

# Niculae Sebe

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/5553001/publications.pdf>

Version: 2024-02-01

269  
papers

16,805  
citations

38660

50  
h-index

30848

102  
g-index

272  
all docs

272  
docs citations

272  
times ranked

11078  
citing authors

#	ARTICLE	IF	CITATIONS
1	Content-based multimedia information retrieval. ACM Transactions on Multimedia Computing, Communications and Applications, 2006, 2, 1-19.	3.0	1,232
2	Facial expression recognition from video sequences: temporal and static modeling. Computer Vision and Image Understanding, 2003, 91, 160-187.	3.0	715
3	Multimodal human-computer interaction: A survey. Computer Vision and Image Understanding, 2007, 108, 116-134.	3.0	693
4	A Survey on Learning to Hash. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40, 769-790.	9.7	650
5	Detecting anomalous events in videos by learning deep representations of appearance and motion. Computer Vision and Image Understanding, 2017, 156, 117-127.	3.0	345
6	Learning Deep Representations of Appearance and Motion for Anomalous Event Detection. , 2015, , .		310
7	Deformable GANs for Pose-Based Human Image Generation. , 2018, , .		309
8	Abnormal event detection in videos using generative adversarial nets. , 2017, , .		305
9	ASCERTAIN: Emotion and Personality Recognition Using Commercial Sensors. IEEE Transactions on Affective Computing, 2018, 9, 147-160.	5.7	295
10	Multi-scale Continuous CRFs as Sequential Deep Networks for Monocular Depth Estimation. , 2017, , .		277
11	Binary neural networks: A survey. Pattern Recognition, 2020, 105, 107281.	5.1	276
12	PAD-Net: Multi-tasks Guided Prediction-and-Distillation Network for Simultaneous Depth Estimation and Scene Parsing. , 2018, , .		267
13	AMIGOS: A Dataset for Affect, Personality and Mood Research on Individuals and Groups. IEEE Transactions on Affective Computing, 2021, 12, 479-493.	5.7	264
14	Combining Head Pose and Eye Location Information for Gaze Estimation. IEEE Transactions on Image Processing, 2012, 21, 802-815.	6.0	255
15	Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces.. Emotion, 2009, 9, 260-264.	1.5	238
16	DECAF: MEG-Based Multimodal Database for Decoding Affective Physiological Responses. IEEE Transactions on Affective Computing, 2015, 6, 209-222.	5.7	236
17	Structured Attention Guided Convolutional Neural Fields for Monocular Depth Estimation. , 2018, , .		212
18	Semisupervised learning of classifiers: theory, algorithms, and their application to human-computer interaction. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26, 1553-1566.	9.7	210

#	ARTICLE	IF	CITATIONS
19	Self-Adaptive Matrix Completion for Heart Rate Estimation from Face Videos under Realistic Conditions. , 2016, , .		204
20	Weakly Supervised Photo Cropping. IEEE Transactions on Multimedia, 2014, 16, 94-107.	5.2	198
21	Dynamic Affinity Graph Construction for Spectral Clustering Using Multiple Features. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29, 6323-6332.	7.2	191
22	Feature Selection for Multimedia Analysis by Sharing Information Among Multiple Tasks. IEEE Transactions on Multimedia, 2013, 15, 661-669.	5.2	187
23	Pattern-Affinitive Propagation Across Depth, Surface Normal and Semantic Segmentation. , 2019, , .		184
24	Web Image Annotation Via Subspace-Sparsity Collaborated Feature Selection. IEEE Transactions on Multimedia, 2012, 14, 1021-1030.	5.2	171
25	Multitask Linear Discriminant Analysis for View Invariant Action Recognition. IEEE Transactions on Image Processing, 2014, 23, 5599-5611.	6.0	156
26	Semisupervised Feature Selection via Spline Regression for Video Semantic Recognition. IEEE Transactions on Neural Networks and Learning Systems, 2015, 26, 252-264.	7.2	156
27	Group Consistent Similarity Learning via Deep CRF for Person Re-identification. , 2018, , .		156
28	Quantization-based hashing: a general framework for scalable image and video retrieval. Pattern Recognition, 2018, 75, 175-187.	5.1	151
29	Multi-Feature Fusion via Hierarchical Regression for Multimedia Analysis. IEEE Transactions on Multimedia, 2013, 15, 572-581.	5.2	146
30	Affective multimodal human-computer interaction. , 2005, , .		140
31	Recurrent Face Aging. , 2016, , .		135
32	Learning Cross-Modal Deep Representations for Robust Pedestrian Detection. , 2017, , .		134
33	Egocentric Daily Activity Recognition via Multitask Clustering. IEEE Transactions on Image Processing, 2015, 24, 2984-2995.	6.0	133
34	Discriminating Joint Feature Analysis for Multimedia Data Understanding. IEEE Transactions on Multimedia, 2012, 14, 1662-1672.	5.2	129
35	Multi-Channel Attention Selection GAN With Cascaded Semantic Guidance for Cross-View Image Translation. , 2019, , .		123
36	Multimedia Event Detection Using A Classifier-Specific Intermediate Representation. IEEE Transactions on Multimedia, 2013, 15, 1628-1637.	5.2	121

#	ARTICLE	IF	CITATIONS
37	An Eye Fixation Database for Saliency Detection in Images. Lecture Notes in Computer Science, 2010, , 30-43.	1.0	112
38	Plug-and-Play CNN for Crowd Motion Analysis: An Application in Abnormal Event Detection. , 2018, , .		111
39	Optimized Graph Learning Using Partial Tags and Multiple Features for Image and Video Annotation. IEEE Transactions on Image Processing, 2016, 25, 4999-5011.	6.0	107
40	SALSA: A Novel Dataset for Multimodal Group Behavior Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38, 1707-1720.	9.7	107
41	MULTIMODAL EMOTION RECOGNITION. , 2005, , 387-409.		106
42	Unsupervised Adversarial Depth Estimation Using Cycled Generative Networks. , 2018, , .		106
43	Event Oriented Dictionary Learning for Complex Event Detection. IEEE Transactions on Image Processing, 2015, 24, 1867-1878.	6.0	105
44	A Multi-Task Learning Framework for Head Pose Estimation under Target Motion. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38, 1070-1083.	9.7	104
45	The Death and Life of Great Italian Cities. , 2016, , .		103
46	A Distance-Computation-Free Search Scheme for Binary Code Databases. IEEE Transactions on Multimedia, 2016, 18, 484-495.	5.2	101
47	Spatio-Temporal Attention Networks for Action Recognition and Detection. IEEE Transactions on Multimedia, 2020, 22, 2990-3001.	5.2	98
48	Looking at the viewer: analysing facial activity to detect personal highlights of multimedia contents. Multimedia Tools and Applications, 2011, 51, 505-523.	2.6	97
49	Human-centered computing. , 2006, , .		90
50	No Matter Where You Are: Flexible Graph-Guided Multi-task Learning for Multi-view Head Pose Classification under Target Motion. , 2013, , .		90
51	Unsupervised Domain Adaptation Using Feature-Whitening and Consensus Loss. , 2019, , .		89
52	Refine and Distill: Exploiting Cycle-Inconsistency and Knowledge Distillation for Unsupervised Monocular Depth Estimation. , 2019, , .		88
53	Fast and robust dynamic hand gesture recognition via key frames extraction and feature fusion. Neurocomputing, 2019, 331, 424-433.	3.5	88
54	Comparing salient point detectors. Pattern Recognition Letters, 2003, 24, 89-96.	2.6	82

#	ARTICLE	IF	CITATIONS
55	We are not All Equal. , 2014, , .		82
56	XingGAN for Person Image Generation. Lecture Notes in Computer Science, 2020, , 717-734.	1.0	81
57	Connecting Meeting Behavior with Extraversionâ€”A Systematic Study. IEEE Transactions on Affective Computing, 2012, 3, 443-455.	5.7	78
58	Attention-Guided Generative Adversarial Networks for Unsupervised Image-to-Image Translation. , 2019, , .		78
59	Reverse Perspective Network for Perspective-Aware Object Counting. , 2020, , .		78
60	Please, tell me about yourself. , 2011, , .		76
61	Learning Personalized Models for Facial Expression Analysis and Gesture Recognition. IEEE Transactions on Multimedia, 2016, 18, 775-788.	5.2	74
62	Knowledge Adaptation with Partially Shared Features for Event Detection Using Few Exemplars. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36, 1789-1802.	9.7	73
63	Video classification with Densely extracted HOG/HOF/MBH features: an evaluation of the accuracy/computational efficiency trade-off. International Journal of Multimedia Information Retrieval, 2015, 4, 33-44.	3.6	73
64	Cross-Domain Car Detection Using Unsupervised Image-to-Image Translation: From Day to Night. , 2019, , .		72
65	Recognizing Emotions from Abstract Paintings Using Non-Linear Matrix Completion. , 2016, , .		71
66	AttentionGAN: Unpaired Image-to-Image Translation Using Attention-Guided Generative Adversarial Networks. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34, 1972-1987.	7.2	71
67	How to complete performance graphs in content-based image retrieval: add generality and normalize scope. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27, 245-251.	9.7	68
68	Distance Learning for Similarity Estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30, 451-462.	9.7	67
69	Self Paced Deep Learning for Weakly Supervised Object Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41, 712-725.	9.7	67
70	Where am I in the dark: Exploring active transfer learning on the use of indoor localization based on thermal imaging. Neurocomputing, 2016, 173, 83-92.	3.5	66
71	The Many Shades of Negativity. IEEE Transactions on Multimedia, 2017, 19, 1558-1568.	5.2	64
72	Transformer-Based Attention Networks for Continuous Pixel-Wise Prediction. , 2021, , .		64

#	ARTICLE	IF	CITATIONS
73	Harnessing Lab Knowledge for Real-World Action Recognition. International Journal of Computer Vision, 2014, 109, 60-73.	10.9	63
74	Curriculum Learning: A Survey. International Journal of Computer Vision, 2022, 130, 1526-1565.	10.9	63
75	In the eye of the beholder. , 2012, , .		61
76	Money walks. , 2014, , .		61
77	Analyzing Free-standing Conversational Groups. , 2015, , .		61
78	Cycle In Cycle Generative Adversarial Networks for Keypoint-Guided Image Generation. , 2019, , .		61
79	Guest Editors' Introduction: Human-Centered Computing--Toward a Human Revolution. Computer, 2007, 40, 30-34.	1.2	60
80	Knowledge adaptation for ad hoc multimedia event detection with few exemplars. , 2012, , .		58
81	Optimal graph learning with partial tags and multiple features for image and video annotation. , 2015, , .		57
82	On the relationship between head pose, social attention and personality prediction for unstructured and dynamic group interactions. , 2013, , .		55
83	The more the merrier: Analysing the affect of a group of people in images. , 2015, , .		55
84	Collaborative Sparse Coding for Multiview Action Recognition. IEEE MultiMedia, 2016, 23, 80-87.	1.5	55
85	What your Facebook Profile Picture Reveals about your Personality. , 2017, , .		55
86	GestureGAN for Hand Gesture-to-Gesture Translation in the Wild. , 2018, , .		54
87	Exploiting the entire feature space with sparsity for automatic image annotation. , 2011, , .		53
88	Unsupervised Anomaly Detection and Localization Based on Deep Spatiotemporal Translation Network. IEEE Access, 2020, 8, 50312-50329.	2.6	53
89	Complex Event Detection via Multi-source Video Attributes. , 2013, , .		52
90	Exploring Transfer Learning Approaches for Head Pose Classification from Multi-view Surveillance Images. International Journal of Computer Vision, 2014, 109, 146-167.	10.9	52

#	ARTICLE	IF	CITATIONS
91	Monocular Depth Estimation Using Multi-Scale Continuous CRFs as Sequential Deep Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41, 1426-1440.	9.7	51
92	When Dictionary Learning Meets Deep Learning: Deep Dictionary Learning and Coding Network for Image Recognition With Limited Data. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32, 2129-2141.	7.2	50
93	Metric-Learning-Based Deep Hashing Network for Content-Based Retrieval of Remote Sensing Images. IEEE Geoscience and Remote Sensing Letters, 2021, 18, 226-230.	1.4	50
94	Category Specific Dictionary Learning for Attribute Specific Feature Selection. IEEE Transactions on Image Processing, 2016, 25, 1465-1478.	6.0	49
95	Emotion modulates eye movement patterns and subsequent memory for the gist and details of movie scenes. Journal of Vision, 2014, 14, 31-31.	0.1	48
96	GLocal tells you more: Coupling GLocal structural for feature selection with sparsity for image and video classification. Computer Vision and Image Understanding, 2014, 124, 99-109.	3.0	48
97	Flexible Manifold Learning With Optimal Graph for Image and Video Representation. IEEE Transactions on Image Processing, 2018, 27, 2664-2675.	6.0	47
98	Dual Attention GANs for Semantic Image Synthesis. , 2020, , .		45
99	Deep appearance and motion learning for egocentric activity recognition. Neurocomputing, 2018, 275, 438-447.	3.5	44
100	Feature Weighting via Optimal Thresholding for Video Analysis. , 2013, , .		42
101	Regressing a 3D Face Shape from a Single Image. , 2015, , .		42
102	Image Attribute Adaptation. IEEE Transactions on Multimedia, 2014, 16, 1115-1126.	5.2	41
103	Multimodal interfaces: Challenges and perspectives. Journal of Ambient Intelligence and Smart Environments, 2009, 1, 23-30.	0.8	39
104	Faved! Biometrics: Tell Me Which Image You Like and I'll Tell You Who You Are. IEEE Transactions on Information Forensics and Security, 2014, 9, 364-374.	4.5	39
105	Sparse Color Interest Points for Image Retrieval and Object Categorization. IEEE Transactions on Image Processing, 2012, 21, 2681-2692.	6.0	38
106	Event-based media processing and analysis: A survey of the literature. Image and Vision Computing, 2016, 53, 3-19.	2.7	37
107	Spatio-Temporal Vector of Locally Max Pooled Features for Action Recognition in Videos. , 2017, , .		36
108	Multi-task linear discriminant analysis for multi-view action recognition. , 2013, , .		34

#	ARTICLE	IF	CITATIONS
109	Every Smile is Unique: Landmark-Guided Diverse Smile Generation. , 2018, , .		34
110	Unsupervised Domain Adaptation for Personalized Facial Emotion Recognition. , 2014, , .		33
111	Are Safer Looking Neighborhoods More Lively?. , 2016, , .		33
112	Who's Afraid of Itten. , 2015, , .		33
113	A Prototype Learning Framework Using EMD: Application to Complex Scenes Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35, 513-526.	9.7	32
114	Supervised Hashing with Pseudo Labels for Scalable Multimedia Retrieval. , 2015, , .		32
115	Affective Analysis of Professional and Amateur Abstract Paintings Using Statistical Analysis and Art Theory. ACM Transactions on Interactive Intelligent Systems, 2015, 5, 1-27.	2.6	32
116	Embedding Perspective Analysis Into Multi-Column Convolutional Neural Network for Crowd Counting. IEEE Transactions on Image Processing, 2021, 30, 1395-1407.	6.0	32
117	The SocioMetric Badges Corpus: A Multilevel Behavioral Dataset for Social Behavior in Complex Organizations. , 2012, , .		31
118	Joint Attributes and Event Analysis for Multimedia Event Detection. IEEE Transactions on Neural Networks and Learning Systems, 2017, 29, 1-10.	7.2	31
119	Viewpoint and Scale Consistency Reinforcement for UAV Vehicle Re-Identification. International Journal of Computer Vision, 2021, 129, 719-735.	10.9	31
120	The S-HOCK dataset: Analyzing crowds at the stadium. , 2015, , .		30
121	Unified Generative Adversarial Networks for Controllable Image-to-Image Translation. IEEE Transactions on Image Processing, 2020, 29, 8916-8929.	6.0	30
122	Implicit User-centric Personality Recognition Based on Physiological Responses to Emotional Videos. , 2015, , .		29
123	Multi-View Spatial Attention Embedding for Vehicle Re-Identification. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31, 816-827.	5.6	29
124	A modified vector of locally aggregated descriptors approach for fast video classification. Multimedia Tools and Applications, 2016, 75, 9045-9072.	2.6	28
125	Compact Image Fingerprint Via Multiple Kernel Hashing. IEEE Transactions on Multimedia, 2015, 17, 1006-1018.	5.2	27
126	Inference of personality traits and affect schedule by analysis of spontaneous reactions to affective videos. , 2015, , .		26



#	ARTICLE	IF	CITATIONS
127	Joint Representation Learning and Keypoint Detection for Cross-View Geo-Localization. IEEE Transactions on Image Processing, 2022, 31, 3780-3792.	6.0	26
128	The First 3D Face Alignment in the Wild (3DFAW) Challenge. Lecture Notes in Computer Science, 2016, , 511-520.	1.0	24
129	Efficient human action recognition using histograms of motion gradients and VLAD with descriptor shape information. Multimedia Tools and Applications, 2017, 76, 22445-22472.	2.6	24
130	Building descriptive and discriminative visual codebook for large-scale image applications. Multimedia Tools and Applications, 2011, 51, 441-477.	2.6	23
131	Can computers learn from humans to see better?. , 2011, , .		23
132	UX_Mate. , 2012, , .		23
133	Distributional semantics with eyes. , 2012, , .		23
134	Robust Real-Time Extreme Head Pose Estimation. , 2014, , .		23
135	Perceptual Attributes Optimization for Multivideo Summarization. IEEE Transactions on Cybernetics, 2016, 46, 2991-3003.	6.2	23
136	Deep and fast: Deep learning hashing with semi-supervised graph construction. Image and Vision Computing, 2016, 55, 101-108.	2.7	23
137	Semantic-Fusion Gans for Semi-Supervised Satellite Image Classification. , 2018, , .		23
138	Real-life violent social interaction detection. , 2015, , .		22
139	Histograms of Motion Gradients for real-time video classification. , 2016, , .		22
140	Multimodal Personality Recognition in Collaborative Goal-Oriented Tasks. IEEE Transactions on Multimedia, 2016, 18, 659-673.	5.2	22
141	Automatic Group Affect Analysis in Images via Visual Attribute and Feature Networks. , 2018, , .		22
142	What Are You Looking at?. International Journal of Computer Vision, 2012, 98, 324-334.	10.9	21
143	Spatio-Temporal VLAD Encoding for Human Action Recognition in Videos. Lecture Notes in Computer Science, 2017, , 365-378.	1.0	21
144	Appearance and Pose-Conditioned Human Image Generation Using Deformable GANs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43, 1156-1171.	9.7	21

#	ARTICLE	IF	CITATIONS
145	A proto-object-based computational model for visual saliency. <i>Journal of Vision</i> , 2013, 13, 27.	0.1	20
146	Budget-Aware Adapters for Multi-Domain Learning. , 2019, , .		20
147	Tell Me What You Like and I'll Tell You What You Are: Discriminating Visual Preferences on Flickr Data. <i>Lecture Notes in Computer Science</i> , 2013, , 45-56.	1.0	20
148	SF-UDA <sup>3D</sup> : Source-Free Unsupervised Domain Adaptation for LiDAR-Based 3D Object Detection. , 2020, , .		19
149	Putting the pieces together. , 2010, , .		18
150	How to Make an Image More Memorable?. , 2017, , .		18
151	Attribute-Guided Sketch Generation. , 2019, , .		18
152	Localize Me Anywhere, Anytime: A Multi-task Point-Retrieval Approach. , 2015, , .		17
153	Indoor localization via multi-view images and videos. <i>Computer Vision and Image Understanding</i> , 2017, 161, 145-160.	3.0	17
154	Progressive Fusion for Unsupervised Binocular Depth Estimation Using Cycled Networks. <i>IEEE Transactions on Pattern Analysis and Machine Intelligence</i> , 2020, 42, 2380-2395.	9.7	17
155	Personalization in multimedia retrieval: A survey. <i>Multimedia Tools and Applications</i> , 2011, 51, 247-277.	2.6	16
156	User-centric Affective Video Tagging from MEG and Peripheral Physiological Responses. , 2013, , .		16
157	Memory efficient large-scale image-based localization. <i>Multimedia Tools and Applications</i> , 2015, 74, 479-503.	2.6	16
158	A Quality Adaptive Multimodal Affect Recognition System for User-Centric Multimedia Indexing. , 2016, , .		16
159	Viraliency: Pooling Local Virality. , 2017, , .		16
160	Depression Severity Estimation from Multiple Modalities. , 2018, , .		16
161	FontMatcher. , 2018, , .		16
162	Recurrent Face Aging with Hierarchical AutoRegressive Memory. <i>IEEE Transactions on Pattern Analysis and Machine Intelligence</i> , 2019, 41, 654-668.	9.7	16

#	ARTICLE	IF	CITATIONS
163	Exploiting sample correlation for crowd counting with multi-expert network. , 2021, , .		16
164	Learning to Group Objects. , 2014, , .		15
165	Facial expression recognition under a wide range of head poses. , 2015, , .		15
166	Fisher Kernel Temporal Variation-based Relevance Feedback for video retrieval. Computer Vision and Image Understanding, 2016, 143, 38-51.	3.0	15
167	Dual Generator Generative Adversarial Networks for Multi-domain Image-to-Image Translation. Lecture Notes in Computer Science, 2019, , 3-21.	1.0	15
168	Time matters!. , 2013, , .		14
169	Joint Graph Learning and Video Segmentation via Multiple Cues and Topology Calibration. , 2016, , .		14
170	Emotion recognition in the wild. Journal on Multimodal User Interfaces, 2016, 10, 95-97.	2.0	14
171	Viewpoint-Consistent 3D Face Alignment. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40, 2250-2264.	9.7	14
172	Deep Metric and Hash-Code Learning for Content-Based Retrieval of Remote Sensing Images. , 2018, , .		14
173	Cross-Paced Representation Learning With Partial Curricula for Sketch-Based Image Retrieval. IEEE Transactions on Image Processing, 2018, 27, 4410-4421.	6.0	14
174	Expression Conditional Gan for Facial Expression-to-Expression Translation. , 2019, , .		14
175	Predicting Group Cohesiveness in Images. , 2019, , .		14
176	Knowing Where I Am: Exploiting Multi-Task Learning for Multi-view Indoor Image-based Localization. , 2014, , .		14
177	Video anomaly detection using deep residual-spatiotemporal translation network. Pattern Recognition Letters, 2022, 155, 143-150.	2.6	14
178	CV-C3D: Action Recognition on Compressed Videos with Convolutional 3D Networks. , 2019, , .		13
179	Multimodal Engagement Classification for Affective Cinema. , 2013, , .		12
180	GLocal structural feature selection with sparsity for multimedia data understanding. , 2013, , .		12

#	ARTICLE	IF	CITATIONS
181	Clustered Multi-task Linear Discriminant Analysis for View Invariant Color-Depth Action Recognition. , 2014, , .		12
182	Recurrent Convolutional Shape Regression. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40, 2569-2582.	9.7	12
183	Learning How to Smile: Expression Video Generation With Conditional Adversarial Recurrent Nets. IEEE Transactions on Multimedia, 2020, 22, 2808-2819.	5.2	12
184	We are not equally negative. , 2013, , .		11
185	Unsupervised Tube Extraction Using Transductive Learning and Dense Trajectories. , 2015, , .		11
186	Academic Coupled Dictionary Learning for Sketch-based Image Retrieval. , 2016, , .		11
187	Total Generate: Cycle in Cycle Generative Adversarial Networks for Generating Human Faces, Hands, Bodies, and Natural Scenes. IEEE Transactions on Multimedia, 2022, 24, 2963-2974.	5.2	11
188	Dual In-painting Model for Unsupervised Gaze Correction and Animation in the Wild. , 2020, , .		11
189	Cross-view panorama image synthesis with progressive attention GANs. Pattern Recognition, 2022, 131, 108884.	5.1	11
190	Evaluating Multi-task Learning for Multi-view Head-Pose Classification in Interactive Environments. , 2014, , .		10
191	Boosting VLAD with double assignment using deep features for action recognition in videos. , 2016, , .		10
192	The S-Hock dataset: A new benchmark for spectator crowd analysis. Computer Vision and Image Understanding, 2017, 159, 47-58.	3.0	10
193	Structured Coupled Generative Adversarial Networks for Unsupervised Monocular Depth Estimation. , 2019, , .		10
194	PET: An eye-tracking dataset for animal-centric Pascal object classes. , 2015, , .		9
195	Increasing Image Memorability with Neural Style Transfer. ACM Transactions on Multimedia Computing, Communications and Applications, 2019, 15, 1-22.	3.0	9
196	Multimodal behavior analysis in the wild: An introduction. , 2019, , 1-8.		9
197	TriGAN: image-to-image translation for multi-source domain adaptation. Machine Vision and Applications, 2021, 32, 1.	1.7	9
198	Layout-to-Image Translation With Double Pooling Generative Adversarial Networks. IEEE Transactions on Image Processing, 2021, 30, 7903-7913.	6.0	9

#	ARTICLE	IF	CITATIONS
199	Daily Living Activities Recognition via Efficient High and Low Level Cues Combination and Fisher Kernel Representation. Lecture Notes in Computer Science, 2013, , 431-441.	1.0	9
200	Visual Gaze Estimation by Joint Head and Eye Information. , 2010, , .		8
201	Computer vision for ambient intelligence. Journal of Ambient Intelligence and Smart Environments, 2011, 3, 187-191.	0.8	8
202	Fisher kernel based relevance feedback for multimodal video retrieval. , 2013, , .		8
203	Attribute Guided Dictionary Learning. , 2015, , .		8
204	Brain and music: Music genre classification using brain signals. , 2016, , .		8
205	Anomaly Event Detection Using Generative Adversarial Network for Surveillance Videos. , 2019, , .		8
206	Gesture-to-Gesture Translation in the Wild via Category-Independent Conditional Maps. , 2019, , .		8
207	On the Eigenvalues of Global Covariance Pooling for Fine-grained Visual Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, , 1-1.	9.7	8
208	Contextual Modeling of Personality States' Dynamics in Face-to-Face Interactions. , 2011, , .		7
209	Deep traffic light detection by overlaying synthetic context on arbitrary natural images. Computers and Graphics, 2021, 94, 76-86.	1.4	7
210	Relation Regularized Scene Graph Generation. IEEE Transactions on Cybernetics, 2022, 52, 5961-5972.	6.2	7
211	Motion-supervised Co-Part Segmentation. , 2021, , .		7
212	Exploiting Sparse Representations for Robust Analysis of Noisy Complex Video Scenes. Lecture Notes in Computer Science, 2012, , 199-213.	1.0	7
213	An Adaptation Framework for Head-Pose Classification in Dynamic Multi-view Scenarios. Lecture Notes in Computer Science, 2013, , 652-666.	1.0	7
214	Probabilistic Graph Attention Network with Conditional Kernels for Pixel-Wise Prediction. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, PP, 1-1.	9.7	7
215	Classifier-specific intermediate representation for multimedia tasks. , 2012, , .		6
216	Nobody likes Mondays. , 2013, , .		6

#	ARTICLE	IF	CITATIONS
217	You Talkin' to Me?. , 2014, , .		6
218	Temporal Dropout of Changes Approach to Convolutional Learning of Spatio-Temporal Features. , 2014, , .		6
219	Beyond Bag-of-Words: Fast video classification with Fisher Kernel Vector of Locally Aggregated Descriptors. , 2015, , .		6
220	Active domain adaptation with noisy labels for multimedia analysis. World Wide Web, 2016, 19, 199-215.	2.7	6
221	Cross-domain object detection using unsupervised image translation. Expert Systems With Applications, 2022, 192, 116334.	4.4	6
222	Decoding affect in videos employing the MEG brain signal. , 2013, , .		5
223	It's all about habits: Exploiting multi-task clustering for activities of daily living analysis. , 2014, , .		5
224	The Mystery of Faces. , 2014, , .		5
225	FaceCept3D: Real Time 3D Face Tracking and Analysis. , 2015, , .		5
226	Simple, Efficient and Effective Encodings of Local Deep Features for Video Action Recognition. , 2017, , .		5
227	Deep Micro-Dictionary Learning and Coding Network. , 2019, , .		5
228	Movie Genre Classification by Exploiting MEG Brain Signals. Lecture Notes in Computer Science, 2015, , 683-693.	1.0	5
229	Impact of Facial Landmark Localization on Facial Expression Recognition. IEEE Transactions on Affective Computing, 2023, 14, 1267-1279.	5.7	5
230	Emotional Valence Recognition, Analysis of Saliency and Eye Movements. , 2014, , .		4
231	Multiple Features But Few Labels?. , 2014, , .		4
232	Computational Modeling of Affective Qualities of Abstract Paintings. IEEE MultiMedia, 2016, 23, 44-54.	1.5	4
233	Recognizing Daily Activities from First-Person Videos with Multi-task Clustering. Lecture Notes in Computer Science, 2015, , 522-537.	1.0	4
234	Unsupervised Domain Adaptation Using Full-Feature Whitening and Colouring. Lecture Notes in Computer Science, 2019, , 225-236.	1.0	4

#	ARTICLE	IF	CITATIONS
235	Simultaneous Ground Metric Learning and Matrix Factorization with Earth Mover's Distance. , 2014, , .		3
236	A Multi-task Learning Framework for Time-continuous Emotion Estimation from Crowd Annotations. , 2014, , .		3
237	Guest Editorial: Big Media Data: Understanding, Search, and Mining. IEEE Transactions on Big Data, 2015, 1, 82-83.	4.4	3
238	Guest Editorsâ€™ Introduction to the Special Section on Learning with Shared Information for Computer Vision and Multimedia Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40, 1029-1031.	9.7	3
239	Automatic Prediction of Group Cohesiveness in Images. IEEE Transactions on Affective Computing, 2022, 13, 1677-1690.	5.7	3
240	Guest Editorial: Humanâ€™Computer Interaction: Real-Time Vision Aspects of Natural User Interfaces. International Journal of Computer Vision, 2013, 101, 401-402.	10.9	2
241	Unified Dictionary Learning and Region Tagging with Hierarchical Sparse Representation. Computer Vision and Image Understanding, 2013, 117, 934-946.	3.0	2
242	Multi-Paced Dictionary Learning for cross-domain retrieval and recognition. , 2016, , .		2
243	Sparse-coded cross-domain adaptation from the visual to the brain domain. , 2016, , .		2
244	MUSA2. , 2017, , .		2
245	New Signals in Multimedia Systems and Applications. IEEE MultiMedia, 2018, 25, 12-13.	1.5	2
246	Emotions in Abstract Art: Does Texture Matter?. Lecture Notes in Computer Science, 2015, , 671-682.	1.0	2
247	Guest Editorial: Introduction to the Special Section on Fine-Grained Visual Categorization. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44, 560-562.	9.7	2
248	Minimizing dataset bias: Discriminative multi-task sparse coding through shared subspace learning for image classification. , 2014, , .		1
249	Guest Editorial: Special Section on Socio-Mobile Media Analysis and Retrieval. IEEE Transactions on Multimedia, 2014, 16, 586-587.	5.2	1
250	Cluster encoding for modelling temporal variation in video. , 2015, , .		1
251	Emerging Topics in Learning from Noisy and Missing Data. , 2016, , .		1
252	SALSA: A Multimodal Dataset for the Automated Analysis of Free-Standing Social Interactions. , 2017, , 321-340.		1

#	ARTICLE	IF	CITATIONS
253	Recurrent Convolutional Face Alignment. Lecture Notes in Computer Science, 2017, , 104-120.	1.0	1
254	Enhancing Perceptual Attributes with Bayesian Style Generation. Lecture Notes in Computer Science, 2019, , 483-498.	1.0	1
255	Low-budget label query through domain alignment enforcement. Computer Vision and Image Understanding, 2022, 222, 103485.	3.0	1
256	Societally connected multimedia across cultures. Journal of Zhejiang University: Science C, 2012, 13, 875-880.	0.7	0
257	Exploiting visual search theory to infer social interactions. , 2013, , .		0
258	Large-Scale Geosocial Multimedia [Guest editorial]. IEEE MultiMedia, 2014, 21, 7-9.	1.5	0
259	Special issue on contextual vision computing. Machine Vision and Applications, 2014, 25, 1631-1633.	1.7	0
260	Special Issue on Event-based Media Processing and Analysis. Image and Vision Computing, 2016, 53, 1-2.	2.7	0
261	A cross-modal adaptation approach for brain decoding. , 2017, , .		0
262	SAWACMMM'17. , 2017, , .		0
263	Exploring Multitask and Transfer Learning Algorithms for Head Pose Estimation in Dynamic Multiview Scenarios. , 2017, , 67-87.		0
264	Special Section on Multimodal Understanding of Social, Affective, and Subjective Attributes. ACM Transactions on Multimedia Computing, Communications and Applications, 2019, 15, 1-3.	3.0	0
265	2nd Workshop on Computational Models of Social Interactions. , 2015, , .		0
266	Projective Unsupervised Flexible Embedding with Optimal Graph. , 2016, , .		0
267	Sparse Code Filtering for Action Pattern Mining. Lecture Notes in Computer Science, 2017, , 3-18.	1.0	0
268	Human-Centered Computing: Application to Multimedia. , 2018, , 1729-1733.		0
269	EE-USAD. , 2018, , .		0