Exploring Faithful Rationale for Multi-hop Fact Verification via Salience-Aware Graph Learning

Jiasheng Si, Yingjie Zhu, Deyu Zhou*

School of Computer Science and Engineering, Key Laboratory of Computer Network and Information Integration, Ministry of Education, Southeast University, China
{jasenchn, yj_zhu, d.zhou}@seu.edu.cn

AAAI2023

code: none

Reported by Xiaoke Li

Introduction

Claim: Olympic athlete May Wafic Sardouk represented Lebanon at the 1988 Summer Olympics in Seoul, Korea, landing in the 6th position in the Heat 4 event.

Evidence:

S_1(wiki/May Sardouk): *May Wafic Sardouk* (Arabic: وفيق مي سردوك; born June 4, 1963) is a *Lebanese* Olympic athlete.

S_2(wiki/May Sardouk): She *represented Lebanon in 1988 Summer Olympics in Seoul*.

S_3(wiki/May Sardouk): Sardouk and Nancy Khalaf were the only female participants for Lebanon in that tournament among a total of 21 participant for Lebanon.

S_4(wiki/Seoul): *Seoul*, officially the Seoul Special City, is the capital and largest metropolis of *South Korea*.

S_5(wiki/1988 Summer Olympics): The 1988 Summer Olympics, ···, was an international multi-sport event held from 17 September to 2 October 1988 in Seoul, South Korea.

T_6(wiki/May Sardouk):

Heat 4

Rank	Athlete	Time
1	Diane Dixon (USA)	52.45
2	Ute Thimm (FRG)	52.79
	.....	
6	May Sardouk (LIB)	1:00.01

Label: SUPPORTS

Figure 1: An example from FEVEROUS dataset, where *S1, S2, S4* and two table cells in *T6* are considered as rationales.

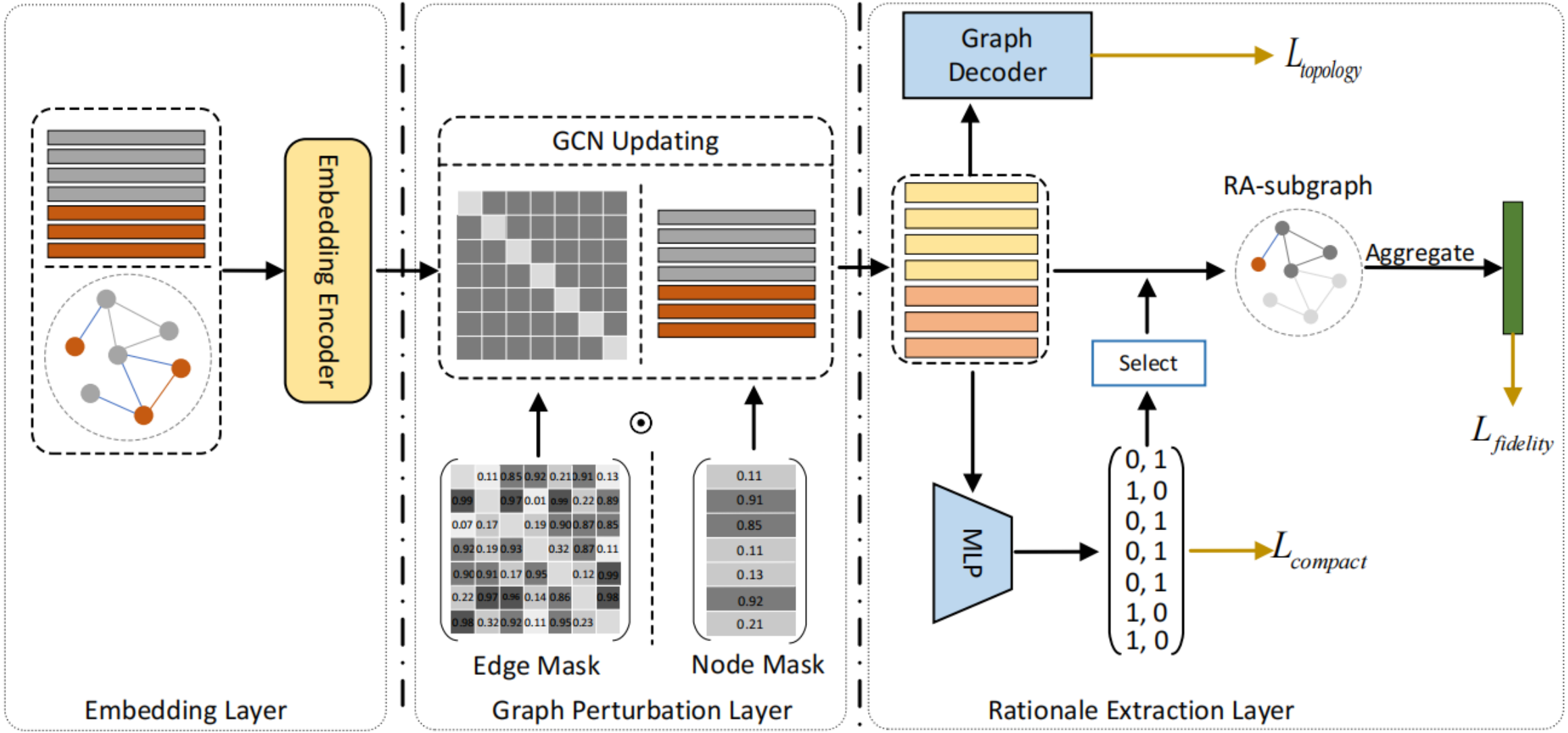

Figure 2: The overall framework of the proposed SaGP.

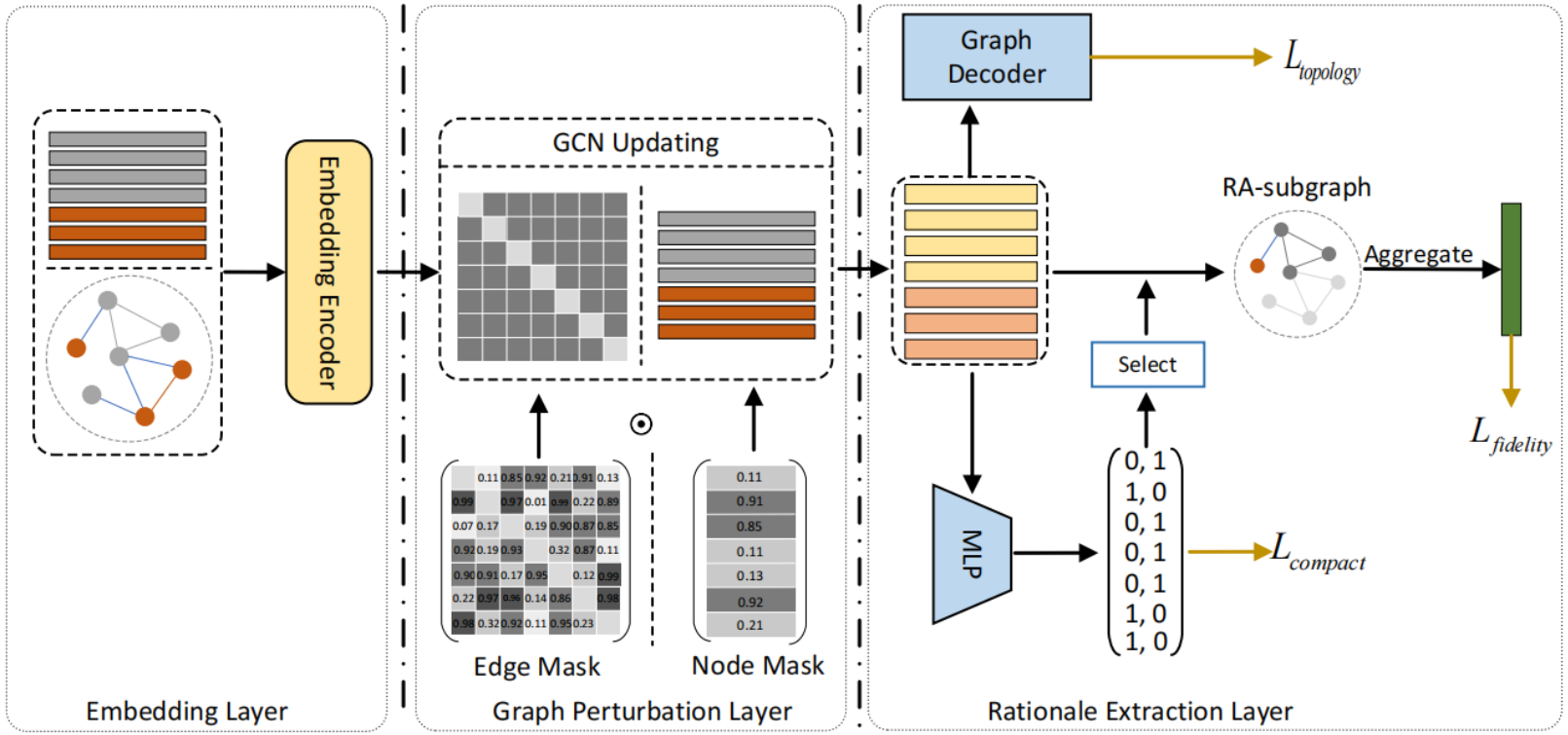

Figure 2: The overall framework of the proposed SaGP.

Let f be a trained GCN layer for node representation learning,

$$f(A, H; W) = relu(\widetilde{D}^{-1/2}\widetilde{A}\widetilde{D}^{-1/2}HW), \quad (1)$$

where $\widetilde{A} = A + I$, I is the identity matrix, $\widetilde{D}$ is the degree matrix, H denotes the evidence embeddings and W denotes the parameters of GCN.

mating 1: (I) the sum regulation $\mathcal{L}_{sum}$ of all entries in the perturbation matrix to constrain the size of the perturbation; (II) the information entropy regulation $\mathcal{L}_{entropy}$ to reduce the uncertainty of the perturbation matrix.

$$\widetilde{f}(A, H; W) = relu(((\widetilde{D}^{-1/2}\widetilde{A}\widetilde{D}^{-1/2}) \odot \sigma(P))HW), \quad (2)$$

where $\odot$ denotes the element product.

$$\widetilde{f}(A, H; W) = relu(\widetilde{D}^{-1/2}\widetilde{A}\widetilde{D}^{-1/2}(H \odot \sigma(M))W) \quad (3)$$

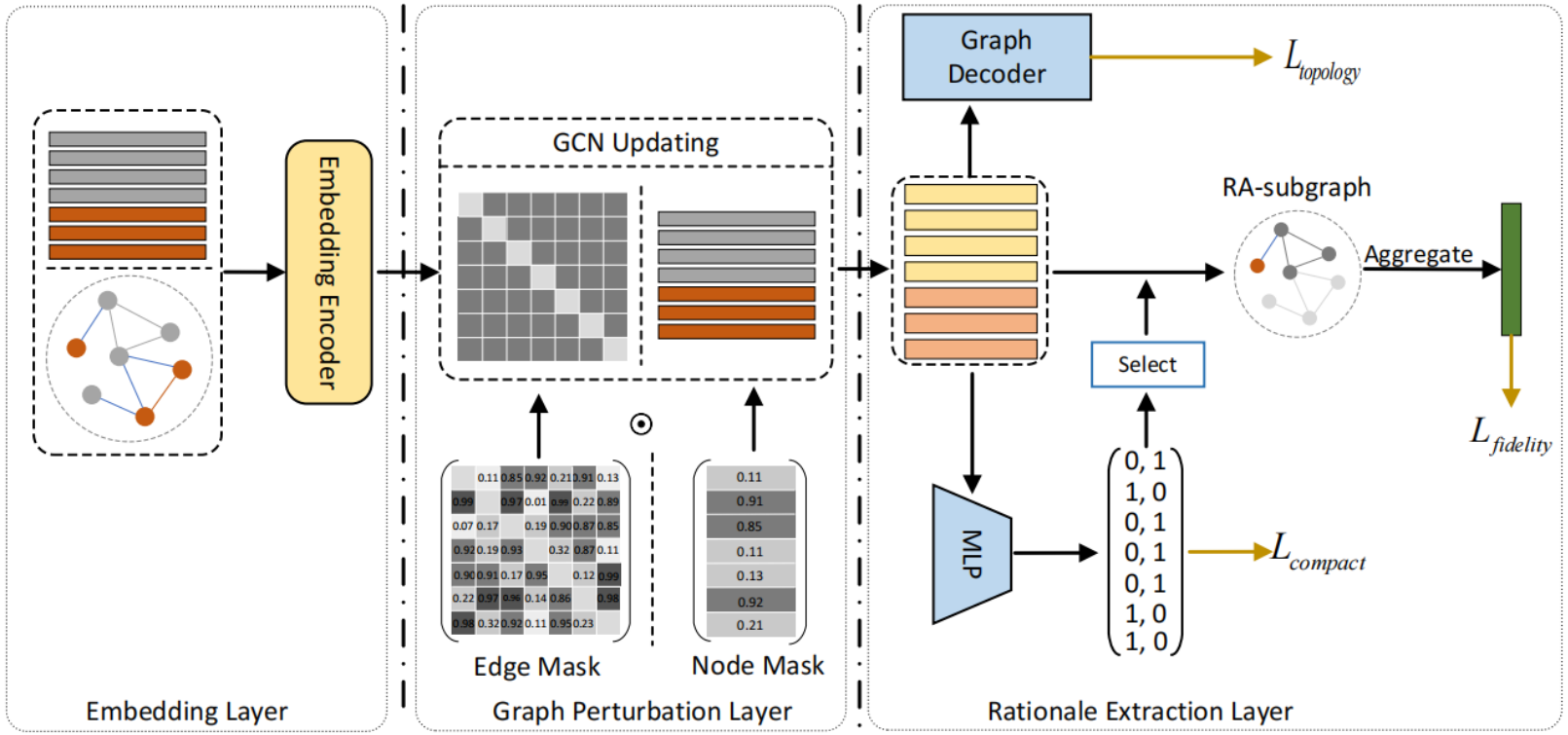

Figure 2: The overall framework of the proposed SaGP.

$$L_{compact} = \|norm(S^T AS) - I_2\|_F, \tag{6}$$

$$L_{topology} = CrossEntropy(\hat{A}, A), \quad \hat{A} = \sigma(\widetilde{U}\widetilde{U}^T) \tag{7}$$

$$\mathcal{L} = \lambda_1 \mathcal{L}_{fidelity} + \lambda_2 \mathcal{L}_{compact} + \lambda_3 \mathcal{L}_{topology} + \lambda_4 \mathcal{L}_{sum} + \lambda_5 \mathcal{L}_{entropy}, \tag{8}$$

$$S = softmax(MLP(\widetilde{U}; W_{sub})), \tag{4}$$

$$L_{fidelity} = CrossEntropy(\hat{y}^c_{sub}(\widetilde{U}), \hat{y}^c_{full}(U)), \tag{5}$$

Experiments

FEVEROUS	Num.Sup	Num.Ref	Avg.Ra	Avg.S	Avg.C
Train	41,835	27,215	4.85	1.43	3.42
Test	3,908	3,481	4.26	1.43	2.83

Table 1: Statistics of the FEVEROUS dataset. *Num.Sup* and *Num.Ref* are the number of claims with *SUPPORT* label and *REFUTE* label. *Avg.Ra*, *Avg.S*, and *Avg.C* denote the average number of *rationales*, *sentence rationales*, *table cell rationales* per claim, respectively.

Experiments

Model		Claim		Rationale				Claim & Rationale	
		F1.c	Acc.c	F1.r	Ext.acc.r	P.r	R.r	Acc.Part	Acc.Full
Unsupervised									
TSS-U		34.61	52.93	18.75	16.83	36.57	14.59	23.77	1.13
DeClarE		68.23	69.18	27.59	13.63	31.46	31.71	43.85	9.81
IB-U		77.30	77.30	65.28	20.08	78.01	67.30	75.36	15.76
	SaGP	**85.05**±0.02	**85.15**±0.02	80.08±0.01	45.33±0.05	79.15±0.03	88.30±0.01	**82.92**±0.03	**41.17**±0.05
Edge Mask	-T.	85.04±0.02	**85.15**±0.02	80.01±0.01	45.30±0.06	79.14±0.01	88.30±0.01	82.82±0.03	40.11±0.05
	-C.	85.04±0.05	85.15±0.07	**80.25**±0.16	**46.22**±1.41	**79.80**±0.97	87.68±1.09	82.85±0.06	41.14±1.57
	-T.&C.	85.01±0.04	85.11±0.04	80.15±0.01	45.23±0.03	79.14±0.01	**88.46**±0.01	82.92±0.05	40.01±0.01
	SaGP	82.24±0.13	**82.26**±0.13	70.47±0.08	38.56±0.13	**75.19**±0.12	76.40±0.05	75.03±0.08	33.61±0.01
Node Mask	-T.	**82.25**±0.12	82.25±0.12	**70.50**±0.09	**38.60**±0.10	**75.19**±0.12	76.37±0.07	75.04±0.06	**33.65**±0.04
	-C.	81.80±0.19	81.81±0.19	70.34±0.26	36.97±0.55	73.60±1.05	**78.28**±1.72	**75.36**±0.54	32.18±0.56
	-T.&C.	81.85±0.15	81.85±0.16	70.17±0.12	37.50±0.18	74.27±0.05	77.01±0.18	74.78±0.22	32.64±0.04
	SaGP	**82.06**±0.12	**82.08**±0.12	70.40±0.21	38.66±0.27	74.99±0.27	76.27±0.14	75.27±0.81	33.90±0.25
All	-T.	81.77±0.11	81.78±0.11	70.14±0.20	37.40±0.36	74.23±0.21	76.95±0.16	74.67±0.15	32.66±0.33
	-C.	81.89±0.09	81.90±0.09	**73.64**±4.80	**40.17**±3.60	**75.81**±2.09	**81.11**±5.70	**76.59**±2.54	**34.99**±3.03
	-T.&C.	82.03±0.11	82.05±0.11	70.38±0.23	38.60±0.26	74.93±0.28	76.30±0.15	74.64±0.05	33.84±0.24
Supervised									
BERT Blackbox		64.72	65.20	-	-	-	-	-	-
Pipeline		69.76	69.80	77.56	44.83	76.87	86.75	62.77	31.23
TSS-S		72.99	74.36	44.15	19.42	85.67	34.12	67.75	11.76
IB-S		79.14	79.17	65.68	20.08	78.91	67.31	76.70	16.37
Transformer-XH		74.05	74.33	76.70	49.10	79.43	80.47	69.17	40.22
	Edge Mask	**85.12**±0.01	**85.25**±0.01	80.49±0.02	48.22±0.01	81.18±0.02	86.14±0.02	**82.77**±0.01	43.36±0.01
SaGP	Node Mask	81.53±0.06	81.54±0.06	84.50±0.66	56.23±0.23	85.51±0.06	86.48±0.02	78.10±0.11	47.67±0.29
	All	82.10±0.04	82.15±0.03	**85.80**±0.07	**61.94**±0.26	**87.89**±0.07	**87.05**±0.06	78.76±0.11	**53.19**±0.30

Experiments

Model		FEVEROUS		
		Fidelity (↓)	Size (↑)	Sparsity (↓)
SaGP	Edge Mask	1.95±0.59	367.40±0.89	3.31±0.23
	-C.	1.53±0.02	361.12±0.01	4.81±0.22
	-T.	1.42±0.01	361.44±1.24	4.88±0.05
	-C. & T.	1.42±0.00	361.45±1.21	4.80±0.05

Table 3: Evaluation of the edge mask matrix. ↓ denotes the lower is better.

Claim: Vitaly Malkin was one of the founders of Rossiysky Kredit, which was converted into a joint stock company in 1997, but then in 2004, he officially retired from business.

Label: SUPPORTS

Rationales: 4 sentences

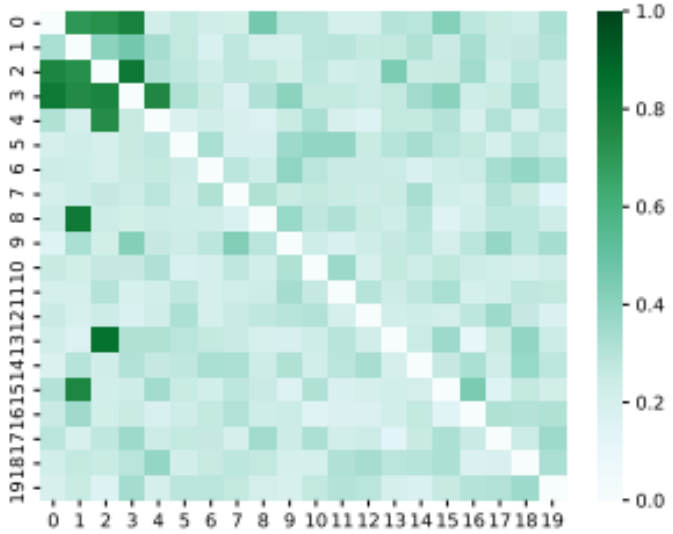

Figure 3: An example with visualization of the edge mask.

Experiments

Claim: The Victoria Falls are a flat plateau extending hundreds of kilometres, formed as the full width of the Zambezi River plummets in a single vertical drop, with the river's course dotted with islands as the river approaches the falls.

Rationales:

S_1(wiki/Victoria Falls): There are no mountains, escarpments, or deep valleys; only a flat plateau extending hundreds of kilometres in all directions. (score: 0.9333)

S_2(wiki/Victoria Falls):The falls are formed as the full width of the river plummets in a single vertical drop into a transverse chasm 1,708 metres (5,604 ft) wide, carved by its waters along a fracture zone in the basalt plateau. (score: 0.8963)

S_3(wiki/Victoria Falls): The river's course is dotted with numerous tree-covered islands, which increase in number as the river approaches the falls. (score: 0.9382)

T_4(wiki/Victoria Falls):

Victoria Falls	
Type	Cataract waterfall
Watercourse	Zambezi River
Total height	108 m (355 ft)

(score: 0.0383)

Label: SUPPORTS

Figure 4: A case with failing to identify rationales within $T4$.

Thanks