ICCV
OCT 19-23, 2025
Highlight

UCF Center for Research in Computer Vision

DisenQ: Disentangling Q-Former For Activity-Biometrics

Shehreen Azad

Yogesh Rawat

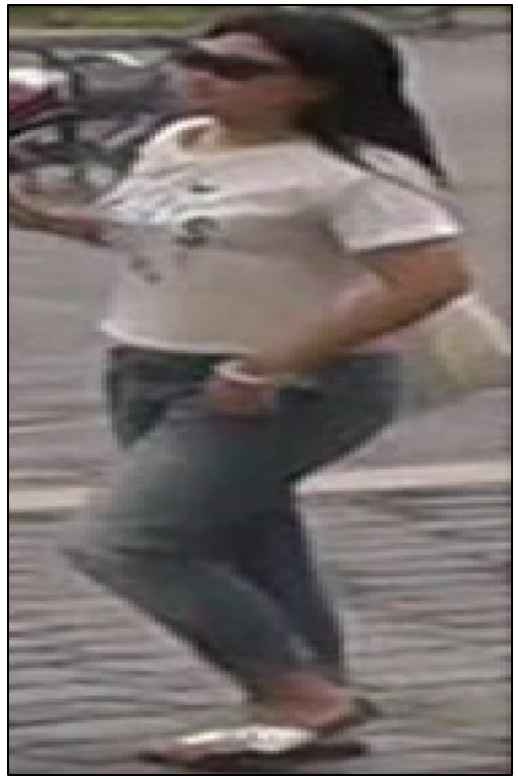
Center for Research in Computer Vision, University of Central Florida

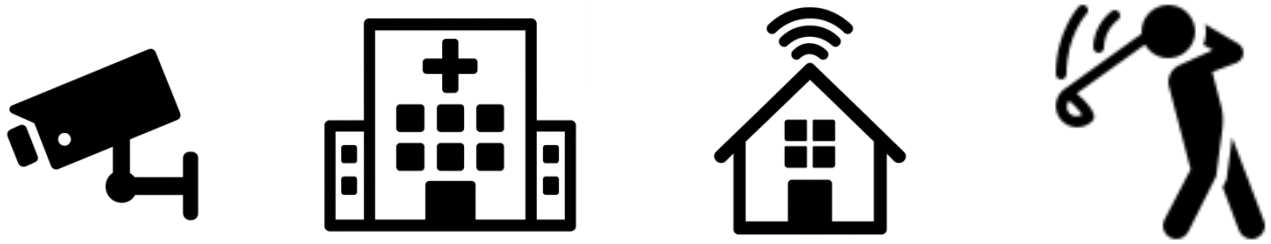
ICCV
OCT 19-23, 2025
★ Highlight

UCF Center for Research in Computer Vision

Traditional Person Identification





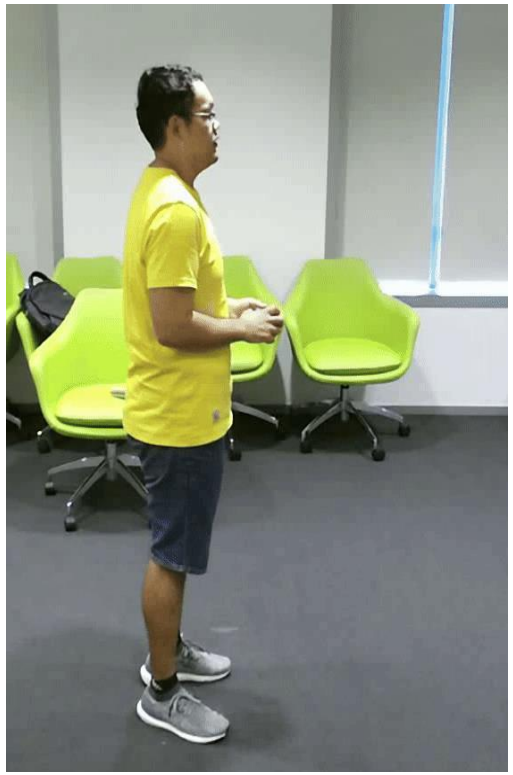
But people **don’t just walk**, they live.

ICCV OCT 19-23, 2025
★ Highlight

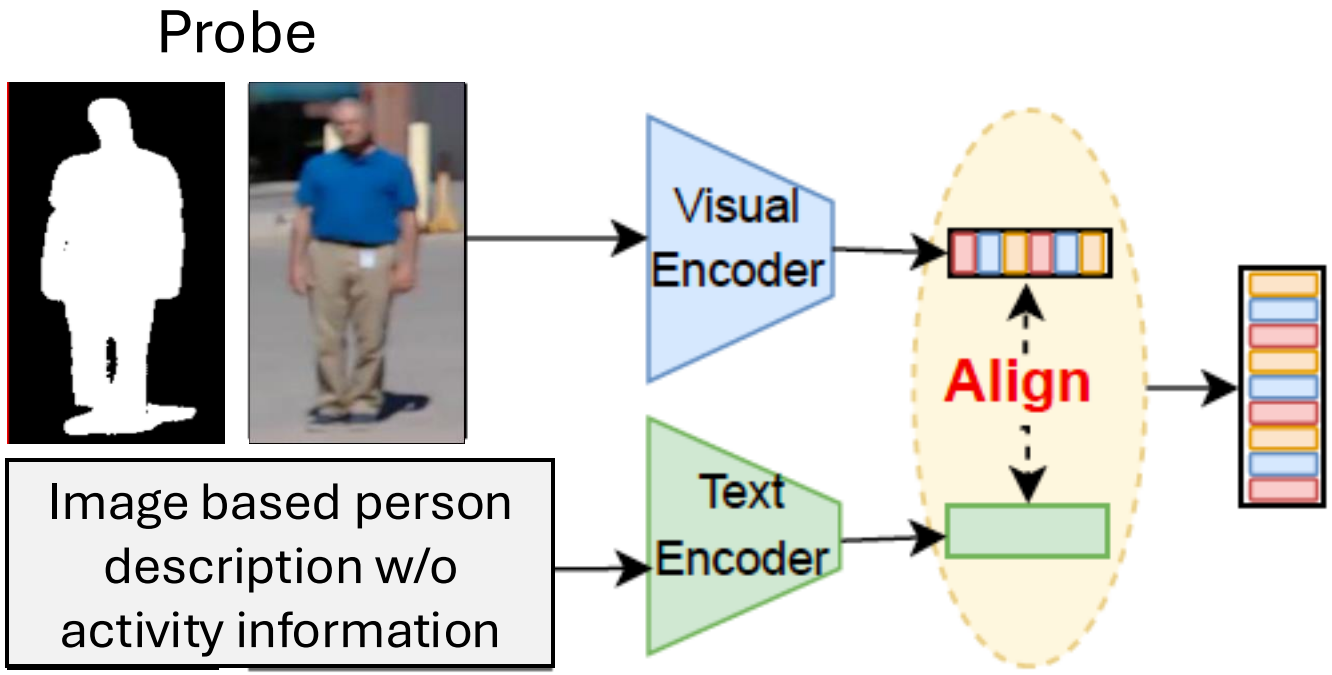
UCF Center for Research in Computer Vision

From Gait to Daily Activities






Challenges in Current Multimodal Models



ICCV OCT 19-23, 2025
★ Highlight

UCF Center for Research in Computer Vision

Reliance on Additional Visual Modality

Probe

✗ Mismatch from noisy silhouettes

Biometrics:
<body shape>,
<build>

Motion:
<action label>,
<description>

Non-biometrics:
<clothing>,
<footwear>

✓ Text supervision replaces silhouettes

Identity Corrupted by Appearance

Probe

✗ Mismatch from **appearance bias**

Biometrics disentangled:
✓ ignores appearance bias

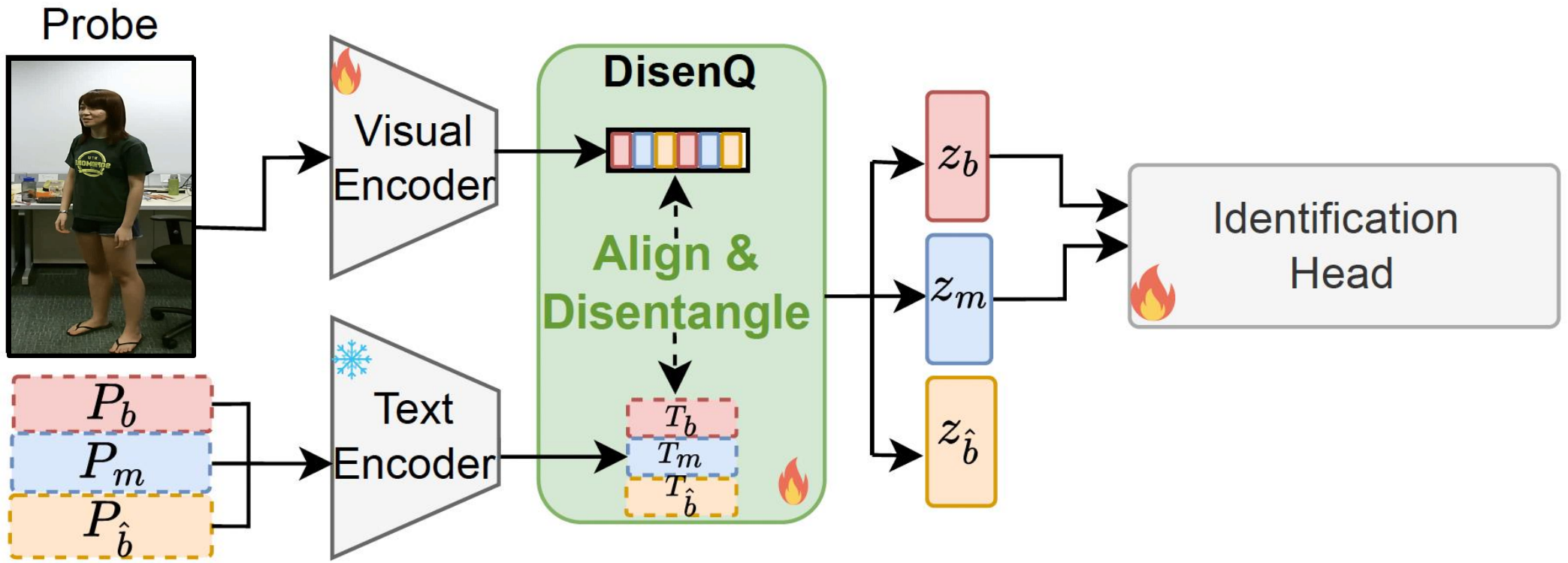
Over Reliance on Motion Cues

Probe

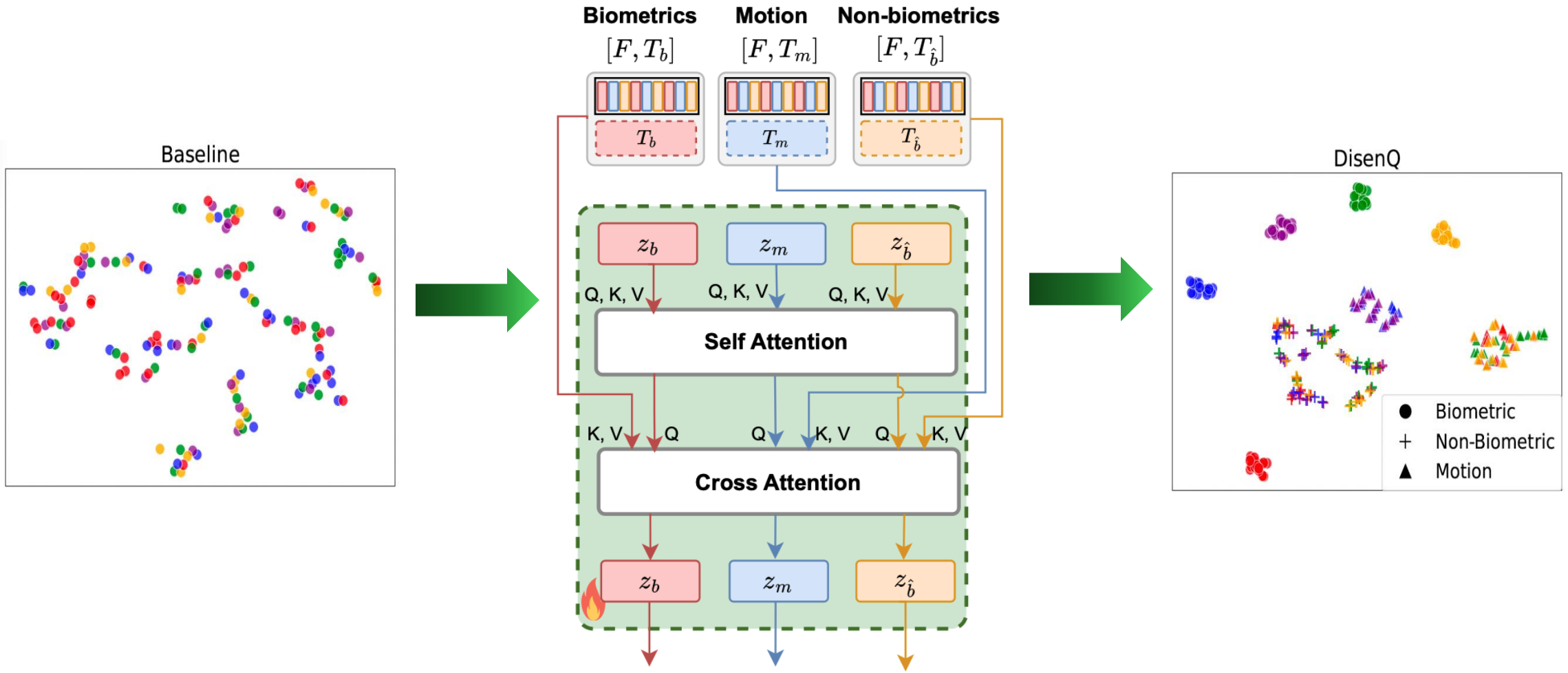
✗ Mismatch from over **motion reliance**

Motion disentangled:
✓ action relevance guides identity

Our Approach



DisenQ Framework



ICCV OCT 19-23, 2025

★ Highlight

UCF Center for Research in Computer Vision

Benchmark Performance

Methods	Venue	NTU RGB-AB				PKU MMD-AB				Charades-AB			
		Same		Cross		Same		Cross		Same		Cross	
		R@1	mAP	R@1	mAP	R@1	mAP	R@1	mAP	R@1	mAP	R@1	mAP
Models with only visual modality													
TSF [28]	AAAI 20	71.8	31.8	67.8	26.9	76.4	37.5	71.6	33.2	35.4	21.9	30.2	19.0
VKD [44]	ECCV 20	67.4	35.6	66.3	31.5	78.4	38.5	72.2	34.3	36.3	20.7	31.9	18.8
BiCnet-TKS [25]	CVPR 21	72.7	34.5	69.1	30.2	80.8	38.5	77.1	33.3	40.3	27.3	38.3	23.3
PSTA [51]	ICCV 21	67.4	34.8	65.1	31.4	77.4	50.4	72.4	47.4	42.9	28.3	38.7	24.8
STMN [15]	ICCV 21	73.0	35.1	70.2	30.1	76.6	47.9	71.5	42.2	38.7	24.5	33.9	20.8
SINet [6]	CVPR 22	69.4	30.7	66.2	27.8	79.6	40.8	74.1	26.2	40.3	26.9	37.3	21.9
CAL [19]	CVPR 22	73.8	28.4	70.3	24.0	81.3	49.4	78.3	43.4	43.8	25.8	40.1	21.2
Video-CAL [19]	CVPR 22	75.5	39.9	73.3	31.7	79.6	49.4	77.3	45.7	43.9	28.5	41.5	25.8
PSTR [7]	CVPR 22	69.1	34.1	68.3	32.5	84.3	47.5	78.0	41.2	37.2	24.7	35.1	20.3
AIM [55]	CVPR 23	71.4	35.4	72.8	30.2	82.5	48.9	79.2	44.9	40.1	28.3	35.6	26.7
SCNet [20]	ACM MM 23	69.9	31.5	68.8	26.3	79.5	43.6	73.9	39.7	31.7	21.9	27.4	17.6
ABNet [3]	CVPR 24	78.8	40.3	77.0	37.6	86.8	57.3	81.4	51.8	45.8	31.6	44.8	28.8
Models with visual + language modality													
CLIP ReID [32] †	AAAI 23	77.1	40.2	75.2	33.7	82.3	52.1	81.2	50.8	44.2	31.3	42.1	27.7
CCLNet [10] †	ACM MM 23	75.2	36.1	74.3	33.1	83.2	51.4	80.1	47.5	42.1	29.3	38.8	23.4
TF-CLIP [58] †	AAAI 24	77.3	41.2	74.8	31.3	83.4	52.3	80.8	50.1	40.2	28.1	39.7	26.0
TVI-LFM [26] †	NeurIPS 24	76.2	38.1	75.9	34.1	85.2	53.9	81.5	52.1	45.7	30.1	42.8	28.3
Instruct-ReID [22] †	CVPR 24	78.2	41.5	75.9	33.4	84.3	53.1	81.7	52.3	44.8	28.3	40.1	25.3
EVA-CLIP [49] †		71.2	35.1	69.1	28.3	73.8	46.2	67.4	39.4	38.1	26.1	31.3	21.8
Ours		**82.2**	**43.8**	**80.9**	**41.3**	**89.2**	**59.3**	**84.1**	**56.9**	**49.9**	**34.8**	**48.4**	**32.5**

ICCV OCT 19-23, 2025

★ Highlight

UCF Center for Research in Computer Vision

Generalization Ability

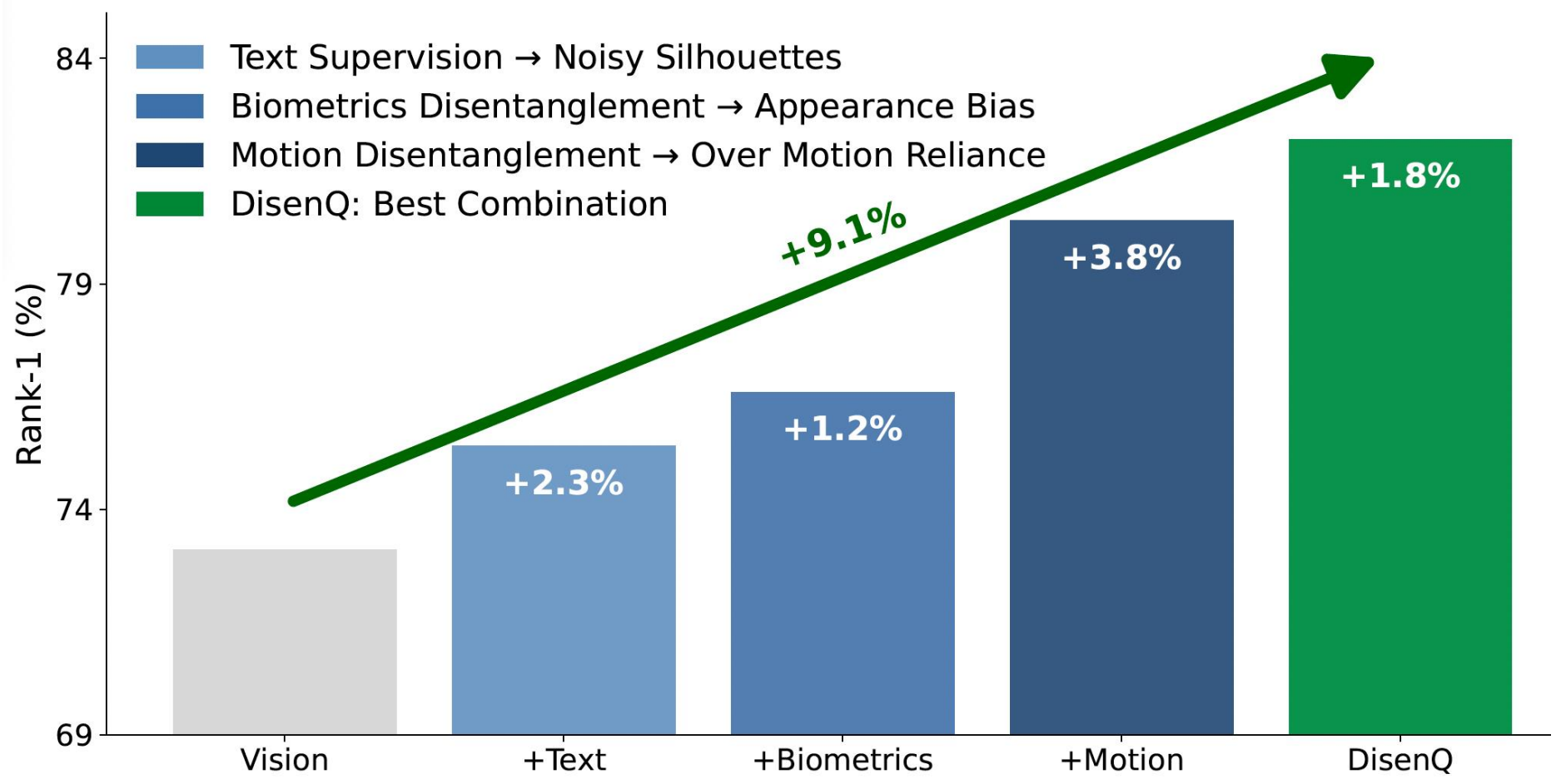
MEVID

Methods	Venue	R@1	R@5	mAP
Models with only visual modality				
Attn-CL [43]	AAAI 20	42.1	56.0	18.6
Attn-CL + rerank [43]	AAAI 20	46.5	59.8	25.9
AP3D [18]	ECCV 20	39.0	56.0	15.9
TCLNet [24]	ECCV 20	48.1	60.1	23.0
BiCnet-TKS [25]	CVPR 21	19.0	35.1	6.3
STMN [14]	ICCV 21	31.0	54.4	11.3
PSTA [51]	ICCV 21	46.9	60.8	21.2
PiT [59]	TII 22	34.2	55.4	13.6
CAL [19]	CVPR 23	52.5	66.5	27.1
ShARc [62]	WACV 24	<u>59.5</u>	**70.3**	29.6
ABNet [3] †	CVPR 24	58.3	68.4	<u>30.1</u>
Models with visual + language modality				
CLIP ReID [32] †	AAAI 23	51.2	64.2	28.3
CCLNet [10] †	ACM MM 23	50.8	60.3	27.1
TVI-LFM [26] †	NeurIPS 24	49.2	61.8	23.7
Instruct-ReID [22] †	CVPR 24	53.8	59.4	28.4
EVA-CLIP [49] †		53.1	59.2	26.9
Ours		**60.7**	**70.3**	**30.4**

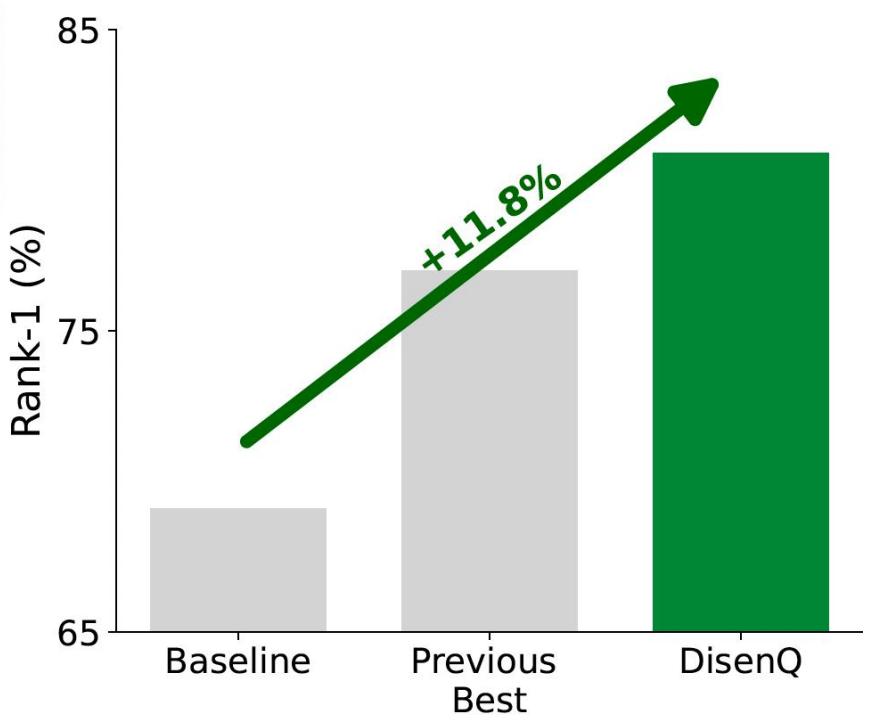
ICCV OCT 19-23, 2025
★ Highlight

UCF Center for Research in Computer Vision

Ablation Studies



Superior Cross-Activity Robustness



Probe

✓ Correct match across different activity

ICCV
OCT 19-23, 2025
★ Highlight

UCF Center for Research in Computer Vision

Thank You

Corresponding author:
Shehreen.Azad@ucf.edu