上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

SJTU

Deep Learning of Partial Graph Matching via Differentiable Top-K

Runzhong Wang, Ziao Guo, Shaofei Jiang, Xiaokang Yang, Junchi Yan

TUE-PM-205

饮 水 思 源 • 爱 国 荣 校

Overview

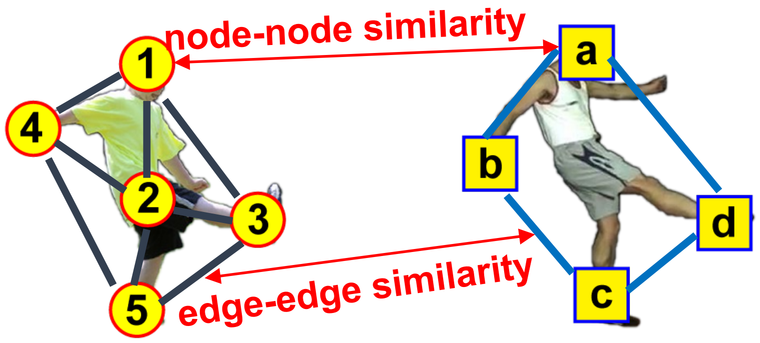
In this paper, we focus on **Partial Graph Matching** problem.

Contributions:

- Propose a top-k formulation and a differentiable top-k-based framework;
- Devise an attention-based graph neural network;
- Collect and remake a new visual graph matching benchmark named IMC-PT-SparseGM.

Background

NP-hard graph matching problem:



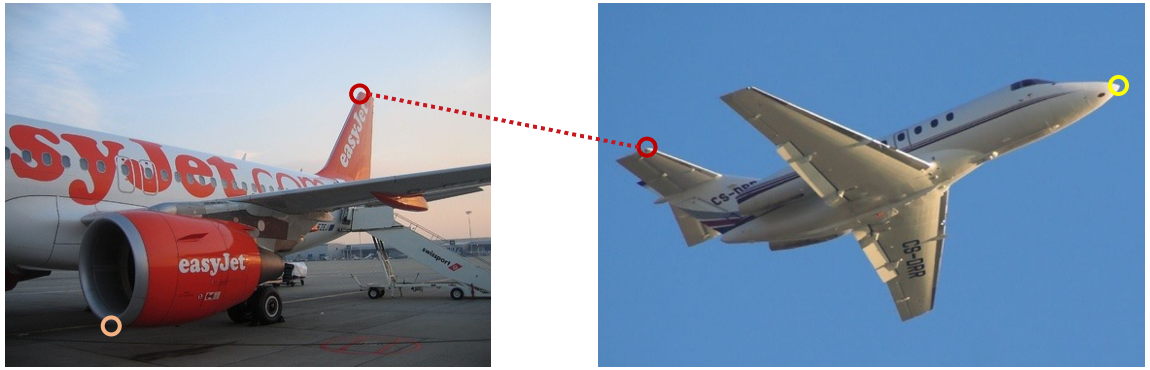
$$\max_{\mathbf{X}} \quad \mathrm{vec}(\mathbf{X})^T \mathbf{K} \, \mathrm{vec}(\mathbf{X})$$

$$\text{s.t.} \quad \mathbf{X} \in \{0,1\}^{5\times 4}$$

$$\mathbf{X}\mathbf{1} \leq \mathbf{1}, \quad \mathbf{X}^T\mathbf{1} = \mathbf{1}$$

Background

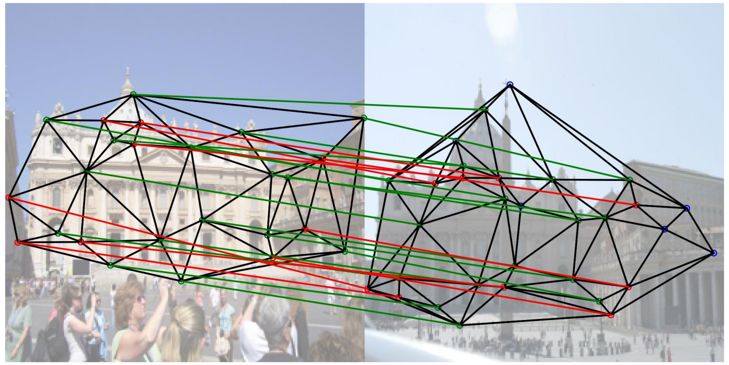
Partial graph matching:



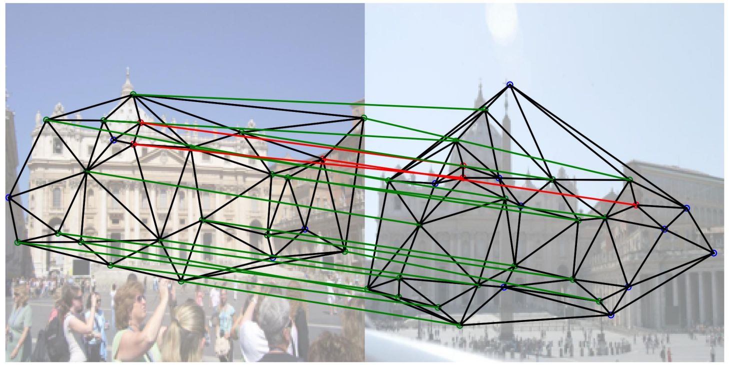
- Unavoidable due to the errors in keypoint detectors and occlusion of objects.
- In existing affinity-maximization graph matching pipeline, this problem is less handled. Also, an ‘intersection’ setting is usually adopted.

Motivation

- Existing partial matching handling (PMH) methods: thresholding, adding dummy rows and columns. **Inflexible**.
- Discard the matchings with small matching confidence. —> Preserve the matchings with **top-k** confidence values.
- k can be estimated by evaluating the **graph-level similarity**.

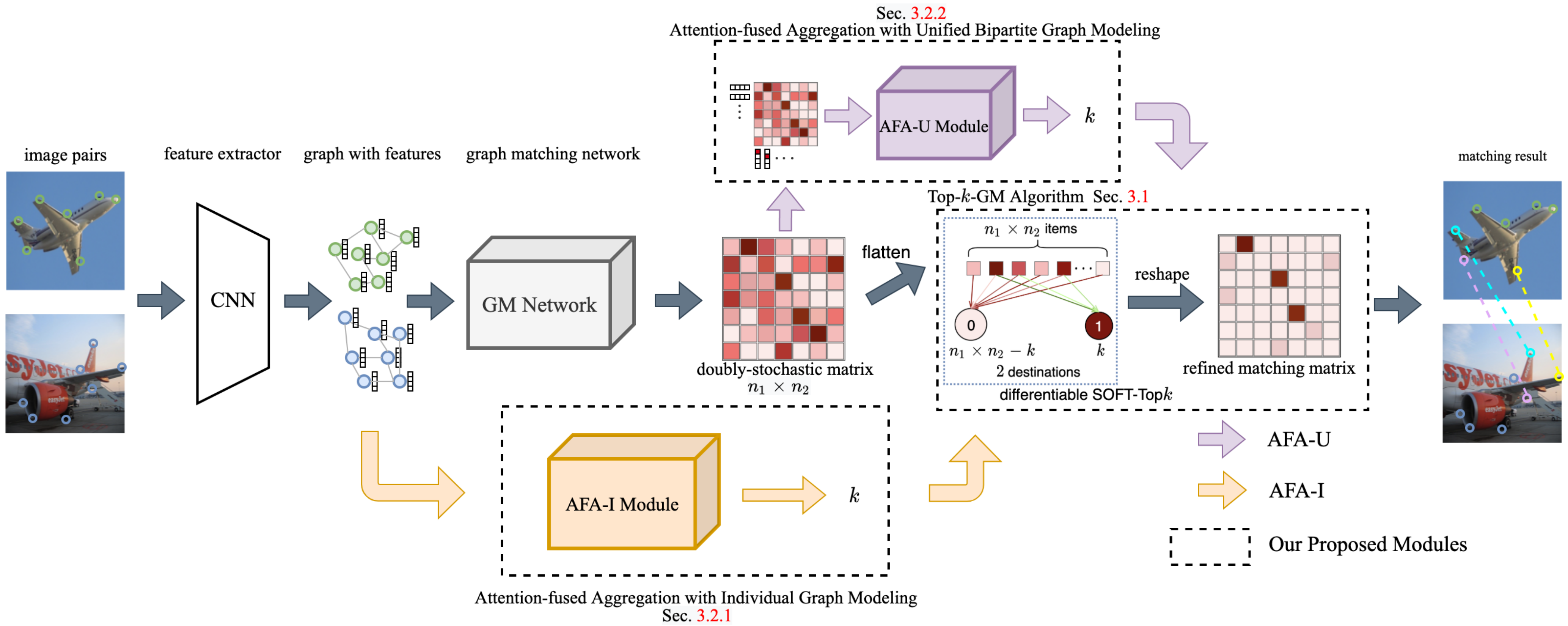


NGMv2-noPMH: precision:16/25 recall:16/20



NGMv2-AFAT: precision:18/21 recall:18/20

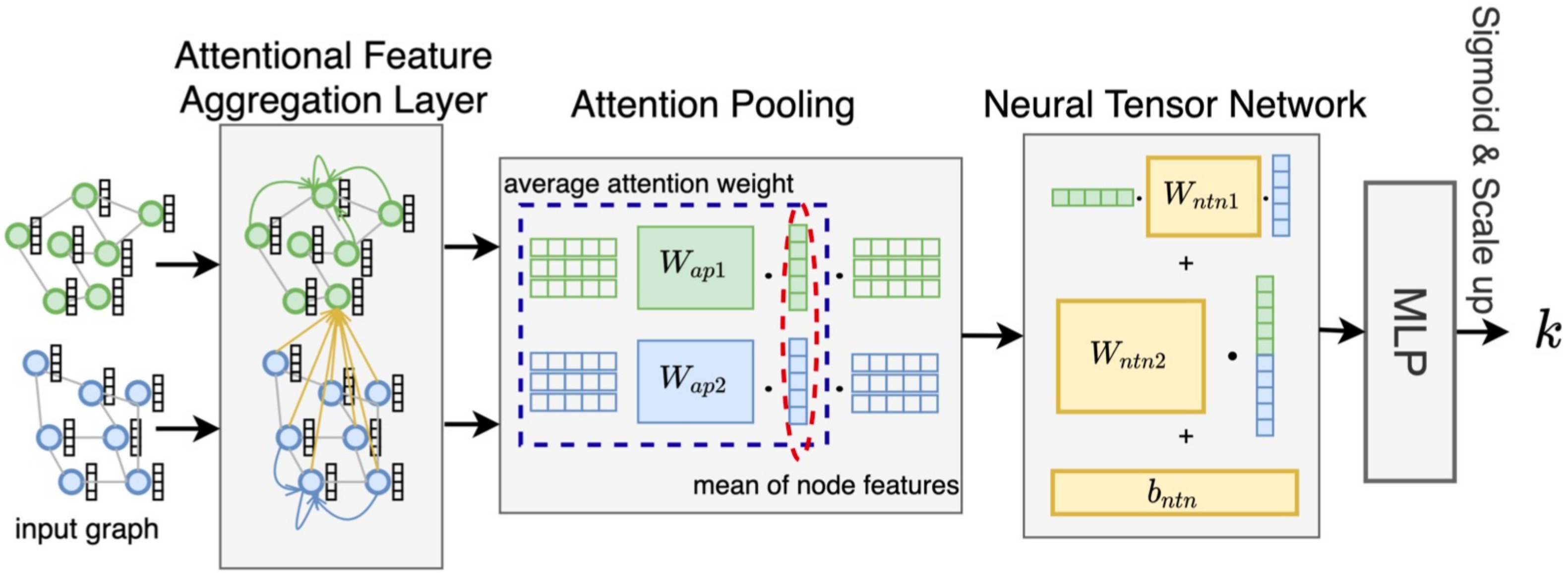
Pipeline



Sec. 3.2.2
Attention-fused Aggregation with Unified Bipartite Graph Modeling
AFA-U Module
k
image pairs
feature extractor
graph with features
graph matching network
CNN
GM Network
doubly-stochastic matrix
$n_1 \times n_2$
flatten
Top-k-GM Algorithm Sec. 3.1
$n_1 \times n_2$ items
0
1
$n_1 \times n_2 - k$
k
2 destinations
differentiable SOFT-Top$k$
reshape
refined matching matrix
matching result
AFA-I Module
k
Attention-fused Aggregation with Individual Graph Modeling
Sec. 3.2.1
AFA-U
AFA-I
Our Proposed Modules

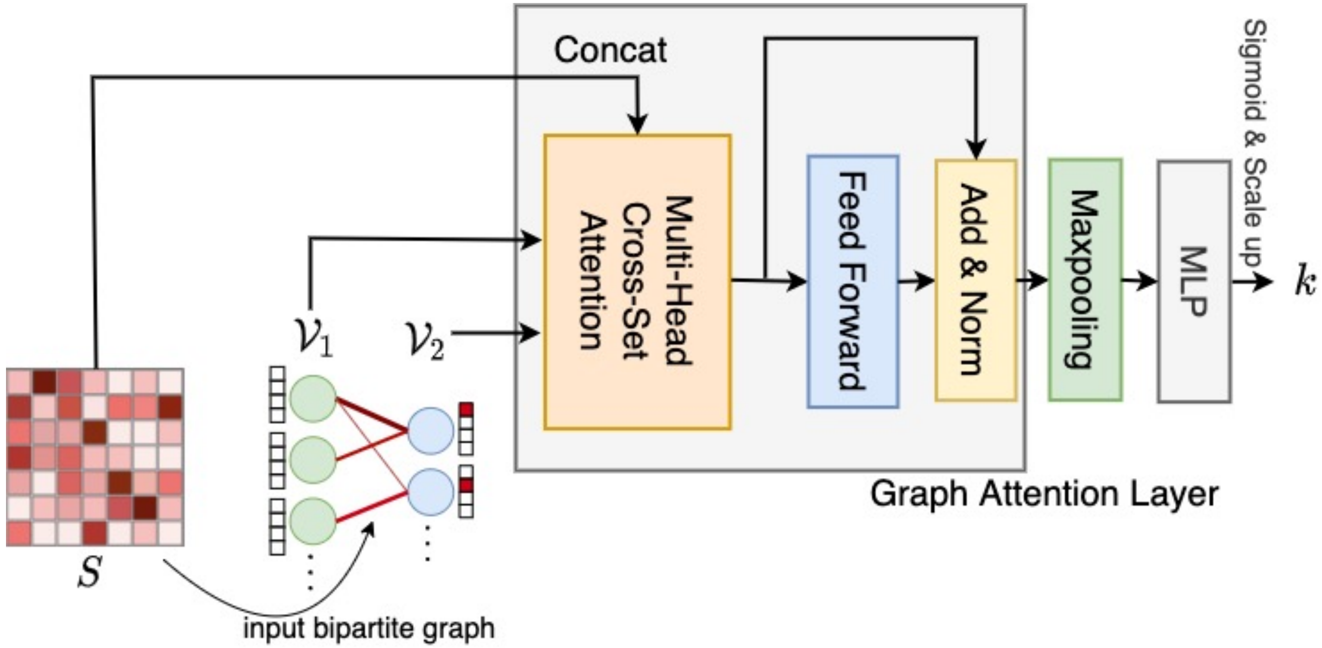
AFA-I: Individual Graph Modeling

Overview of AFA-I



AFA-U: Unified Bipartite Graph Modeling

Overview of AFA-U



Implementation Details

- Train GM networks and AFA in different stages.
  - Stage1: Train the GM network (using ground truth k)
  - Stage2: Train AFA modules
  - Stage3: Jointly train the GM network and AFA modules

- Sparse implementation of quadratic GM network NGMv2

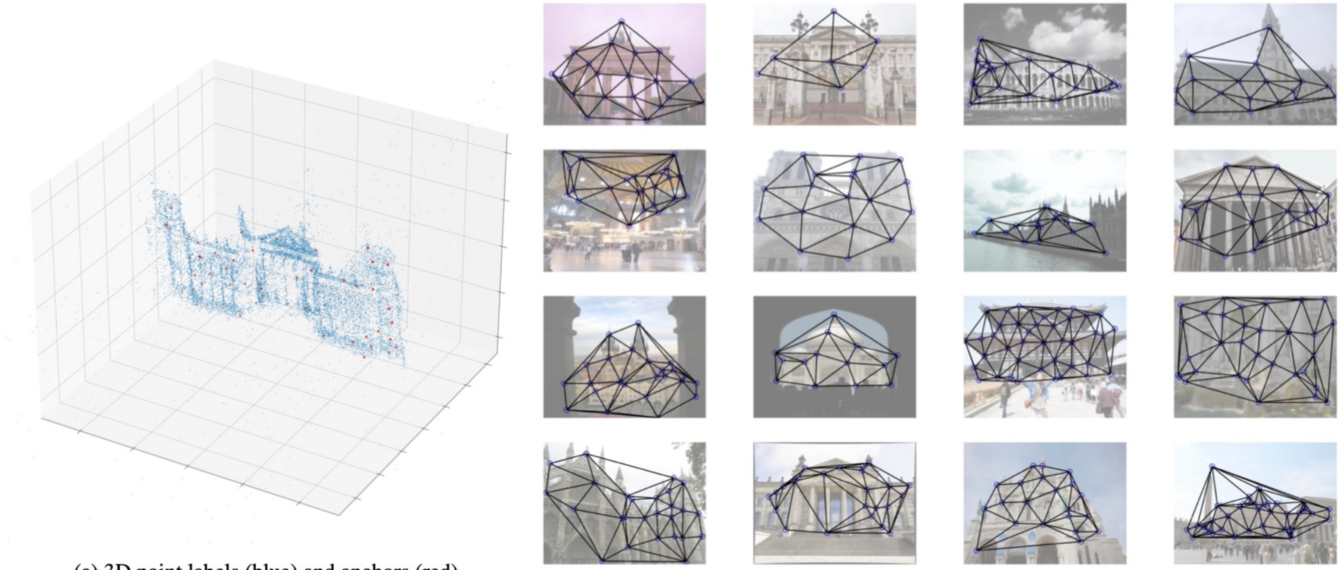
Table 3. The GPU memory cost (GB) of matching two graphs (batch size=1) w.r.t. dense and sparse NGMv2. The original dense version exceeds memory limit of RTX 3090 (24GB) at 110 nodes, whereas our sparse version scales-up more efficiently.

Number of Nodes	40	50	60	70	80	90	100	110
Dense NGMv2 [44]	2.82	3.05	4.42	6.26	8.80	12.48	17.59	<u>26.06</u>
Sparse NGMv2 (ours)	2.73	2.73	2.74	2.75	2.75	2.76	2.77	2.78

SJTU

Remade IMC-PT-SparseGM Benchmark

- We provide a new visual graph matching benchmark IMC-PT-SparseGM, based on the novel stereo benchmark Image Matching Challenge PhotoTourism (IMC-PT) 2020.



(a) 3D point labels (blue) and anchors (red)

(b) Examples (raw images with built graphs) for visual graph matching

Figure 6. The 3D points in (a) are detected by colmap [31, 32] which are available as labels in IMC-PT [18]. The blue points denote our selected anchors, based on which our IMC-PT SparseGM-50 is built, as shown in (b). The lines connecting anchors are the edges we build through Delaunay triangulation.

Table 1. Visual GM datasets. "partial rate" means the mean percentage of occluded keypoints w.r.t. the universe of all keypoints.

dataset name	# images	avg # nodes	# universe	partial rate
Willow Object Class [6]	404	10	10	0.0%
Pascal VOC Keypoint [5]	8702	9.07	6 to 23	28.5%
IMC-PT-SparseGM-50 (ours)	25667	21.36	50	57.3%
IMC-PT-SparseGM-100 (ours)	25667	44.48	100	55.5%

Experiments

- Unfiltered setting
- 2 GM backbone
  - NGMv2: SOTA quadratic matching network
  - GCAN: SOTA linear matching network
- Our PMH method: **A**ttention-**F**used **A**ggregation with **T**op-k-GM (**AFAT**)

Table 2. F1 (%) on Pascal VOC Keypoints (unfiltered). PMH means Partial Matching Handling. Our methods are marked as gray.

GM Network	PMH	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	mbike	person	plant	sheep	sofa	train	tv	mean
ZACR [43]	ZACR [43]	12.2	31.8	31.7	23.0	35.0	28.3	21.8	32.6	19.6	23.8	33.8	29.9	28.8	21.4	10.8	39.0	26.9	15.5	55.8	82.5	30.2
PCA-GM [44]	None	35.6	60.3	43.7	34.5	81.5	54.9	30.1	47.8	30.4	46.4	43.9	44.5	46.1	52.4	29.4	78.7	40.7	30.4	58.6	81.2	48.6
BBGM [29]	LPMP [38]	42.2	66.7	54.9	46.1	85.7	66.5	39.8	60.3	38.9	65.1	60.1	58.4	58.1	62.4	41.3	96.1	53.5	26.3	75.9	82.6	59.0
NGMv2 [45]	None	45.5	65.3	55.3	45.8	88.4	64.3	45.9	58.6	43.3	59.1	39.2	55.7	58.0	65.3	44.4	95.4	50.3	41.2	72.4	81.8	58.8
	Thresholding	48.3	65.4	55.3	48.6	87.6	63.0	51.1	61.1	39.6	63.3	33.6	59.2	59.3	63.4	46.9	95.2	53.5	45.5	73.4	81.4	59.4±0.4
	Dummy node	44.7	61.9	57.1	41.9	83.9	63.9	54.1	60.8	40.5	64.2	36.2	60.6	60.8	61.9	48.7	91.2	56.2	37.4	63.2	82.2	58.6±0.5
	AFAT-U (ours)	45.7	67.7	57.3	44.9	90.1	65.5	49.9	59.3	44.0	62.0	54.9	58.4	58.6	63.8	45.9	94.8	50.9	37.3	74.2	82.8	**60.2**±0.4
	AFAT-I (ours)	45.0	67.3	55.9	45.6	90.3	64.6	48.7	58.0	44.7	60.2	54.8	57.2	57.5	63.4	45.2	95.3	49.3	41.6	73.6	82.4	59.9±0.3
GCAN [17]	Dummy node	46.3	67.7	57.4	45.0	87.1	64.8	57.5	61.2	40.8	61.6	37.3	59.9	59.2	64.6	49.7	95.1	54.5	28.5	77.9	83.1	59.7±0.3
	AFAT-U (ours)	47.1	70.8	58.1	45.8	90.8	66.5	49.6	58.8	50.6	64.6	47.2	60.5	62.3	65.7	46.3	95.4	52.7	47.4	74.2	83.8	**62.0**±0.2
	AFAT-I (ours)	46.1	69.9	56.1	46.6	90.7	66.1	48.1	57.9	49.9	63.9	50.4	59.0	61.6	65.0	44.7	95.5	50.9	49.2	74.0	83.8	61.6±0.3

Experiments

Table 4. F1 (%) on Willow Object Class (+random outliers) and IMC-PT-SparseGM (50/100 anchors). Our methods are marked as gray.

Dataset name		Willow Object Class						IMC-PT-SparseGM (50 anchors)				IMC-PT-SparseGM (100 anchors)			
GM Network	PMH	car	duck	face	mbike	bottle	mean	reichstag	sacre_coeur	st_peters_square	mean	reichstag	sacre_coeur	st_peters_square	mean
ZACR [43]	ZACR [43]	47.3	44.7	77.7	39.9	53.6	52.6	72.1	33.7	29.5	45.1	39.4	33.1	30.4	34.3
PCA-GM [44]	None	55.8	56.5	81.2	46.4	58.1	59.6	83.4	47.5	58.5	63.1	70.7	43.1	58.8	57.5
BBGM [29]	LPMP [38]	65.1	60.7	85.5	71.6	65.5	69.7	85.4	55.1	59.3	66.6	88.1	55.0	56.4	66.5
NGMv2 [45]	None	78.9	66.6	84.3	63.1	76.0	73.8	90.8	55.9	64.3	70.3	78.4	54.9	69.3	67.6
	Thresholding	86.8	74.5	91.2	71.0	83.8	81.4±0.2	91.4	56.8	65.8	71.3±0.3	80.3	56.9	71.6	69.6±0.3
	Dummy node	83.3	69.7	95.7	68.8	86.7	80.8±0.4	88.5	56.1	63.0	69.2±0.5	80.0	57.0	71.3	69.5±0.3
	AFAT-U(ours)	82.6	74.5	90.6	73.9	87.0	81.7±0.5	90.5	58.7	66.9	72.0±0.3	81.7	57.0	72.2	**70.3**±0.2
	AFAT-I(ours)	84.6	75.7	92.0	74.5	88.6	**83.1**±0.2	92.3	58.7	66.7	**72.8**±0.4	82.0	57.0	71.4	70.1±0.3
GCAN [17]	Dummy node	74.8	75.7	92.8	77.1	83.5	80.8±0.2	87.2	55.1	63.0	68.4±0.5	80.4	55.7	72.8	69.6±0.4
	AFAT-U(ours)	80.1	78.0	90.6	76.0	87.0	82.3±0.3	86.9	59.4	67.1	71.1±0.4	82.6	58.2	73.8	**71.5**±0.2
	AFAT-I(ours)	82.2	77.7	92.7	77.2	88.6	**83.7**±0.3	91.0	60.3	67.3	**72.9**±0.6	82.7	57.8	72.4	70.9±0.4

Table 9. F1 (%) on SPair-71k (unfiltered setting). Our methods are marked as gray.

GM Network	PMH	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	dog	horse	mbike	person	plant	sheep	train	tv	mean
ZACR	ZACR	32.9	33.3	45.7	24.6	62.0	13.5	36.0	56.2	17.4	47.5	32.7	19.0	40.7	42.7	37.3	34.8	52.5	60.0	38.3
PCA-GM	None	36.5	25.6	48.9	24.7	50.7	29.1	19.2	54.6	30.1	39.1	42.9	34.0	31.3	27.1	70.5	31.1	56.6	75.2	40.4
BBGM	None	42.9	43.8	65.3	34.6	62.6	47.6	25.6	68.0	38.6	62.0	57.8	42.8	44.1	36.0	83.2	45.4	86.7	90.3	54.3
Ngmv2	None	45.4	42.3	61.0	31.2	62.2	53.3	34.2	65.3	37.0	59.5	54.7	41.3	44.8	38.9	77.5	44.2	77.8	89.9	53.4
	Thresholding	50.2	42.9	63.4	29.9	62.1	53.9	34.8	65.7	37.3	62.7	56.1	43.8	45.7	41.8	77.1	45.2	79.0	90.4	54.6±0.5
	Dummy node	47.7	41.6	62.1	30.3	59.0	49.7	27.4	68.3	33.9	62.4	57.3	46.7	46.4	42.7	78.7	43.5	80.5	89.5	53.8±0.4
	AFAT-U(ours)	50.3	43.5	63.8	32.4	59.0	60.1	39.7	68.6	36.1	63.6	56.5	46.3	51.4	43.3	77.0	51.2	81.1	89.4	56.3±0.4
	AFAT-I(ours)	50.4	43.6	63.9	32.1	61.2	58.5	38.0	68.4	35.7	62.7	56.4	47.7	51.9	44.3	78.5	50.7	79.2	91.2	**56.4**±0.6
GCAN	Dummy node	49.0	41.3	64.0	30.3	57.3	55.0	37.4	64.8	36.6	63.0	58.0	44.4	46.4	42.6	68.4	42.3	83.2	91.9	54.2±0.3
	AFAT-U(ours)	46.7	43.3	65.8	33.3	61.5	54.9	35.2	68.4	37.7	59.9	56.0	47.6	47.2	43.5	80.3	47.7	83.8	89.0	**55.7**±0.4
	AFAT-I(ours)	46.8	44.3	65.9	32.4	61.5	53.8	33.7	68.4	38.1	60.1	56.3	47.9	48.3	43.8	81.2	48.4	82.9	88.0	**55.7**±0.4

Ablation Study

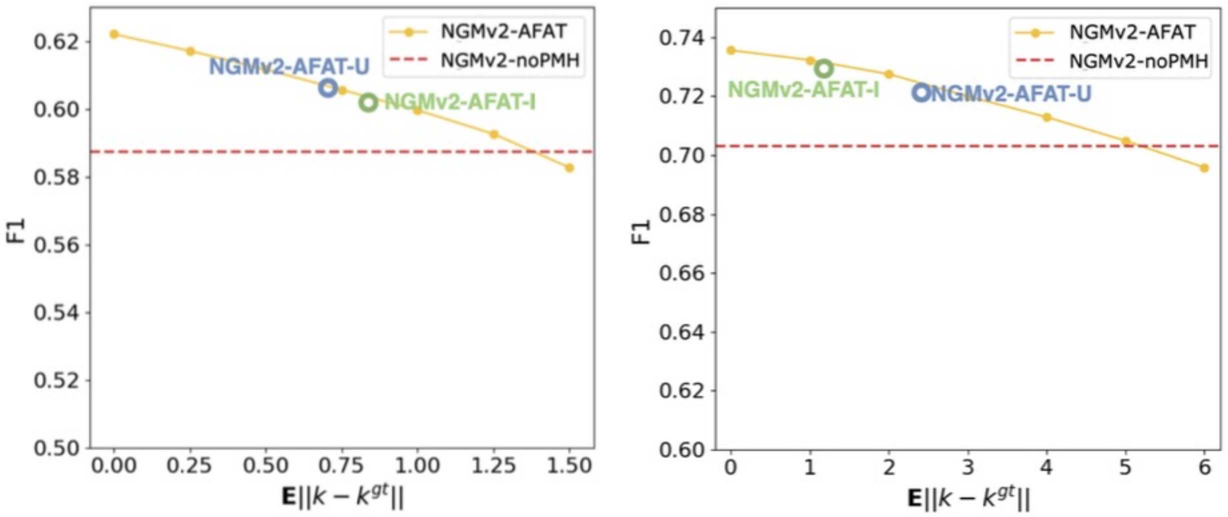
Necessity of learning in top-k-GM

Table 6. Ablation study addressing the necessity of learning. We report mean F1 (%), assuming that the ground truth k is given.

Dataset	Solver	Alg. 1 in test	Train	Alg. 1 in train	F1
PascalVOC	RRWM	✗	✗	✗	20.3
	RRWM	✓	✗	✗	19.4
	NGMv2	✓	✓	✗	60.7
	NGMv2	✓	✓	✓	**62.3**
WillowObject	RRWM	✗	✗	✗	20.1
	RRWM	✓	✗	✗	18.9
	NGMv2	✓	✓	✗	85.2
	NGMv2	✓	✓	✓	**85.5**
IMC-PT-SparseGM-50	RRWM	✗	✗	✗	39.6
	RRWM	✓	✗	✗	39.1
	NGMv2	✓	✓	✗	72.3
	NGMv2	✓	✓	✓	**73.6**
IMC-PT-SparseGM-100	RRWM	✗	✗	✗	34.6
	RRWM	✓	✗	✗	34.0
	NGMv2	✓	✓	✗	70.7
	NGMv2	✓	✓	✓	**71.8**

Ablation Study

Impact of the accuracy of k prediction



(a) Pascal VOC Keypoints (b) IMC-PT-SparseGM-50

Figure 5. Relations between matching F1 score and k error.

Conclusion

We propose a top-k-based framework to tackle the partial graph matching problem, which is ubiquitous in vision.

A new benchmark based on IMC-PT 2020, which is better suited for partial graph matching problem is remade and released.

Extensive experimental results on both classic and our new benchmarks show the effectiveness and significance of our work.

SJTU

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

THANK YOU

饮 水 思 源　爱 国 荣 校