Paper	Source	Datasets
MV-Adapter: Multimodal Video Transfer Learning for Video Text Retrieval.	[CVPR 2024](#)	MSR-VTT, MSVD, LSMDC, DiDeMo, ActivityNet Caption
Holistic Features are Almost Sufficient for Text-to-Video Retrieval.	[CVPR 2024](#)	MSR-VTT, MSVD, VATEX, ActivityNet Caption, DiDeMo
Text Is MASS: Modeling as Stochastic Embedding for Text-Video Retrieval.	[CVPR 2024](#)	MSR-VTT, LSMDC, DiDeMo, Charades, VATEX
Dual Alignment Unsupervised Domain Adaptation for Video-Text Retrieval.	[CVPR 2023](#)	MSR-VTT, MSVD, TGIF
VoP: Text-Video Co-Operative Prompt Tuning for Cross-Modal Retrieval.	[CVPR 2023](#)	MSR-VTT, DiDeMo, ActivityNet Caption, LSMDC
Towards Generalisable Video Moment Retrieval: Visual-Dynamic Injection to Image-Text Pre-Training.	[CVPR 2023](#)	MSR-VTT, DiDeMo, VATEX, MSVD
DiffusionRet: Generative Text-Video Retrieval with Diffusion Model.	[ICCV 2023](#)	MSR-VTT, LSMDC, MSVD, ActivityNet Caption, DiDeMo
Prompt Switch: Efficient CLIP Adaptation for Text-Video Retrieval.	[ICCV 2023](#)	MSR-VTT, MSVD, LSMDC
UATVR: Uncertainty-Adaptive Text-Video Retrieval.	[ICCV 2023](#)	MSR-VTT, MSVD, DiDeMo, VATEX
PIDRo: Parallel Isomeric Attention with Dynamic Routing for Text-Video Retrieval.	[ICCV 2023](#)	MSR-VTT, MSVD, DiDeMo, ActivityNet Caption
Audio-Enhanced Text-to-Video Retrieval using Text-Conditioned Feature Alignment.	[ICCV 2023](#)	MSR-VTT, LSMDC, VATEX, Charades
Progressive Spatio-Temporal Prototype Matching for Text-Video Retrieval.	[ICCV 2023](#)	MSR-VTT, DiDeMo, VATEX, LSMDC
In-Style: Bridging Text and Uncurated Videos with Style Transfer for Text-Video Retrieval.	[ICCV 2023](#)	MSR-VTT, YouCook2, MSVD, LSMDC, DiDeMo, VATEX, HowTo100M
Unified Coarse-to-Fine Alignment for Video-Text Retrieval.	[ICCV 2023](#)	MSR-VTT, ActivityNet Caption, DiDeMo, MSVD, VATEX
Multi-event Video-Text Retrieval.	[ICCV 2023](#)	ActivityNet Caption, Charades
Rethinking Video-Text Understanding: Retrieval from Counterfactually Augmented Data.	[ECCV 2024](#)	MSR-VTT, VATEX

Paper	Source	Datasets
KDProR: A Knowledge-Decoupling Probabilistic Framework for Video-Text Retrieval.	ECCV 2024	MSR-VTT, DiDeMo, LSMDC, ActivityNet Caption

Dataset	Number of Uses
MSR-VTT	16
DiDeMo	12
MSVD	10
LSMDC	9
VATEX	9
ActivityNet Caption	8
Charades	3
YouCook2	1
HowTo100M	1
TGIF	1