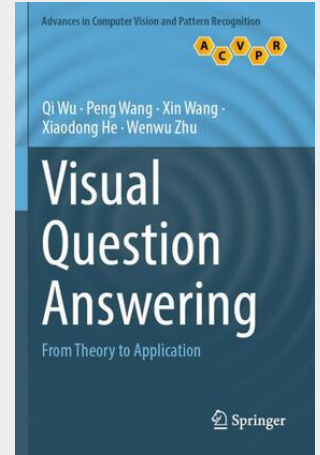


Visual Question Answering

From Theory to Application

Visual Question Answering (VQA) usually combines visual inputs like image and video with a natural language question concerning the input and generates a natural language answer as the output. This is by nature a multi-disciplinary research problem, involving computer vision (CV), natural language processing (NLP), knowledge representation and reasoning (KR), etc. Further, VQA is an ambitious undertaking, as it must overcome the challenges of general image understanding and the question-answering task, as well as the difficulties entailed by using large-scale databases with mixed-quality inputs. However, with the advent of deep learning (DL) and driven by the existence of advanced techniques in both CV and NLP and the availability of relevant large-scale datasets, we have recently seen enormous strides in VQA, with more systems and promising results emerging. This book provides a comprehensive overview of VQA, covering fundamental theories, models, datasets, and promising future directions. Given its scope, it can be used as a textbook on computer vision and natural language processing, especially for researchers and students in the area of visual question answering. It also highlights the key models used in VQA.

Visual Question Answering (VQA) usually combines visual inputs like image and video with a natural language question concerning the input and generates a natural language answer as the output. This is by nature a multi-disciplinary research problem, involving computer vision (CV), natural language processing (NLP), knowledge representation and reasoning (KR), etc. Further, VQA is an ambitious undertaking, as it must overcome the challenges of general image understanding and the question-answering task, as well as the difficulties entailed by using large-scale databases with mixed-quality inputs. However, with the advent of deep learning (DL) and driven by the existence of advanced techniques in both CV and NLP and the availability of relevant large-scale datasets, we have recently seen enormous strides in VQA, with more systems and promising results emerging. This book provides a comprehensive overview of VQA, covering fundamental theories, models, datasets, and promising future directions. Given its scope, it can be used as a textbook on computer vision and natural language processing, especially for researchers and students in the area of visual question answering. It also highlights the key models used in VQA.



106,99 €

99,99 € (zzgl. MwSt.)

Lieferfrist: bis zu 10 Tage

Artikelnummer: 9789811909665

Medium: Buch

ISBN: 978-981-19-0966-5

Verlag: Springer Nature Singapore

Erscheinungstermin: 15.05.2023

Sprache(n): Englisch

Auflage: 1. Auflage 2022

Serie: Advances in Computer Vision and Pattern Recognition

Produktform: Kartoniert

Gewicht: 388 g

Seiten: 238

Format (B x H): 155 x 235 mm

