# Qi Wu

## List of Publications by Year in Descending Order



| | | | |
|---|---|---|---|
| 72 papers | 1,753 citations | 19 h-index | 41 g-index |
| 83 ext. papers | 2,621 ext. citations | 3.9 avg, IF | 5.47 L-index |

| # | Paper | IF | Citations |
|---|---|---|---|
| 72 | A proposal-free one-stage framework for referring expression comprehension and generation via dense cross-attention. *IEEE Transactions on Multimedia*, **2022**, 1-1 | 6.6 | 2 |
| 71 | Medical VQA. *Advances in Computer Vision and Pattern Recognition*, **2022**, 165-176 | 1.1 | 0 |
| 70 | Vision-and-Language Pretraining for VQA. *Advances in Computer Vision and Pattern Recognition*, **2022**, 91-107 | 1.1 | |
| 69 | Knowledge-Based VQA. *Advances in Computer Vision and Pattern Recognition*, **2022**, 73-90 | 1.1 | |
| 68 | Visual Question Generation. *Advances in Computer Vision and Pattern Recognition*, **2022**, 189-197 | 1.1 | |
| 67 | Classical Visual Question Answering. *Advances in Computer Vision and Pattern Recognition*, **2022**, 35-72 | 1.1 | |
| 66 | Video Representation Learning. *Advances in Computer Vision and Pattern Recognition*, **2022**, 111-117 | 1.1 | |
| 65 | Text-Based VQA. *Advances in Computer Vision and Pattern Recognition*, **2022**, 177-187 | 1.1 | |
| 64 | Visual Dialogue. *Advances in Computer Vision and Pattern Recognition*, **2022**, 199-218 | 1.1 | |
| 63 | Video Question Answering. *Advances in Computer Vision and Pattern Recognition*, **2022**, 119-133 | 1.1 | |
| 62 | Question Answering (QA) Basics. *Advances in Computer Vision and Pattern Recognition*, **2022**, 27-31 | 1.1 | |
| 61 | Embodied VQA. *Advances in Computer Vision and Pattern Recognition*, **2022**, 147-164 | 1.1 | |
| 60 | Deep Learning Basics. *Advances in Computer Vision and Pattern Recognition*, **2022**, 15-26 | 1.1 | 0 |
| 59 | Advanced Models for Video Question Answering. *Advances in Computer Vision and Pattern Recognition*, **2022**, 135-143 | 1.1 | |
| 58 | Referring Expression Comprehension. *Advances in Computer Vision and Pattern Recognition*, **2022**, 219-230 | 1.1 | |
| 57 | Non-Salient Region Object Mining for Weakly Supervised Semantic Segmentation **2021**, | | 24 |
| 56 | Jo-SRC: A Contrastive Approach for Combating Noisy Labels **2021**, | | 14 |

| | | | |
|---|---|---|---|
| 55 | Room-and-Object Aware Knowledge Reasoning for Remote Embodied Referring Expression **2021**, | | 5 |
| 54 | . *IEEE Open Journal of Intelligent Transportation Systems*, **2021**, 1-1 | 1.7 | 1 |
| 53 | . *IEEE Transactions on Multimedia*, **2021**, 1-1 | 6.6 | 3 |
| 52 | Modular Graph Attention Network for Complex Visual Relational Reasoning. *Lecture Notes in Computer Science*, **2021**, 137-153 | 0.9 | |
| 51 | . *IEEE Transactions on Multimedia*, **2021**, 1-1 | 6.6 | |
| 50 | Learning Dual Encoding Model for Adaptive Visual Understanding in Visual Dialogue. *IEEE Transactions on Image Processing*, **2021**, 30, 220-233 | 8.7 | 1 |
| 49 | Optimistic Agent: Accurate Graph-Based Value Estimation for More Successful Visual Navigation **2021**, | | 2 |
| 48 | . *IEEE Transactions on Multimedia*, **2021**, 1-1 | 6.6 | 3 |
| 47 | Medical Data Inquiry Using a Question Answering Model **2020**, | | 2 |
| 46 | . *IEEE Transactions on Multimedia*, **2020**, 22, 3196-3209 | 6.6 | 17 |
| 45 | Give Me Something to Eat: Referring Expression Comprehension with Commonsense Knowledge **2020**, | | 2 |
| 44 | Data-driven Meta-set Based Fine-Grained Visual Recognition **2020**, | | 13 |
| 43 | DualVD: An Adaptive Dual Encoding Model for Deep Visual Understanding in Visual Dialogue. *Proceedings of the AAAI Conference on Artificial Intelligence*, **2020**, 34, 11125-11132 | 5 | 9 |
| 42 | Visual-Semantic Graph Matching for Visual Grounding **2020**, | | 3 |
| 41 | Length-Controllable Image Captioning. *Lecture Notes in Computer Science*, **2020**, 712-729 | 0.9 | 5 |
| 40 | . *IEEE Transactions on Multimedia*, **2020**, 1-1 | 6.6 | 1 |
| 39 | Soft Expert Reward Learning for Vision-and-Language Navigation. *Lecture Notes in Computer Science*, **2020**, 126-141 | 0.9 | 4 |
| 38 | Object-and-Action Aware Model for Visual Language Navigation. *Lecture Notes in Computer Science*, **2020**, 303-317 | 0.9 | 5 |

| | | | |
|---|---|---|---|
| 37 | Scripted Video Generation With a Bottom-Up Generative Adversarial Network. *IEEE Transactions on Image Processing*, **2020**, 29, 7454-7467 | 8.7 | 4 |
| 36 | Visual Grounding via Accumulated Attention. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **2020**, PP, | 13.3 | 1 |
| 35 | Overcoming Language Priors in VQA via Decomposed Linguistic Representations. *Proceedings of the AAAI Conference on Artificial Intelligence*, **2020**, 34, 11181-11188 | 5 | 10 |
| 34 | REVERIE: Remote Embodied Visual Referring Expression in Real Indoor Environments **2020**, | | 19 |
| 33 | Cops-Ref: A New Dataset and Task on Compositional Referring Expression Comprehension **2020**, | | 4 |
| 32 | . *IEEE Transactions on Circuits and Systems for Video Technology*, **2020**, 1-1 | 6.4 | 6 |
| 31 | Image and Sentence Matching via Semantic Concepts and Order Learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **2020**, 42, 636-650 | 13.3 | 11 |
| 30 | Heritage image annotation via collective knowledge. *Pattern Recognition*, **2019**, 93, 204-214 | 7.7 | 4 |
| 29 | Medical image classification using synergic deep learning. *Medical Image Analysis*, **2019**, 54, 10-19 | 15.4 | 117 |
| 28 | . *IEEE Transactions on Multimedia*, **2019**, 21, 1971-1981 | 6.6 | 30 |
| 27 | What Is to Know? Uncertainty as a Guide to Asking Goal-Oriented Questions **2019**, | | 4 |
| 26 | Neighbourhood Watch: Referring Expression Comprehension via Language-Guided Graph Attention Networks **2019**, | | 52 |
| 25 | Mind Your Neighbours: Image Annotation With Metadata Neighbourhood Graph Co-Attention Networks **2019**, | | 9 |
| 24 | An Attribute-Based High-Level Image Representation for Scene Classification. *IEEE Access*, **2019**, 7, 4629-4640 | 3.5 | 3 |
| 23 | FVQA: Fact-based Visual Question Answering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **2018**, 40, 2413-2427 | 13.3 | 97 |
| 22 | . *IEEE Transactions on Multimedia*, **2018**, 20, 2801-2813 | 6.6 | 76 |
| 21 | Image Captioning and Visual Question Answering Based on Attributes and External Knowledge. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **2018**, 40, 1367-1381 | 13.3 | 147 |
| 20 | Goal-Oriented Visual Question Generation via Intermediate Rewards. *Lecture Notes in Computer Science*, **2018**, 189-204 | 0.9 | 7 |

| | | | |
|---|---|---|---|
| 19 | Learning Semantic Concepts and Order for Image and Sentence Matching **2018**, | | 111 |
| 18 | **2018**, | | 59 |
| 17 | Vision-and-Language Navigation: Interpreting Visually-Grounded Navigation Instructions in Real Environments **2018**, | | 170 |
| 16 | Visual Question Answering with Memory-Augmented Networks **2018**, | | 34 |
| 15 | Are You Talking to Me? Reasoned Visual Dialog Generation Through Adversarial Learning **2018**, | | 31 |
| 14 | Parallel Attention: A Unified Framework for Visual Object Discovery Through Dialogs and Queries **2018**, | | 32 |
| 13 | Skin Lesion Classification in Dermoscopy Images Using Synergic Deep Learning. *Lecture Notes in Computer Science*, **2018**, 12-20 | 0.9 | 29 |
| 12 | Research of UAV target detection and flight control based on deep learning **2018**, | | 4 |
| 11 | Visual question answering: A survey of methods and datasets. *Computer Vision and Image Understanding*, **2017**, 163, 21-40 | 4.3 | 108 |
| 10 | . *IEEE Signal Processing Magazine*, **2017**, 34, 63-75 | 9.4 | 18 |
| 9 | The VQA-Machine: Learning How to Use Existing Vision Algorithms to Answer New Questions **2017**, | | 30 |
| 8 | Explicit Knowledge-based Reasoning for Visual Question Answering **2017**, | | 35 |
| 7 | What Value Do Explicit High Level Concepts Have in Vision to Language Problems? **2016**, | | 194 |
| 6 | Ask Me Anything: Free-Form Visual Question Answering Based on Knowledge from External Sources **2016**, | | 126 |
| 5 | Cross-depiction problem: Recognition and synthesis of photographs and artwork. *Computational Visual Media*, **2015**, 1, 91-103 | 3.9 | 16 |
| 4 | Beyond Photo-Domain Object Recognition: Benchmarks for the Cross-Depiction Problem **2015**, | | 9 |
| 3 | Learning Graphs to Model Visual Objects across Different Depictive Styles. *Lecture Notes in Computer Science*, **2014**, 313-328 | 0.9 | 9 |
| 2 | Modelling Visual Objects Invariant to Depictive Style **2013**, | | 2 |

1 Prime Shapes in Natural Images **2012**, 2