieval. IEEE Transactions on Cybernetics, 2017, 47, 4275-4288.	9.5	101
42	Robust discrete code modeling for supervised hashing. Pattern Recognition, 2018, 75, 128-135.	8.1	101
43	Learning Binary Codes for Maximum Inner Product Search. , 2015, , .		94
44	Graph Convolutional Network Hashing. IEEE Transactions on Cybernetics, 2020, 50, 1460-1472.	9.5	90
45	Giving meanings to WWW images. , 2000, , .		88
46	Deep Asymmetric Pairwise Hashing. , 2017, , .		88
47	Matching Images and Text with Multi-modal Tensor Fusion and Re-ranking. , 2019, , .		88
48	Temporal Pyramid Pooling-Based Convolutional Neural Network for Action Recognition. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 27, 2613-2622.	8.3	85
49	Matrix Tri-Factorization with Manifold Regularizations for Zero-Shot Learning. , 2017, , .		85
50	Hashing with Angular Reconstructive Embeddings. IEEE Transactions on Image Processing, 2018, 27, 545-555.	9.8	85
51	Tag localization with spatial correlations and joint group sparsity. , 2011, , .		83
52	Efficient Semantic-Based Content Search in P2P Network. IEEE Transactions on Knowledge and Data Engineering, 2004, 16, 813-826.	5.7	82
53	Robust Discrete Spectral Hashing for Large-Scale Image Semantic Indexing. IEEE Transactions on Big Data, 2015, 1, 162-171.	6.1	82
54	Remote Sensing Image Super-Resolution via Mixed High-Order Attention Network. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59, 5183-5196.	6.3	82

#	ARTICLE	IF	CITATIONS
55	SK-LSH. Proceedings of the VLDB Endowment, 2014, 7, 745-756.	3.8	69
56	Bounded coordinate system indexing for real-time video clip search. ACM Transactions on Information Systems, 2009, 27, 1-33.	4.9	67
57	Discrete Nonnegative Spectral Clustering. IEEE Transactions on Knowledge and Data Engineering, 2017, 29, 1834-1845.	5.7	65
58	More is Better: Precise and Detailed Image Captioning Using Online Positive Recall and Missing Concepts Mining. IEEE Transactions on Image Processing, 2019, 28, 32-44.	9.8	65
59	Heterogeneous data fusion for predicting mild cognitive impairment conversion. Information Fusion, 2021, 66, 54-63.	19.1	60
60	The Gap of Semantic Parsing: A Survey on Automatic Math Word Problem Solvers. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42, 2287-2305.	13.9	59
61	Universal Weighting Metric Learning for Cross-Modal Matching. , 2020, , .		58
62	Optimal graph learning with partial tags and multiple features for image and video annotation. , 2015, , .		57
63	Discriminative Nonnegative Spectral Clustering with Out-of-Sample Extension. IEEE Transactions on Knowledge and Data Engineering, 2013, 25, 1760-1771.	5.7	56
64	A Survey of Human Action Analysis in HRI Applications. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 30, 2114-2128.	8.3	54
65	IF-Matching: Towards Accurate Map-Matching with Information Fusion. IEEE Transactions on Knowledge and Data Engineering, 2017, 29, 114-127.	5.7	52
66	MRA-Net: Improving VQA Via Multi-Modal Relation Attention Network. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44, 318-329.	13.9	51
67	Generative Domain-Migration Hashing for Sketch-to-Image Retrieval. Lecture Notes in Computer Science, 2018, , 304-321.	1.3	51
68	Effective transfer tagging from image to video. ACM Transactions on Multimedia Computing, Communications and Applications, 2013, 9, 1-20.	4.3	50
69	Video-to-Shot Tag Propagation by Graph Sparse Group Lasso. IEEE Transactions on Multimedia, 2013, 15, 633-646.	7.2	49
70	TBN: Convolutional Neural Network with Ternary Inputs and Binary Weights. Lecture Notes in Computer Science, 2018, , 322-339.	1.3	49
71	Efficient Supervised Discrete Multi-View Hashing for Large-Scale Multimedia Search. IEEE Transactions on Multimedia, 2020, 22, 2048-2060.	7.2	48
72	Half-Quadratic Minimization for Unsupervised Feature Selection on Incomplete Data. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32, 3122-3135.	11.3	48

#	ARTICLE	IF	CITATIONS
73	Towards Automatic Construction of Diverse, High-Quality Image Datasets. IEEE Transactions on Knowledge and Data Engineering, 2020, 32, 1199-1211.	5.7	47
74	Unified Binary Generative Adversarial Network for Image Retrieval and Compression. International Journal of Computer Vision, 2020, 128, 2243-2264.	15.6	47
75	Practical Online Near-Duplicate Subsequence Detection for Continuous Video Streams. IEEE Transactions on Multimedia, 2010, 12, 386-398.	7.2	46
76	Bidirectional Long-Short Term Memory for Video Description. , 2016, , .		46
77	Face image classification by pooling raw features. Pattern Recognition, 2016, 54, 94-103.	8.1	46
78	Learnable Aggregating Net with Diversity Learning for Video Question Answering. , 2019, , .		46
79	On Both Cold-Start and Long-Tail Recommendation with Social Data. IEEE Transactions on Knowledge and Data Engineering, 2021, 33, 194-208.	5.7	46
80	Attention-based LSTM with Semantic Consistency for Videos Captioning. , 2016, , .		45
81	Deep appearance and motion learning for egocentric activity recognition. Neurocomputing, 2018, 275, 438-447.	5.9	44
82	Temporal Reasoning Graph for Activity Recognition. IEEE Transactions on Image Processing, 2020, 29, 5491-5506.	9.8	44
83	Robust Web Image Annotation via Exploring Multi-Facet and Structural Knowledge. IEEE Transactions on Image Processing, 2017, 26, 4871-4884.	9.8	43
84	Effective and Efficient Query Processing for Video Subsequence Identification. IEEE Transactions on Knowledge and Data Engineering, 2009, 21, 321-334.	5.7	39
85	Local image tagging via graph regularized joint group sparsity. Pattern Recognition, 2013, 46, 1358-1368.	8.1	39
86	Hierarchical Representation Network With Auxiliary Tasks for Video Captioning and Video Question Answering. IEEE Transactions on Image Processing, 2022, 31, 202-215.	9.8	39
87	Mining multi-tag association for image tagging. World Wide Web, 2011, 14, 133-156.	4.0	38
88	Hierarchical Indexing Structure for Efficient Similarity Search in Video Retrieval. IEEE Transactions on Knowledge and Data Engineering, 2006, 18, 1544-1559.	5.7	37
89	Hierarchical Multi-Clue Modelling for POI Popularity Prediction with Heterogeneous Tourist Information. IEEE Transactions on Knowledge and Data Engineering, 2019, 31, 757-768.	5.7	37
90	Arbitrary-view human action recognition via novel-view action generation. Pattern Recognition, 2021, 118, 108043.	8.1	37

#	ARTICLE	IF	CITATIONS
91	VChunkJoin: An Efficient Algorithm for Edit Similarity Joins. IEEE Transactions on Knowledge and Data Engineering, 2013, 25, 1916-1929.	5.7	35
92	Exploring Consistent Preferences. , 2017, , .		35
93	Deliberate Attention Networks for Image Captioning. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33, 8320-8327.	4.9	35
94	Inductive Structure Consistent Hashing via Flexible Semantic Calibration. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32, 4514-4528.	11.3	35
95	An adaptive and dynamic dimensionality reduction method for high-dimensional indexing. VLDB Journal, 2007, 16, 219-234.	4.1	33
96	Instance optimal query processing in spatial networks. VLDB Journal, 2009, 18, 675-693.	4.1	33
97	A Large-scale RGB-D Database for Arbitrary-view Human Action Recognition. , 2018, , .		33
98	Structured Two-Stream Attention Network for Video Question Answering. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33, 6391-6398.	4.9	33
99	A Context Knowledge Map Guided Coarse-to-Fine Action Recognition. IEEE Transactions on Image Processing, 2020, 29, 2742-2752.	9.8	33
100	Distributed shortest path query processing on dynamic road networks. VLDB Journal, 2017, 26, 399-419.	4.1	32
101	Classification by Retrieval. , 2017, , .		32
102	Adaptively Attending to Visual Attributes and Linguistic Knowledge for Captioning. , 2017, , .		32
103	Adaptive Square Attack: Fooling Autonomous Cars With Adversarial Traffic Signs. IEEE Internet of Things Journal, 2021, 8, 6337-6347.	8.7	32
104	Radial Graph Convolutional Network for Visual Question Generation. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32, 1654-1667.	11.3	32
105	Adaptive Component Embedding for Domain Adaptation. IEEE Transactions on Cybernetics, 2021, 51, 3390-3403.	9.5	31
106	From Pixels to Objects: Cubic Visual Attention for Visual Question Answering. , 2018, , .		31
107	Scalable Zero-Shot Learning via Binary Visual-Semantic Embeddings. IEEE Transactions on Image Processing, 2019, 28, 3662-3674.	9.8	30
108	Attribute hashing for zero-shot image retrieval. , 2017, , .		29

#	ARTICLE	IF	CITATIONS
109	Perceptual Pyramid Adversarial Networks for Text-to-Image Synthesis. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33, 8312-8319.	4.9	29
110	Statistical Karyotype Analysis Using CNN and Geometric Optimization. IEEE Access, 2019, 7, 179445-179453.	4.2	29
111	Exploiting Web Images for Multi-Output Classification: From Category to Subcategories. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31, 1-13.	11.3	29
112	Highly-Economized Multi-view Binary Compression for Scalable Image Clustering. Lecture Notes in Computer Science, 2018, , 731-748.	1.3	29
113	Webly Supervised Fine-Grained Recognition: Benchmark Datasets and An Approach. , 2021, , .		29
114	Two Birds One Stone. , 2017, , .		28
115	Deep Collaborative Discrete Hashing With Semantic-Invariant Structure Construction. IEEE Transactions on Multimedia, 2021, 23, 1274-1286.	7.2	28
116	Online Near-Duplicate Video Clip Detection and Retrieval: An Accurate and Fast System. Proceedings - International Conference on Data Engineering, 2009, , .	0.0	25
117	Lightweight dynamic conditional GAN with pyramid attention for text-to-image synthesis. Pattern Recognition, 2021, 110, 107384.	8.1	25
118	Cross-Modal Hybrid Feature Fusion for Image-Sentence Matching. ACM Transactions on Multimedia Computing, Communications and Applications, 2021, 17, 1-23.	4.3	25
119	One-shot learning based pattern transition map for action early recognition. Signal Processing, 2018, 143, 364-370.	3.7	24
120	Learning Cross-Modal Common Representations by Privateâ€œShared Subspaces Separation. IEEE Transactions on Cybernetics, 2022, 52, 3261-3275.	9.5	24
121	Investigating the Bilateral Connections in Generative Zero-Shot Learning. IEEE Transactions on Cybernetics, 2022, 52, 8167-8178.	9.5	24
122	Data-Driven Estimation of Remaining Useful Lifetime and State of Charge for Lithium-Ion Battery. IEEE Transactions on Transportation Electrification, 2022, 8, 356-367.	7.8	24
123	Joint Feature Synthesis and Embedding: Adversarial Cross-Modal Retrieval Revisited. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44, 3030-3047.	13.9	24
124	Order-aware convolutional pooling for video based action recognition. Pattern Recognition, 2019, 91, 357-365.	8.1	23
125	Cross-Modal Dynamic Networks for Video Moment Retrieval With Text Query. IEEE Transactions on Multimedia, 2022, 24, 1221-1232.	7.2	23
126	Pseudo Transfer with Marginalized Corrupted Attribute for Zero-shot Learning. , 2018, , .		22

#	ARTICLE	IF	CITATIONS
127	What Machines See Is Not What They Get: Fooling Scene Text Recognition Models With Adversarial Text Images. , 2020, , .		22
128	Zero-shot Cross-modal Retrieval by Assembling AutoEncoder and Generative Adversarial Network. ACM Transactions on Multimedia Computing, Communications and Applications, 2021, 17, 1-17.	4.3	22
129	Tag Features for Geo-Aware Image Classification. IEEE Transactions on Multimedia, 2015, 17, 1058-1067.	7.2	21
130	Hierarchical Latent Concept Discovery for Video Event Detection. IEEE Transactions on Image Processing, 2017, 26, 2149-2162.	9.8	21
131	Surface k-NN Query Processing. , 2006, , .		20
132	Bidirectional Discrete Matrix Factorization Hashing for Image Search. IEEE Transactions on Cybernetics, 2020, 50, 4157-4168.	9.5	20
133	Statistical summarization of content features for fast near-duplicate video detection. , 2007, , .		19
134	Augmented keyword search on spatial entity databases. VLDB Journal, 2018, 27, 225-244.	4.1	19
135	Visual Spatial Attention Network for Relationship Detection. , 2018, , .		19
136	Alleviating Domain Shift via Discriminative Learning for Generalized Zero-Shot Learning. IEEE Transactions on Multimedia, 2022, 24, 1325-1337.	7.2	19
137	A multi-resolution surface distance model for k-NN query processing. VLDB Journal, 2008, 17, 1101-1119.	4.1	18
138	Instant Advertising in Mobile Peer-to-Peer Networks. Proceedings - International Conference on Data Engineering, 2009, , .	0.0	18
139	Automatic tagging by exploring tag information capability and correlation. World Wide Web, 2012, 15, 233-256.	4.0	18
140	Action-Centric Relation Transformer Network for Video Question Answering. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32, 63-74.	8.3	18
141	Batch Nearest Neighbor Search for Video Retrieval. IEEE Transactions on Multimedia, 2008, 10, 409-420.	7.2	17
142	Scalable Video Event Retrieval by Visual State Binary Embedding. IEEE Transactions on Multimedia, 2016, 18, 1590-1603.	7.2	17
143	Web Video Event Recognition by Semantic Analysis From Ubiquitous Documents. IEEE Transactions on Image Processing, 2016, 25, 5689-5701.	9.8	16
144	Salience-Guided Iterative Asymmetric Mutual Hashing for Fast Person Re-Identification. IEEE Transactions on Image Processing, 2021, 30, 7776-7789.	9.8	16

#	ARTICLE	IF	CITATIONS
145	One-shot Scene Graph Generation. , 2020, , .		16
146	MR-NET: Exploiting Mutual Relation for Visual Relationship Detection. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33, 8110-8117.	4.9	15
147	Kernel Attention Network for Single Image Super-Resolution. ACM Transactions on Multimedia Computing, Communications and Applications, 2020, 16, 1-15.	4.3	15
148	Locality condensation. , 2008, , .		14
149	Robust Cross-view Hashing for Multimedia Retrieval. IEEE Signal Processing Letters, 2016, 23, 893-897.	3.6	14
150	Examine before You Answer. , 2018, , .		14
151	Push & Pull: Transferable Adversarial Examples With Attentive Attack. IEEE Transactions on Multimedia, 2022, 24, 2329-2338.	7.2	14
152	Speed up interactive image retrieval. VLDB Journal, 2009, 18, 329-343.	4.1	13
153	Transfer tagging from image to video. , 2011, , .		13
154	Large Factor Image Super-Resolution With Cascaded Convolutional Neural Networks. IEEE Transactions on Multimedia, 2021, 23, 2172-2184.	7.2	13
155	Play and rewind: Context-aware video temporal action proposals. Pattern Recognition, 2020, 107, 107477.	8.1	13
156	Reducing bias to source samples for unsupervised domain adaptation. Neural Networks, 2021, 141, 61-71.	5.9	13
157	Correlated Features Synthesis and Alignment for Zero-shot Cross-modal Retrieval. , 2020, , .		13
158	Interclass-Relativity-Adaptive Metric Learning for Cross-Modal Matching and Beyond. IEEE Transactions on Multimedia, 2021, 23, 3073-3084.	7.2	12
159	Multi-Branch Networks for Video Super-Resolution With Dynamic Reconstruction Strategy. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31, 3954-3966.	8.3	12
160	Bilinear Optimized Product Quantization for Scalable Visual Content Analysis. IEEE Transactions on Image Processing, 2017, 26, 5057-5069.	9.8	11
161	MetaMixUp: Learning Adaptive Interpolation Policy of MixUp With Metalearning. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33, 3050-3064.	11.3	11
162	Fusing functional connectivity with network nodal information for sparse network pattern learning of functional brain networks. Information Fusion, 2021, 75, 131-139.	19.1	11

#	ARTICLE	IF	CITATIONS
163	Robust regression based face recognition with fast outlier removal. Multimedia Tools and Applications, 2016, 75, 12535-12546.	3.9	10
164	Fusion by synthesizing: A multi-view deep neural network for zero-shot recognition. Signal Processing, 2019, 164, 354-367.	3.7	10
165	Deep quantization generative networks. Pattern Recognition, 2020, 105, 107338.	8.1	10
166	Similarity preserving feature generating networks for zero-shot learning. Neurocomputing, 2020, 406, 333-342.	5.9	10
167	Residual Graph Convolutional Networks for Zero-Shot Learning. , 2019, , .		10
168	KTN: Knowledge Transfer Network for Multi-person DensePose Estimation. , 2020, , .		10
169	Challenges and techniques for effective and efficient similarity search in large video databases. Proceedings of the VLDB Endowment, 2008, 1, 1598-1603.	3.8	9
170	Efficient and Continuous Near-duplicate Video Detection. , 2010, , .		9
171	Quick identification of near-duplicate video sequences with cut signature. World Wide Web, 2012, 15, 355-382.	4.0	9
172	Attention Transfer (ANT) Network for View-invariant Action Recognition. , 2019, , .		9
173	Answer Again: Improving VQA with Cascaded-Answering Model. IEEE Transactions on Knowledge and Data Engineering, 2020, , 1-1.	5.7	9
174	View-invariant action recognition via Unsupervised Attention Transfer (UANT). Pattern Recognition, 2021, 113, 107807.	8.1	9
175	Indexing and Integrating Multiple Features for WWW Images. World Wide Web, 2006, 9, 343-364.	4.0	8
176	On the Influence Propagation of Web Videos. IEEE Transactions on Knowledge and Data Engineering, 2014, 26, 1961-1973.	5.7	7
177	Relation Regularized Scene Graph Generation. IEEE Transactions on Cybernetics, 2022, 52, 5961-5972.	9.5	7
178	Indexing Text and Visual Features for WWW Images. Lecture Notes in Computer Science, 2005, , 885-899.	1.3	7
179	Lab2Pix: Label-Adaptive Generative Adversarial Network for Unsupervised Image Synthesis. , 2020, , .		7
180	Dynamic Batch Nearest Neighbor Search in Video Retrieval. , 2007, , .		6

#	ARTICLE	IF	CITATIONS
181	Compact Indexing and Judicious Searching for Billion-Scale Microblog Retrieval. ACM Transactions on Information Systems, 2017, 35, 1-24.	4.9	6
182	Structured Learning of Binary Codes with Column Generation for Optimizing Ranking Measures. International Journal of Computer Vision, 2017, 123, 287-308.	15.6	6
183	Efficient Binary Coding for Subspace-based Query-by-Image Video Retrieval. , 2017, , .		6
184	Leveraging Weak Semantic Relevance for Complex Video Event Classification. , 2017, , .		6
185	Learning binary codes with local and inner data structure. Neurocomputing, 2018, 282, 32-41.	5.9	6
186	One-Shot Image-to-Image Translation via Part-Global Learning With a Multi-Adversarial Framework. IEEE Transactions on Multimedia, 2022, 24, 480-491.	7.2	6
187	Deep Semantic Indexing Using Convolutional Localization Network with Region-Based Visual Attention for Image Database. Lecture Notes in Computer Science, 2017, , 261-272.	1.3	6
188	Temporal Denoising Mask Synthesis Network for Learning Blind Video Temporal Consistency. , 2020, , .		6
189	Interactive near-duplicate video retrieval and detection. , 2009, , .		5
190	Continuous Proximity Detection via Predictive Safe Region Construction. , 2018, , .		5
191	Combine Early and Late Fusion Together: A Hybrid Fusion Framework for Image-Text Matching. , 2021, , .		5
192	Hybrid Fusion with Intra- and Cross-Modality Attention for Image-Recipe Retrieval. , 2021, , .		5
193	Max-margin adaptive model for complex video pattern recognition. Multimedia Tools and Applications, 2015, 74, 505-521.	3.9	4
194	Cumulative Nets for Edge Detection. , 2018, , .		4
195	Embedding and predicting the event at early stage. World Wide Web, 2019, 22, 1055-1074.	4.0	4
196	Multi-scale Dynamic Network for Temporal Action Detection. , 2021, , .		4
197	Web-based SBRL method of multimedia tools for computer-aided drawing. Multimedia Tools and Applications, 2019, 78, 799-816.	3.9	3
198	Dimensionality Reduction. , 2009, , 843-846.		3

#	ARTICLE	IF	CITATIONS
199	3D Self-Attention for Unsupervised Video Quantization. , 2020, , .		3
200	ICICLE: A semantic-based retrieval system for WWW images. Multimedia Systems, 2006, 11, 438-454.	4.7	2
201	A Multiresolution Terrain Model for Efficient Visualization Query Processing. IEEE Transactions on Knowledge and Data Engineering, 2006, 18, 1382-1396.	5.7	2
202	High-dimensional indexing with oriented cluster representation for multimedia databases. , 2009, , .		2
203	Robust Semantic Video Indexing by Harvesting Web Images. Lecture Notes in Computer Science, 2013, , 70-80.	1.3	2
204	Personalized query evaluation in ring-based P2P networks. Information Sciences, 2013, 220, 463-482.	6.9	2
205	Towards Accurate Georeferenced Video Search With Camera Field of View Modeling. IEEE Transactions on Circuits and Systems for Video Technology, 2019, 29, 1844-1855.	8.3	2
206	Jointly Learning Attentions with Semantic Cross-Modal Correlation for Visual Question Answering. Lecture Notes in Computer Science, 2017, , 248-260.	1.3	2
207	Effective Data Density Estimation in Ring-Based P2P Networks. , 2012, , .		1
208	CC-LSTM: Cross and Conditional Long-Short Time Memory for Video Captioning. Lecture Notes in Computer Science, 2021, , 353-365.	1.3	1
209	Graph-based variational auto-encoder for generalized zero-shot learning. , 2021, , .		1
210	Efficient Histogram-Based Similarity Search in Ultra-High Dimensional Space. Lecture Notes in Computer Science, 2011, , 1-15.	1.3	1
211	Multimodal Disentanglement Variational AutoEncoders for Zero-Shot Cross-Modal Retrieval. , 2022, , .		1
212	SaveRF: Towards Efficient Relevance Feedback Search. , 2006, , .		0
213	Capture local information in shape representation. Multimedia Tools and Applications, 2007, 33, 157-174.	3.9	0
214	Modeling Two-Stream Correspondence for Visual Sound Separation. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32, 3291-3302.	8.3	0
215	Efficient Supervised Hashing via Exploring Local and Inner Data Structure. Lecture Notes in Computer Science, 2017, , 98-109.	1.3	0
216	Video Sequence Indexing. , 2017, , 1-4.		0

#	ARTICLE	IF	CITATIONS
217	Video Sequence Indexing. , 2018, , 4431-4434.		0
218	Dimensionality Reduction. , 2018, , 1106-1107.		0
219	Structure-Aware Semantic-Aligned Network for Universal Cross-Domain Retrieval. , 2022, , .		0