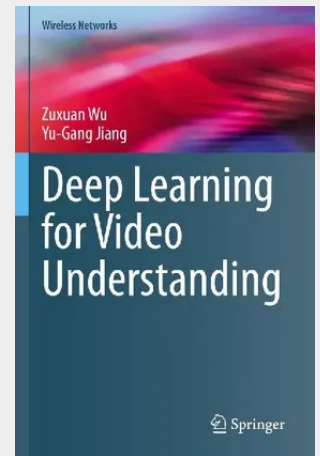


Deep Learning for Video Understanding

This book presents deep learning techniques for video understanding. For deep learning basics, the authors cover machine learning pipelines and notations, 2D and 3D Convolutional Neural Networks for spatial and temporal feature learning. For action recognition, the authors introduce classical frameworks for image classification, and then elaborate both image-based and clip-based 2D/3D CNN networks for action recognition. For action detection, the authors elaborate sliding windows, proposal-based detection methods, single stage and two stage approaches, spatial and temporal action localization, followed by datasets introduction. For video captioning, the authors present language-based models and how to perform sequence to sequence learning for video captioning. For unsupervised feature learning, the authors discuss the necessity of shifting from supervised learning to unsupervised learning and then introduce how to design better surrogate training tasks to learn video representations. Finally, the book introduces recent self-training pipelines like contrastive learning and masked image/video modeling with transformers. The book provides promising directions, with an aim to promote future research outcomes in the field of video understanding with deep learning.

This book presents deep learning techniques for video understanding. For deep learning basics, the authors cover machine learning pipelines and notations, 2D and 3D Convolutional Neural Networks for spatial and temporal feature learning. For action recognition, the authors introduce classical frameworks for image classification, and then elaborate both image-based and clip-based 2D/3D CNN networks for action recognition. For action detection, the authors elaborate sliding windows, proposal-based detection methods, single stage and two stage approaches, spatial and temporal action localization, followed by datasets introduction. For video captioning, the authors present language-based models and how to perform sequence to sequence learning for video captioning. For unsupervised feature learning, the authors discuss the necessity of shifting from supervised learning to unsupervised learning and then introduce how to design better surrogate training tasks to learn video representations. Finally, the book introduces recent self-training pipelines like contrastive learning and masked image/video modeling with transformers. The book provides promising directions, with an aim to promote future research outcomes in the field of video understanding with deep learning.

- Presents an overview of deep learning techniques for video understanding;
- Covers important topics like action recognition, action localization, video captioning, and more;
- Introduces cutting-edge and state-of-the-art video understanding techniques.



149,79 €

139,99 € (zzgl. MwSt.)

Lieferfrist: bis zu 10 Tage

Artikelnummer: 9783031576812

Medium: Buch

ISBN: 978-3-031-57681-2

Verlag: Palgrave Macmillan

Erscheinungstermin: 02.08.2025

Sprache(n): Englisch

Auflage: Erscheinungsjahr 2025

Serie: Wireless Networks

Produktform: Kartoniert

Gewicht: 312 g

Seiten: 188

Format (B x H): 155 x 235 mm

