

1. Record Nr.	UNINA9911049213803321
Autore	Wang Xin
Titolo	Video Grounding and Its Generalization : From I.D. and Task-specific Models to O.O.D. and Large Foundation Models // by Xin Wang, Xiaohan Lan, Wenwu Zhu
Pubbl/distr/stampa	Cham : , : Springer Nature Switzerland : , : Imprint : Springer, , 2025
ISBN	3-031-94837-8
Edizione	[1st ed. 2025.]
Descrizione fisica	1 online resource (262 pages)
Collana	Professional and Applied Computing Series
Disciplina	006.696
Soggetti	Multimedia systems Computer vision Natural language processing (Computer science) Machine learning Multimedia Information Systems Computer Vision Natural Language Processing (NLP) Machine Learning
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Preface -- Introduction -- Traditional Temporal Sentence Grounding in Videos -- Generalized Video Grounding -- Future Research Directions -- References.
Sommario/riassunto	This book consists of two parts: Part I Methodologies for Video Grounding and Part II Generalized Video Grounding and Trending Directions. To make this book self-contained and cutting edge, Part I will cover basic and advanced methodologies for Video Grounding, discussing key comparisons with several representative Vision-Language learning tasks including multimodal understanding and generation. Part II will cover our insights for Generalized Video Grounding and the development of Video Grounding in the era of large foundation models, discussing future directions such as Out-of-Distribution settings which deserve further investigations. Discussions on Video Grounding will cover both the task of Video Grounding and other Vision-Language Task, as well as their relations. The basics and

advances will touch Video Grounding from model to benchmark, from supervised learning to unsupervised pre-training, from single video grounding to video corpus grounding, and from in-distribution setting to out-of-distribution setting. As for Generalized Video Grounding, we discuss cross-modal grounding, event grounding for multi-modal tasks, various distribution shifts in out-of-distribution setting, explainable Video Grounding, and large foundation model for Video Grounding. We deeply hope this book can benefit interested readers from both academy and industry, covering needs from junior starters in research to senior practitioners in IT companies.
