



#### TUE-PM-286

# MISC210K: A Large-Scale Dataset for Multi-Instance Semantic Correspondence

Yixuan Sun<sup>1,\*</sup>, Yiwen Huang<sup>2,\*</sup>, Haijing Guo<sup>2</sup>, Yuzhou Zhao<sup>2</sup>, Runmin Wu<sup>3</sup>, Yizhou Yu<sup>3</sup>, Weifeng Ge<sup>2,†</sup>, Wenqiang Zhang<sup>1,2,†</sup>

<sup>1</sup>Academy of Engineering & Technology, Fudan University, Shanghai, China <sup>2</sup>School of Computer Science, Fudan University, Shanghai, China <sup>3</sup>The University of Hong Kong, Hong Kong, China {wfge, wqzhang}@fudan.edu.cn

#### Introduction

- Existing semantic correspondence datasets mainly focus on one-to-one object matching.
- However, it is not suitable for real-world applications.



Spair-71k (Single to Single)



MISC210K (Multi-to-Multi)

#### Introduction

- MISC210K Dataset:
  - 218,179 image pairs across 34 object categories.
  - Multi-to-multi matching cases.
  - More complicated annotations, larger scale, and more challenging variations.



Overview of dataset construction pipeline

**Statistics** 

Challenging

Example

#### Introduction

- Dual-Path Collaborative Learning (DPCL) Pipeline:
  - Extract discriminative features.
  - Alleviate uncertainty in the number of matching keypoints.
  - Handle occlusion and interlacing.



Task Definition and Design Protocols:



Category and Image Selection:



Category and Image Selection:



Category Keypoint System Definition:



Human-machine Collaborative Annotation:



• Statistics and Examples:





Scale Variation

Heavy Occlusion Shape Inconsistency

# DPCL

- **D**ual-**P**ath **C**ollaborative Learning (DPCL) Pipeline:
  - Feature extraction backbone: ViT-B pre-trained with iBOT strategy.
  - Transformer decoder: 6 cascaded transformer blocks for generating 4D cost volume.
  - Semantic correspondence branch: employs a sigmoid block, non-maximum suppression, and static thresholding to obtain final predictions.
  - Instance segmentation branch: enables the grouping of same-instance matching key-points.



## **Experimental Results**

#### • Baseline evaluation:

Method	α	airplane	base- ballba	t bear	bed	bench	bicycle	bird	boat	bottle	bus	car	cat	chair	clock	cow	cup	dog	all
MMNet [49]	0.05	5.26	5.57	6.38	2.12	5.70	6.72	7.47	6.56	6.30	3.06	1.26	6.86	4.09	11.53	5.68	5.04	6.61	5.68
	0.10	20.38	21.81	25.95	9.86	18.03	26.80	27.98	23.64	20.50	11.85	5.96	24.03	15.10	39.78	23.92	18.79	26.70	21.58
	0.15	37.38	40.60	48.14	20.79	32.30	47.89	51.37	41.46	34.26	25.70	13.26	43.39	28.74	61.37	45.92	35.50	49.48	39.66
	1.00	99.76	99.94	99.98	99.39	98.63	99.71	99.92	100.00	99.99	99.73	99.89	99.57	99.91	99.96	99.85	99.47	99.69	99.76
CATs [4]	0.05	10.95	4.68	10.50	4.64	3.95	9.18	10.76	8.58	5.43	11.12	8.27	10.41	4.53	15.88	12.27	4.86	8.62	10.00
	0.10	25.55	14.47	27.09	13.02	12.40	23.68	24.11	20.86	14.55	27.13	19.87	24.44	13.28	33.29	27.83	15.07	22.79	23.88
	0.15	38.27	25.04	40.56	22.18	21.16	36.94	36.63	31.64	23.27	39.69	27.50	36.46	23.88	46.00	39.05	23.50	34.93	35.45
	1.00	88.02	91.85	88.41	84.60	83.53	88.90	89.56	88.96	91.70	90.74	83.21	89.17	87.97	91.41	90.80	87.53	88.60	89.15
DPCL	0.05	9.81	9.23	17.43	2.06	6.03	14.74	22.01	11.98	12.24	5.08	6.63	15.64	4.30	17.91	16.57	8.66	13.14	11.32
	0.10	22.96	23.50	36.83	7.37	16.24	32.90	43.04	27.93	23.47	14.37	15.93	38.96	13.62	37.37	35.74	20.24	31.45	25.21
	0.15	35.97	35.17	51.52	16.60	23.94	47.66	54.38	40.76	33.46	26.75	25.34	52.18	24.19	48.47	49.03	34.22	44.78	37.01
	1.00	93.90	94.04	97.29	95.25	94.67	92.79	96.77	96.29	96.23	96.32	92.06	95.37	94.17	97.27	94.65	95.46	93.60	95.07
	mIoU	21.39	1.74	44.36	27.81	32.69	24.92	24.59	21.37	4.42	52.27	16.17	33.03	3.94	15.57	37.50	0.93	30.17	22.80
-		22																	
Method	α	fork	giraffe	horse 1	laptop	motor- cycle	mouse	person	sheep	skate- board	skis	stop- sign	tennis- racket	tie	tooth- brush	train	tv	zebra	all
Method	α 0.05	fork	giraffe 3.76	horse 1	laptop	motor- cycle 5.22	mouse	person 3.80	sheep	skate- board 5.96	skis 5.54	stop- sign 7.21	tennis- racket 7.21	tie 5.68	tooth- brush 4.85	train 7.59	tv 4.64	zebra	all
Method MMNet	α 0.05 0.10	fork 4.19 13.63	giraffe 3.76 16.24	horse 1 5.39 20.18	5.20 19.33	motor- cycle 5.22 21.38	mouse 4.26 16.91	person 3.80 16.11	sheep 7.78 30.81	skate- board 5.96 22.68	skis 5.54 23.15	stop- sign 7.21 25.53	tennis- racket 7.21 28.96	tie 5.68 24.08	tooth- brush 4.85 18.50	train 7.59 26.54	tv 4.64 17.16	zebra 4.35 17.11	all 5.68 21.58
Method MMNet [49]	α 0.05 0.10 0.15	fork 4.19 13.63 24.93	giraffe 3.76 16.24 32.49	horse 1 5.39 20.18 38.71	5.20 19.33 34.95	motor- cycle 5.22 21.38 42.19	mouse 4.26 16.91 33.02	person 3.80 16.11 32.80	sheep 7.78 30.81 55.63	skate- board 5.96 22.68 41.13	skis 5.54 23.15 44.39	stop- sign 7.21 25.53 44.90	tennis- racket 7.21 28.96 51.55	tie 5.68 24.08 47.83	tooth- brush 4.85 18.50 35.18	train 7.59 26.54 46.41	tv 4.64 17.16 31.58	zebra 4.35 17.11 34.30	all 5.68 21.58 39.66
Method MMNet [49]	α 0.05 0.10 0.15 1.00	fork 4.19 13.63 24.93 99.71	giraffe 3.76 16.24 32.49 99.82	horse 1 5.39 20.18 38.71 99.97	5.20 19.33 34.95 99.87	motor- cycle 5.22 21.38 42.19 98.34	mouse 4.26 16.91 33.02 99.78	person 3.80 16.11 32.80 99.99	sheep 7.78 30.81 55.63 99.73	skate- board 5.96 22.68 41.13 99.81	skis 5.54 23.15 44.39 99.89	stop- sign 7.21 25.53 44.90 99.44	tennis- racket 7.21 28.96 51.55 99.91	tie 5.68 24.08 47.83 100.00	tooth- brush 4.85 18.50 35.18 99.86	train 7.59 26.54 46.41 99.77	tv 4.64 17.16 31.58 99.96	zebra 4.35 17.11 34.30 99.72	all 5.68 21.58 39.66 99.76
Method MMNet [49]	α 0.05 0.10 0.15 1.00 0.05	fork 4.19 13.63 24.93 99.71 5.34	giraffe 3.76 16.24 32.49 99.82 15.85	horse 1 5.39 20.18 38.71 99.97 14.82	5.20 19.33 34.95 99.87 5.34	motor- cycle 5.22 21.38 42.19 98.34 11.43	4.26 16.91 33.02 99.78 8.19	2.80 3.80 16.11 32.80 99.99 13.22	sheep 7.78 30.81 55.63 99.73 17.82	skate- board 5.96 22.68 41.13 99.81 7.22	skis 5.54 23.15 44.39 99.89 11.47	stop- sign 7.21 25.53 44.90 99.44 18.19	tennis- racket 7.21 28.96 51.55 99.91 8.56	tie 5.68 24.08 47.83 100.00 15.73	tooth- brush 4.85 18.50 35.18 99.86 4.94	train 7.59 26.54 46.41 99.77 11.66	tv 4.64 17.16 31.58 99.96 10.42	zebra 4.35 17.11 34.30 99.72 14.93	all 5.68 21.58 39.66 99.76
Method MMNet [49] CATs	α 0.05 0.10 0.15 1.00 0.05 0.10	fork 4.19 13.63 24.93 99.71 5.34 13.81	giraffe 3.76 16.24 32.49 99.82 15.85 34.07	horse 1 5.39 20.18 38.71 99.97 14.82 35.04	5.20 19.33 34.95 99.87 5.34 15.94	motor- cycle 5.22 21.38 42.19 98.34 11.43 29.47	4.26 16.91 33.02 99.78 8.19 17.91	28.90 person	sheep 7.78 30.81 55.63 99.73 17.82 38.07	skate- board 5.96 22.68 41.13 99.81 7.22 19.67	skis 5.54 23.15 44.39 99.89 11.47 29.50	stop- sign 7.21 25.53 44.90 99.44 18.19 34.89	tennis- racket 7.21 28.96 51.55 99.91 8.56 20.57	tie 5.68 24.08 47.83 100.00 15.73 34.57	tooth- brush 4.85 18.50 35.18 99.86 4.94 12.58	train 7.59 26.54 46.41 99.77 11.66 28.52	tv 4.64 17.16 31.58 99.96 10.42 24.37	zebra 4.35 17.11 34.30 99.72 14.93 34.35	all 5.68 21.58 39.66 99.76 10.00 23.88
Method MMNet [49] CATs [4]	α 0.05 0.10 0.15 1.00 0.05 0.10 0.15	fork 4.19 13.63 24.93 99.71 5.34 13.81 21.97	giraffe 3.76 16.24 32.49 99.82 15.85 34.07 47.42	horse 1 5.39 20.18 38.71 99.97 14.82 35.04 49.38	5.20 19.33 34.95 99.87 5.34 15.94 27.41	motor- cycle 5.22 21.38 42.19 98.34 11.43 29.47 44.62	mouse           4.26           16.91           33.02           99.78           8.19           17.91           26.78	person 3.80 16.11 32.80 99.99 13.22 28.90 41.59	sheep 7.78 30.81 55.63 99.73 17.82 38.07 51.80	skate- board 5.96 22.68 41.13 99.81 7.22 19.67 31.79	skis 5.54 23.15 44.39 99.89 11.47 29.50 42.09	stop- sign 7.21 25.53 44.90 99.44 18.19 34.89 46.48	tennis- racket 7.21 28.96 51.55 99.91 8.56 20.57 32.19	tie 5.68 24.08 47.83 100.00 15.73 34.57 47.00	tooth- brush 4.85 18.50 35.18 99.86 4.94 12.58 19.92	train 7.59 26.54 46.41 99.77 11.66 28.52 41.49	tv 4.64 17.16 31.58 99.96 10.42 24.37 36.20	zebra 4.35 17.11 34.30 99.72 14.93 34.35 48.57	all 5.68 21.58 39.66 99.76 10.00 23.88 35.45
Method MMNet [49] CATs [4]	α 0.05 0.10 0.15 1.00 0.05 0.10 0.15 1.00	fork 4.19 13.63 24.93 99.71 5.34 13.81 21.97 88.63	giraffe 3.76 16.24 32.49 99.82 15.85 34.07 47.42 92.24	horse 1 5.39 20.18 38.71 99.97 14.82 35.04 49.38 91.64	1aptop 5.20 19.33 34.95 99.87 5.34 15.94 27.41 87.54	motor- cycle 5.22 21.38 42.19 98.34 11.43 29.47 44.62 90.34	mouse 4.26 16.91 33.02 99.78 8.19 17.91 26.78 82.90	person 3.80 16.11 32.80 99.99 13.22 28.90 41.59 89.58	sheep 7.78 30.81 55.63 99.73 17.82 38.07 51.80 92.71	skate- board 5.96 22.68 41.13 99.81 7.22 19.67 31.79 87.76	skis 5.54 23.15 44.39 99.89 11.47 29.50 42.09 89.40	stop- sign           7.21           25.53           44.90           99.44           18.19           34.89           46.48           86.95	tennis- racket 7.21 28.96 51.55 99.91 8.56 20.57 32.19 92.45	tie 5.68 24.08 47.83 100.00 15.73 34.57 47.00 89.30	tooth- brush 4.85 18.50 35.18 99.86 4.94 12.58 19.92 87.22	train 7.59 26.54 46.41 99.77 11.66 28.52 41.49 88.21	tv 4.64 17.16 31.58 99.96 10.42 24.37 36.20 91.96	zebra 4.35 17.11 34.30 99.72 14.93 34.35 48.57 91.76	all 5.68 21.58 39.66 99.76 10.00 23.88 35.45 89.15
Method MMNet [49] CATs [4]	α 0.05 0.10 0.15 1.00 0.05 0.10 0.15 1.00 0.05	fork 4.19 13.63 24.93 99.71 5.34 13.81 21.97 88.63 11.66	giraffe 3.76 16.24 32.49 99.82 15.85 34.07 47.42 92.24 14.77	horse 1 5.39 20.18 38.71 99.97 14.82 35.04 49.38 91.64 11.68	laptop 5.20 19.33 34.95 99.87 5.34 15.94 27.41 87.54 4.18	motor- cycle 5.22 21.38 42.19 98.34 11.43 29.47 44.62 90.34 11.41	4.26           16.91           33.02           99.78           8.19           17.91           26.78           82.90           8.74	person           3.80           16.11           32.80           99.99           13.22           28.90           41.59           89.58           8.99	sheep 7.78 30.81 55.63 99.73 17.82 38.07 51.80 92.71 21.38	skate- board 5.96 22.68 41.13 99.81 7.22 19.67 31.79 87.76 10.24	skis 5.54 23.15 44.39 99.89 11.47 29.50 42.09 89.40 6.15	stop- sign           7.21           25.53           44.90           99.44           18.19           34.89           46.48           86.95           6.93	tennis- racket 7.21 28.96 51.55 99.91 8.56 20.57 32.19 92.45 15.13	tie 5.68 24.08 47.83 100.00 15.73 34.57 47.00 89.30 17.30	tooth- brush 4.85 18.50 35.18 99.86 4.94 12.58 19.92 87.22 16.26	train 7.59 26.54 46.41 99.77 11.66 28.52 41.49 88.21 10.10	tv 4.64 17.16 31.58 99.96 10.42 24.37 36.20 91.96 2.48	zebra 4.35 17.11 34.30 99.72 14.93 34.35 48.57 91.76 13.46	all 5.68 21.58 39.66 99.76 10.00 23.88 35.45 89.15 11.32
Method MMNet [49] CATs [4]	α 0.05 0.10 0.15 1.00 0.05 0.10 0.15 1.00 0.05 0.10 0.05 0.10	fork 4.19 13.63 24.93 99.71 5.34 13.81 21.97 88.63 11.66 21.93	giraffe 3.76 16.24 32.49 99.82 15.85 34.07 47.42 92.24 14.77 31.38	horse 1 5.39 20.18 38.71 99.97 14.82 35.04 49.38 91.64 11.68 25.94	laptop 5.20 19.33 34.95 99.87 5.34 15.94 27.41 87.54 4.18 13.92	motor- cycle 5.22 21.38 42.19 98.34 11.43 29.47 44.62 90.34 11.41 29.86	mouse 4.26 16.91 33.02 99.78 8.19 17.91 26.78 82.90 8.74 15.99	person 3.80 16.11 32.80 99.99 13.22 28.90 41.59 89.58 8.99 23.73	sheep 7.78 30.81 55.63 99.73 17.82 38.07 51.80 92.71 21.38 42.47	skate- board 5.96 22.68 41.13 99.81 7.22 19.67 31.79 87.76 10.24 25.12	skis 5.54 23.15 44.39 99.89 11.47 29.50 42.09 89.40 6.15 16.14	stop- sign 7.21 25.53 44.90 99.44 18.19 34.89 46.48 86.95 6.93 15.39	tennis- racket 7.21 28.96 51.55 99.91 8.56 20.57 32.19 92.45 15.13 34.37	tie 5.68 24.08 47.83 100.00 15.73 34.57 47.00 89.30 17.30 27.93	tooth- brush 4.85 18.50 35.18 99.86 4.94 12.58 19.92 87.22 16.26 27.82	train 7.59 26.54 46.41 99.77 11.66 28.52 41.49 88.21 10.10 28.21	tv 4.64 17.16 31.58 99.96 10.42 24.37 36.20 91.96 2.48 8.19	zebra 4.35 17.11 34.30 99.72 14.93 34.35 48.57 91.76 13.46 29.22	all           5.68           21.58           39.66           99.76           10.00           23.88           35.45           89.15           11.32           25.21
Method MMNet [49] CATs [4] DPCL	α 0.05 0.10 0.15 1.00 0.05 0.10 0.15 1.00 0.05 0.10 0.05 0.10 0.15	fork 4.19 13.63 24.93 99.71 5.34 13.81 21.97 88.63 11.66 21.93 28.92	giraffe 3.76 16.24 32.49 99.82 15.85 34.07 47.42 92.24 14.77 31.38 41.82	horse 1 5.39 20.18 38.71 99.97 14.82 35.04 49.38 91.64 11.68 25.94 40.18	laptop 5.20 19.33 34.95 99.87 5.34 15.94 27.41 87.54 4.18 13.92 22.90	motor- cycle 5.22 21.38 42.19 98.34 11.43 29.47 44.62 90.34 11.41 29.86 45.19	mouse 4.26 16.91 33.02 99.78 8.19 17.91 26.78 82.90 8.74 15.99 27.00	person 3.80 16.11 32.80 99.99 13.22 28.90 41.59 89.58 8.99 23.73 38.62	sheep 7.78 30.81 55.63 99.73 17.82 38.07 51.80 92.71 21.38 42.47 54.26	skate- board 5.96 22.68 41.13 99.81 7.22 19.67 31.79 87.76 10.24 25.12 37.44	skis 5.54 23.15 44.39 99.89 11.47 29.50 42.09 89.40 6.15 16.14 24.24	stop- sign           7.21           25.53           44.90           99.44           18.19           34.89           46.48           86.95           6.93           15.39           26.69	tennis- racket 7.21 28.96 51.55 99.91 8.56 20.57 32.19 92.45 15.13 34.37 48.97	tie 5.68 24.08 47.83 100.00 15.73 34.57 47.00 89.30 17.30 27.93 38.57	tooth- brush 4.85 18.50 35.18 99.86 4.94 12.58 19.92 87.22 16.26 27.82 35.24	train 7.59 26.54 46.41 99.77 11.66 28.52 41.49 88.21 10.10 28.21 41.54	tv 4.64 17.16 31.58 99.96 10.42 24.37 36.20 91.96 2.48 8.19 20.10	zebra 4.35 17.11 34.30 99.72 14.93 34.35 48.57 91.76 13.46 29.22 43.36	all           5.68           21.58           39.66           99.76           10.00           23.88           35.45           89.15           11.32           25.21           37.01
Method MMNet [49] CATs [4] DPCL	α 0.05 0.10 0.15 1.00 0.05 0.10 0.15 1.00 0.05 0.10 0.15 1.00	fork 4.19 13.63 24.93 99.71 5.34 13.81 21.97 88.63 11.66 21.93 28.92 96.29	giraffe 3.76 16.24 32.49 99.82 15.85 34.07 47.42 92.24 14.77 31.38 41.82 94.66	horse 1 5.39 20.18 38.71 99.97 14.82 35.04 49.38 91.64 11.68 25.94 40.18 95.18	laptop 5.20 19.33 34.95 99.87 5.34 15.94 27.41 87.54 4.18 13.92 22.90 94.40	motor- cycle 5.22 21.38 42.19 98.34 11.43 29.47 44.62 90.34 11.41 29.86 45.19 91.28	mouse 4.26 16.91 33.02 99.78 8.19 17.91 26.78 82.90 8.74 15.99 27.00 93.06	person 3.80 16.11 32.80 99.99 13.22 28.90 41.59 89.58 8.99 23.73 38.62 95.77	sheep 7.78 30.81 55.63 99.73 17.82 38.07 51.80 92.71 21.38 42.47 54.26 96.21	skate- board 5.96 22.68 41.13 99.81 7.22 19.67 31.79 87.76 10.24 25.12 37.44 94.82	skis 5.54 23.15 44.39 99.89 11.47 29.50 42.09 89.40 6.15 16.14 24.24 94.87	stop- sign           7.21           25.53           44.90           99.44           18.19           34.89           46.48           86.95           6.93           15.39           26.69           93.02	tennis- racket 7.21 28.96 51.55 99.91 8.56 20.57 32.19 92.45 15.13 34.37 48.97 96.97	tie 5.68 24.08 47.83 100.00 15.73 34.57 47.00 89.30 17.30 27.93 38.57 94.53	tooth- brush 4.85 18.50 35.18 99.86 4.94 12.58 19.92 87.22 16.26 27.82 35.24 96.15	train 7.59 26.54 46.41 99.77 11.66 28.52 41.49 88.21 10.10 28.21 41.54 93.58	tv 4.64 17.16 31.58 99.96 10.42 24.37 36.20 91.96 2.48 8.19 20.10 95.61	zebra 4.35 17.11 34.30 99.72 14.93 34.35 48.57 91.76 13.46 29.22 43.36 96.62	all           5.68           21.58           39.66           99.76           10.00           23.88           35.45           89.15           11.32           25.21           37.01           95.07

### **Experimental Results**

Challenges and Visualizations:



Missing Keypoint



Inconsistent Prediction



Perspective Distortion

#### Future Direction

Unseen Key-point Discovery





Training





### Future Direction



Matching Closed-loop Constraint



Correspondence based Recognition Tasks

MISC210K: A Large-Scale Dataset for Multi-Instance Semantic Correspondence

# Thank You

Academy of Engineering & Technology, Fudan University, Shanghai, China School of Computer Science, Fudan University, Shanghai, China The University of Hong Kong, Hong Kong, China {wfge, wqzhang}@fudan.edu.cn