

Andrew Zisserman

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/6676312/publications.pdf>

Version: 2024-02-01

113
papers

43,295
citations

126708

33
h-index

233125

45
g-index

117
all docs

117
docs citations

117
times ranked

22767
citing authors

#	ARTICLE	IF	CITATIONS
1	The Pascal Visual Object Classes (VOC) Challenge. International Journal of Computer Vision, 2010, 88, 303-338.	10.9	12,097
2	Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset. , 2017, , .		4,575
3	Deep Face Recognition. , 2015, , .		2,584
4	Object retrieval with large vocabularies and fast spatial matching. , 2007, , .		2,144
5	Convolutional Two-Stream Network Fusion for Video Action Recognition. , 2016, , .		1,749
6	Automated Flower Classification over a Large Number of Classes. , 2008, , .		1,529
7	Return of the Devil in the Details: Delving Deep into Convolutional Nets. , 2014, , .		1,522
8	VGGFace2: A Dataset for Recognising Faces across Pose and Age. , 2018, , .		1,475
9	Lost in quantization: Improving particular object retrieval in large scale image databases. , 2008, , .		1,054
10	A Statistical Approach to Texture Classification from Single Images. International Journal of Computer Vision, 2005, 62, 61-81.	10.9	998
11	VoxCeleb2: Deep Speaker Recognition. , 0, , .		964
12	Synthetic Data for Text Localisation in Natural Images. , 2016, , .		906
13	Reading Text in the Wild with Convolutional Neural Networks. International Journal of Computer Vision, 2016, 116, 1-20.	10.9	831
14	VoxCeleb: A Large-Scale Speaker Identification Dataset. , 0, , .		786
15	Total Recall: Automatic Query Expansion with a Generative Feature Model for Object Retrieval. , 2007, , .		630
16	The devil is in the details: an evaluation of recent feature encoding methods. , 2011, , .		595
17	Cats and dogs. , 2012, , .		518
18	All About VLAD. , 2013, , .		489

#	ARTICLE	IF	CITATIONS
19	Look, Listen and Learn. , 2017, , .		414
20	Non-uniform Deblurring for Shaken Images. International Journal of Computer Vision, 2012, 98, 168-186.	10.9	390
21	Flowing ConvNets for Human Pose Estimation in Videos. , 2015, , .		350
22	End-to-End Learning of Visual Representations From Uncurated Instructional Videos. , 2020, , .		308
23	A framework for spatiotemporal control in the tracking of visual contours. International Journal of Computer Vision, 1993, 11, 127-145.	10.9	300
24	Geodesic star convexity for interactive image segmentation. , 2010, , .		259
25	Microscopy cell counting and detection with fully convolutional regression networks. Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization, 2018, 6, 283-292.	1.3	251
26	Tabula rasa: Model transfer for object category detection. , 2011, , .		227
27	Video Representation Learning by Dense Predictive Coding. , 2019, , .		203
28	Creating Architectural Models from Images. Computer Graphics Forum, 1999, 18, 39-50.	1.8	194
29	An Exemplar Model for Learning Object Classes. , 2007, , .		193
30	Emotion Recognition in Speech using Cross-Modal Transfer in the Wild. , 2018, , .		180
31	Objects that Sound. Lecture Notes in Computer Science, 2018, , 451-466.	1.0	178
32	Non-uniform deblurring for shaken images. , 2010, , .		174
33	Symbiotic Segmentation and Part Localization for Fine-Grained Categorization. , 2013, , .		170
34	Triangulation Embedding and Democratic Aggregation for Image Search. , 2014, , .		167
35	Utterance-level Aggregation for Speaker Recognition in the Wild. , 2019, , .		161
36	EPIC-Fusion: Audio-Visual Temporal Binding for Egocentric Action Recognition. , 2019, , .		156

#	ARTICLE	IF	CITATIONS
37	Deblurring Shaken and Partially Saturated Images. International Journal of Computer Vision, 2014, 110, 185-201.	10.9	140
38	“Who are you?” Learning person specific classifiers from video. , 2009, , .		137
39	ISSLS PRIZE IN BIOENGINEERING SCIENCE 2017: Automation of reading of radiological features from magnetic resonance images (MRIs) of the lumbar spine without human intervention is comparable with an expert radiologist. European Spine Journal, 2017, 26, 1374-1383.	1.0	131
40	Diagnostically relevant facial gestalt information from ordinary photos. ELife, 2014, 3, e02020.	2.8	129
41	Counting in the Wild. Lecture Notes in Computer Science, 2016, , 483-498.	1.0	127
42	Chimpanzee face recognition from videos in the wild using deep learning. Science Advances, 2019, 5, eaaw0736.	4.7	127
43	Out of Time: Automated Lip Sync in the Wild. Lecture Notes in Computer Science, 2017, , 251-263.	1.0	120
44	Vggsound: A Large-Scale Audio-Visual Dataset. , 2020, , .		119
45	Seeing Voices and Hearing Faces: Cross-Modal Biometric Matching. , 2018, , .		117
46	SpineNet: Automated classification and evidence visualization in spinal MRIs. Medical Image Analysis, 2017, 41, 63-73.	7.0	97
47	A Compact and Discriminative Face Track Descriptor. , 2014, , .		90
48	You Said That?: Synthesising Talking Faces from Audio. International Journal of Computer Vision, 2019, 127, 1767-1779.	10.9	87
49	Interactive Object Counting. Lecture Notes in Computer Science, 2014, , 504-518.	1.0	85
50	Non-contact physiological monitoring of preterm infants in the Neonatal Intensive Care Unit. Npj Digital Medicine, 2019, 2, 128.	5.7	82
51	Learning an Alphabet of Shape and Appearance for Multi-Class Object Detection. International Journal of Computer Vision, 2008, 80, 16-44.	10.9	81
52	Crystal nucleation in metallic alloys using x-ray radiography and machine learning. Science Advances, 2018, 4, eaar4004.	4.7	74
53	Fully-automated alignment of 3D fetal brain ultrasound to a canonical reference space using multi-task learning. Medical Image Analysis, 2018, 46, 1-14.	7.0	72
54	Object Level Grouping for Video Shots. International Journal of Computer Vision, 2006, 67, 189-210.	10.9	71

#	ARTICLE	IF	CITATIONS
55	Self-supervised Learning of Audio-Visual Objects from Video. Lecture Notes in Computer Science, 2020, , 208-224.	1.0	70
56	Memory-Augmented Dense Predictive Coding for Video Representation Learning. Lecture Notes in Computer Science, 2020, , 312-329.	1.0	70
57	Learnable PINs: Cross-modal Embeddings for Person Identity. Lecture Notes in Computer Science, 2018, , 73-89.	1.0	65
58	TeachText: CrossModal Generalized Distillation for Text-Video Retrieval. , 2021, , .		65
59	Object Mining Using a Matching Graph on Very Large Image Collections. , 2008, , .		64
60	Sparse kernel approximations for efficient classification and detection. , 2012, , .		63
61	Detecting overlapping instances in microscopy images using extremal region trees. Medical Image Analysis, 2016, 27, 3-16.	7.0	63
62	Smooth-AP: Smoothing the Path Towards Large-Scale Image Retrieval. Lecture Notes in Computer Science, 2020, , 677-694.	1.0	62
63	Template Adaptation for Face Verification and Identification. , 2017, , .		61
64	BiCoS: A Bi-level co-segmentation method for image classification. , 2011, , .		56
65	Self-supervised Video Object Segmentation by Motion Grouping. , 2021, , .		56
66	Smooth object retrieval using a bag of boundaries. , 2011, , .		52
67	Upper Body Detection and Tracking in Extended Signing Sequences. International Journal of Computer Vision, 2011, 95, 180-197.	10.9	52
68	Geometric Latent Dirichlet Allocation on a Matching Graph for Large-scale Image Datasets. International Journal of Computer Vision, 2011, 95, 138-153.	10.9	51
69	Automatic and Efficient Human Pose Estimation for Sign Language Videos. International Journal of Computer Vision, 2014, 110, 70-90.	10.9	50
70	Seeing the Arrow of Time. , 2014, , .		50
71	Deblurring shaken and partially saturated images. , 2011, , .		49
72	Multi-Task Convolutional Neural Network for Patient Detection and Skin Segmentation in Continuous Non-Contact Vital Sign Monitoring. , 2017, , .		48

#	ARTICLE	IF	CITATIONS
73	Resolving ambiguities in auto-calibration. Philosophical Transactions Series A, Mathematical, Physical, and Engineering Sciences, 1998, 356, 1193-1211.	1.6	44
74	Synthetic Humans for Action Recognition from Unseen Viewpoints. International Journal of Computer Vision, 2021, 129, 2264-2287.	10.9	44
75	LAEO-Net: Revisiting People Looking at Each Other in Videos. , 2019, , .		38
76	Amplifying Key Cues for Human-Object-Interaction Detection. Lecture Notes in Computer Science, 2020, , 248-265.	1.0	38
77	A Statistical Approach to Texture Classification from Single Images. International Journal of Computer Vision, 2005, 62, 61-81.	10.9	34
78	Time-lapse imagery and volunteer classifications from the Zooniverse Penguin Watch project. Scientific Data, 2018, 5, 180124.	2.4	33
79	Automated audiovisual behavior recognition in wild primates. Science Advances, 2021, 7, eabi4883.	4.7	32
80	On-the-fly learning for visual search of large-scale image and video datasets. International Journal of Multimedia Information Retrieval, 2015, 4, 75-93.	3.6	31
81	Humanising GrabCut: Learning to segment humans using the Kinect. , 2011, , .		29
82	Learning to Detect Partially Overlapping Instances. , 2013, , .		29
83	Slow-Fast Auditory Streams for Audio Recognition. , 2021, , .		24
84	Learning equivariant structured output SVM regressors. , 2011, , .		23
85	On-the-fly specific person retrieval. , 2012, , .		22
86	Total Cluster. , 2014, , .		22
87	From Benedict Cumberbatch to Sherlock Holmes: Character Identification in TV series without a Script. , 2017, , .		21
88	Semi-Supervised Learning with Scarce Annotations. , 2020, , .		17
89	3D Shape Attributes. , 2016, , .		15
90	Sight to Sound: An End-to-End Approach for Visual Piano Transcription. , 2020, , .		14

#	ARTICLE	IF	CITATIONS
91	Self-Supervised Learning of Class Embeddings from Video. , 2019, , .		13
92	Identifying Scoliosis in Population-Based Cohorts: Automation of a Validated Method Based on Total Body Dual Energy X-ray Absorptiometry Scans. Calcified Tissue International, 2020, 106, 378-385.	1.5	11
93	Face, Body, Voice: Video Person-Clustering with Multiple Modalities. , 2021, , .		10
94	Learning to Predict 3D Surfaces of Sculptures from Single and Multiple Views. International Journal of Computer Vision, 2019, 127, 1780-1800.	10.9	9
95	Processing citizen science- and machine-annotated time-lapse imagery for biologically meaningful metrics. Scientific Data, 2020, 7, 102.	2.4	9
96	Immediate, Scalable Object Category Detection. , 2014, , .		8
97	From Same Photo: Cheating on Visual Kinship Challenges. Lecture Notes in Computer Science, 2019, , 654-668.	1.0	8
98	Learning to Read by Spelling. , 2018, , .		8
99	LAEO-Net++: Revisiting People Looking at Each Other in Videos. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44, 3069-3081.	9.7	7
100	Automated Video Labelling: Identifying Faces by Corroborative Evidence. , 2021, , .		7
101	Human pose search using deep poselets. , 2015, , .		6
102	QUERYD: A Video Dataset with High-Quality Text and Audio Narrations. , 2021, , .		6
103	LSD-C: Linearly Separable Deep Clusters. , 2021, , .		6
104	Discovery of Rare Phenotypes in Cellular Images Using Weakly Supervised Deep Learning. , 2017, , .		4
105	Count, Crop and Recognise: Fine-Grained Recognition in the Wild. , 2019, , .		4
106	Scaling Up Sign Spotting Through Sign Language Dictionaries. International Journal of Computer Vision, 2022, 130, 1416-1439.	10.9	4
107	Automated Video Face Labelling for Films and TV Material. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42, 780-792.	9.7	3
108	The Use and Reuse of Printed Illustrations in 15th-Century Venetian Editions. Studi Di Storia, 2020, , .	0.0	2

#	ARTICLE	IF	CITATIONS
109	Where Is the Archive?. <i>Photography and Culture</i> , 2010, 3, 331-342.	0.2	1
110	SeeHear: Signer Diarisation and a New Dataset. , 2021, , .		1
111	“Who are you?” - Learning person specific classifiers from video. , 2009, , .		1
112	Illuminance Flow Estimation by Regression. <i>International Journal of Computer Vision</i> , 2010, 90, 304-312.	10.9	0
113	Bounding an archiving: assessing the relative completeness of the Jacques Toussele archive using pattern-matching and face-recognition. <i>Visual Studies</i> , 0, , 1-25.	0.3	0