16-11 July · Seoul, South Korea

FACT: Fuzzy Alignment with Comorbidity Topology for Reliable Multi-Label Medical Image Diagnosis

Yingyu Chen[1], Yongqiang Huang[2], Yang Qin[1], Ziyuan Yang[1], Lang Yuan[2], Maosong Ran[1], Yi Zhang[2,*]

[1]College of Computer Science, Sichuan University
[2]School of Cyber Science and Engineering, Sichuan University

Background

Single-label Classification

Natural images often contain one primary object

The task is to assign a single label (e.g., "cat")

direct extension

suboptimal

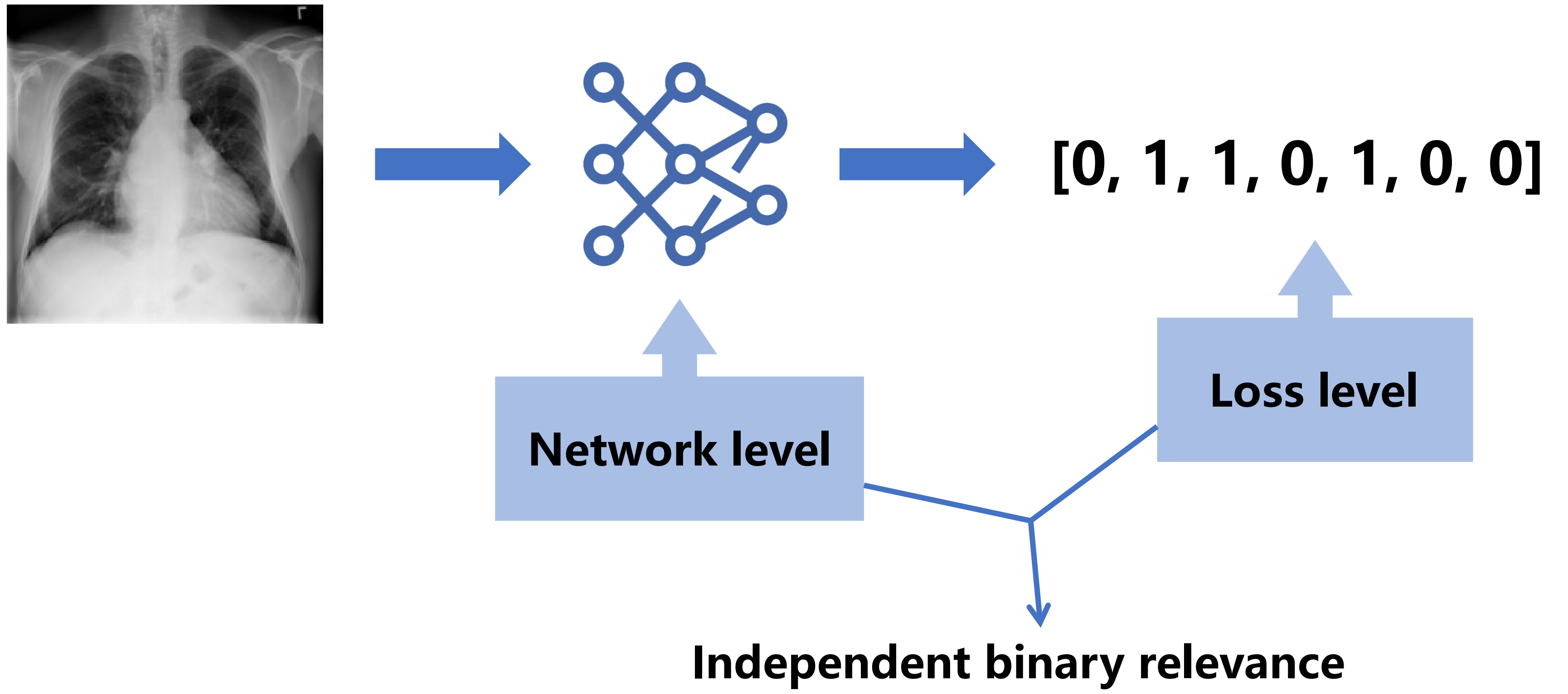
Multi-label Diagnosis (MLD)

Medical images often present **multiple co-occurring conditions** simultaneously, which are inherently more complex

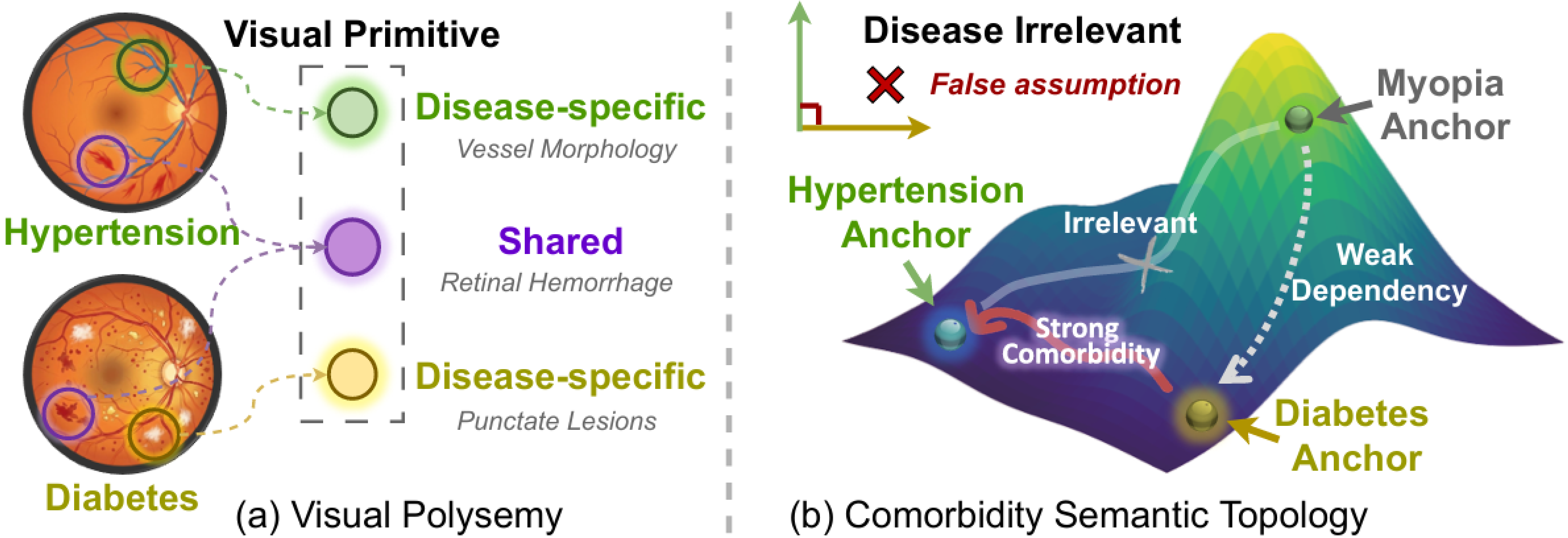
Chest X-ray (e.g., Pneumonia, Effusion, Cardiomegaly).

Background

Current Work

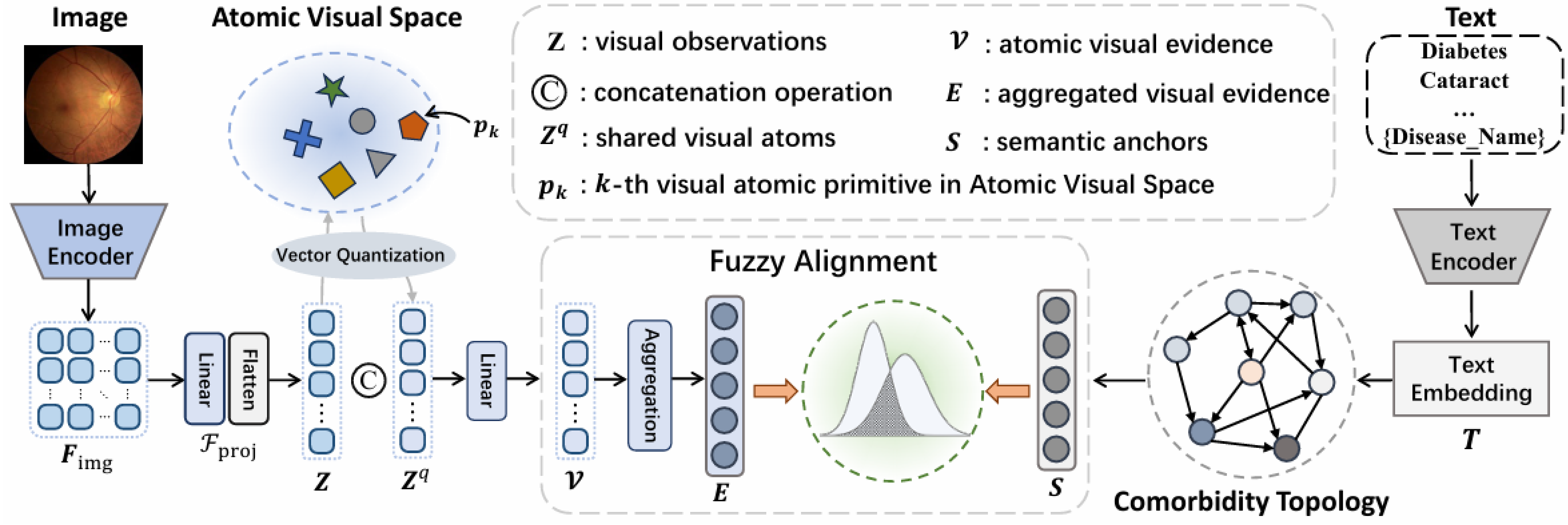


Motivation


Visual Primitive
Hypertension
Diabetes
Disease-specific
Vessel Morphology
Shared
Retinal Hemorrhage
Disease-specific
Punctate Lesions
(a) Visual Polysemy
Disease Irrelevant
False assumption
Hypertension Anchor
Irrelevant
Myopia Anchor
Weak Dependency
Strong Comorbidity
Diabetes Anchor
(b) Comorbidity Semantic Topology

Methodology

(3) Fuzzy Alignment
- **(1) Atomic Visual Space → captures visual polysemy**
- **(2) Comorbidity Topology → encodes disease relationship**



Methodology

(1) Atomic Visual Space

$$\boldsymbol{Z} = \mathcal{F}_{\text{proj}}(\boldsymbol{F}_{\text{img}}) \in \mathbb{R}^{M \times d} \quad M = H \times W$$

Vector Quantization (VQ)

$$z_m^q = p_k, \quad \text{where } k = \underset{j \in \{1,\ldots,K\}}{\arg\min} \|z_m - p_j\|_2.$$

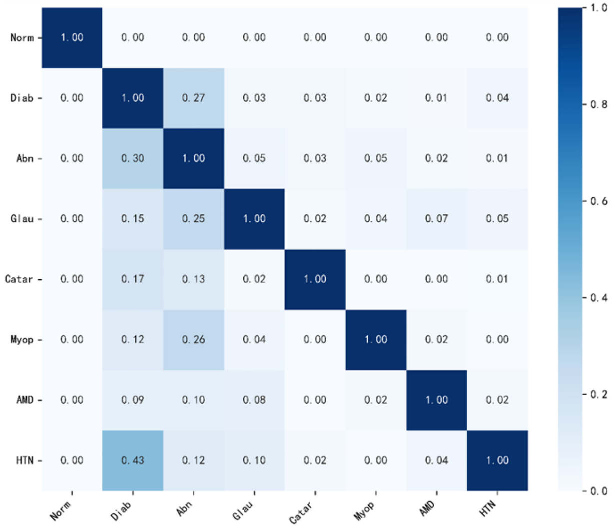
Methodology

(2) Comorbidity Topology

GCN $$\mathbf{S}^{(l+1)} = \mathrm{RELU}\left(\tilde{\mathbf{D}}^{-\frac{1}{2}}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-\frac{1}{2}}\mathbf{S}^{(l)}\mathbf{W}^{(l)}\right).$$

① **Adjacency matrix** A $\leftarrow$ **correlation matrix**

② **Node features** S^0 $\leftarrow$ **disease text embedding**



Methodology

(3) Fuzzy Alignment

Visual feature (VQ) $\mathcal{V} \in \mathbb{R}^{M\times d}$ **Semantic anchor (GCN)** $\mathcal{S} \in \mathbb{R}^{C\times d}$

Semantic interface operator $E = \mathcal{I}(\mathcal{V}) = Q + \mathrm{Sigmoid}\left(\frac{Q\mathcal{V}^\top}{\sqrt{d}}\right)\mathcal{V}$

Fuzzy membership $\mu_c = \exp\left(-\frac{\|e_c - s_c\|^2}{\tau}\right)$

Derived from continuity & semantic consistency (see appendix A.1)

Methodology

Fuzzy loss

$$\mathcal{L}_{\text{fuzzy}} = \mathcal{L}_{\text{point}}(\boldsymbol{\mu}, \mathbf{y}) + \alpha \cdot \mathcal{L}_{\text{set}}(\boldsymbol{\mu}, \mathbf{y}).$$

(see appendix A.3)

Point loss

$$\mathcal{L}_{\text{point}} = \log\left(1 + \sum_{c:y_c=0} r_c\right) + \log\left(1 + \sum_{c:y_c=1} r_c\right)$$

Set loss

$$\mathcal{L}_{\text{set}} = 1 - \mathcal{J}(\boldsymbol{\mu}, \mathbf{y})$$

Total loss

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{fuzzy}} + \gamma \cdot \mathcal{L}_{commit}$$

$$\mathcal{L}_{\text{commit}} = \|\text{sg}[\boldsymbol{Z}] - \boldsymbol{Z}^q\|_2^2 + \beta\|\boldsymbol{Z} - \text{sg}[\boldsymbol{Z}^q]\|_2^2$$

Experiment

Dataset	Modal	Data Scale	Num of Classes	Other Description
ODIR-5K	Color Fundus image	5000 (tiny)	8	N/A
NIH-Chest	X-ray	112,120 (medium)	14	Long-tailed
RSNA-IHD	CT	752,802 (large)	5	Long-tailed
CXR-LT	X-ray	377,110 (large)	40	Extremely Long-tailed & Noisy

Experiment

Methods	Type	Ref.	Backbone	Param (M)	Flops (G)
TransferNet (2021)	Network- centric	BSPC'21	ResNet50	23.52	4.13
ASL (2021)	Objective-centric	CVPR'21	ResNet50	23.52	4.13
RAL (2023)	Objective-centric	ICCV'23	ResNet50	23.52	4.13
TADCL (2023)	Network- centric	MIA'23	ViT	61.36	8.84
Two-Way (2023)	Objective-centric	CVPR'23	ResNet50	23.52	4.13
CTransCNN (2023)	Network- centric	KBS'23	CNN+ViT	33.99	10.57
LDR (2024)	Network- centric	ACMMM'24	ResNet50	25.51	8.42
SupCon (2024)	Objective-centric	AAAI'24	ResNet50	23.52	4.13
MultiCo (2025)	Objective-centric	arxiv'25	ResNet50	23.52	4.13
HydraViT (2025)	Network- centric	BSPC'25	ResNet50	39.22	16.58
FACT (Ours)	–	–	ResNet50	25.87	4.21

Experiment

Methods	ODIR-5K				NIH-Chest				RSNA-IHD			
	mAP ↑	F1↑	AUC↑	Avg.↑	mAP↑	F1↑	AUC↑	Avg.↑	mAP↑	F1↑	AUC↑	Avg.↑
TransferNet (2021)	64.30	86.29	89.01	79.87	51.71	90.12	86.73	76.19	84.53	90.36	95.44	90.11
ASL (2021)	59.69	87.84	90.41	79.31	42.32	86.31	85.41	71.35	83.07	91.97	96.15	90.40
RAL (2023)	59.61	86.86	89.77	78.75	38.63	83.51	83.78	68.64	82.91	92.02	96.24	90.39
TADCL (2023)	65.04	86.57	86.45	79.35	41.05	86.91	84.22	70.73	87.47	92.07	96.32	91.95
Two-Way (2023)	55.93	84.13	91.17	77.08	46.38	84.81	86.70	72.63	57.51	79.17	94.37	77.02
CTransCNN (2023)	69.30	90.29	91.35	83.65	52.40	87.88	85.61	75.30	85.99	90.45	94.69	90.38
LDR (2024)	69.69	89.79	90.81	83.43	57.81	90.01	87.52	78.45	84.85	89.41	94.54	89.60
SupCon (2024)	67.43	86.75	90.62	81.60	58.82	89.28	86.08	78.06	84.59	90.03	95.12	89.91
MultiCo (2025)	68.30	89.14	91.45	82.96	54.32	90.14	87.04	77.17	86.70	92.11	96.22	91.68
HydraViT (2025)	67.92	88.25	90.22	82.13	54.60	90.33	86.96	77.30	86.39	91.22	96.33	91.31
FACT (Ours)	**72.61**	**90.91**	**92.97**	**85.50**	**58.95**	**90.44**	**87.81**	**79.07**	**89.72**	**92.97**	**96.90**	**93.20**
△	+2.92	+0.62	+1.52	+1.85	+0.13	+0.11	+0.29	+0.62	+2.25	+0.86	+0.57	+1.25

"△" represents the difference in metrics between our method and the second-best method.

Experiment

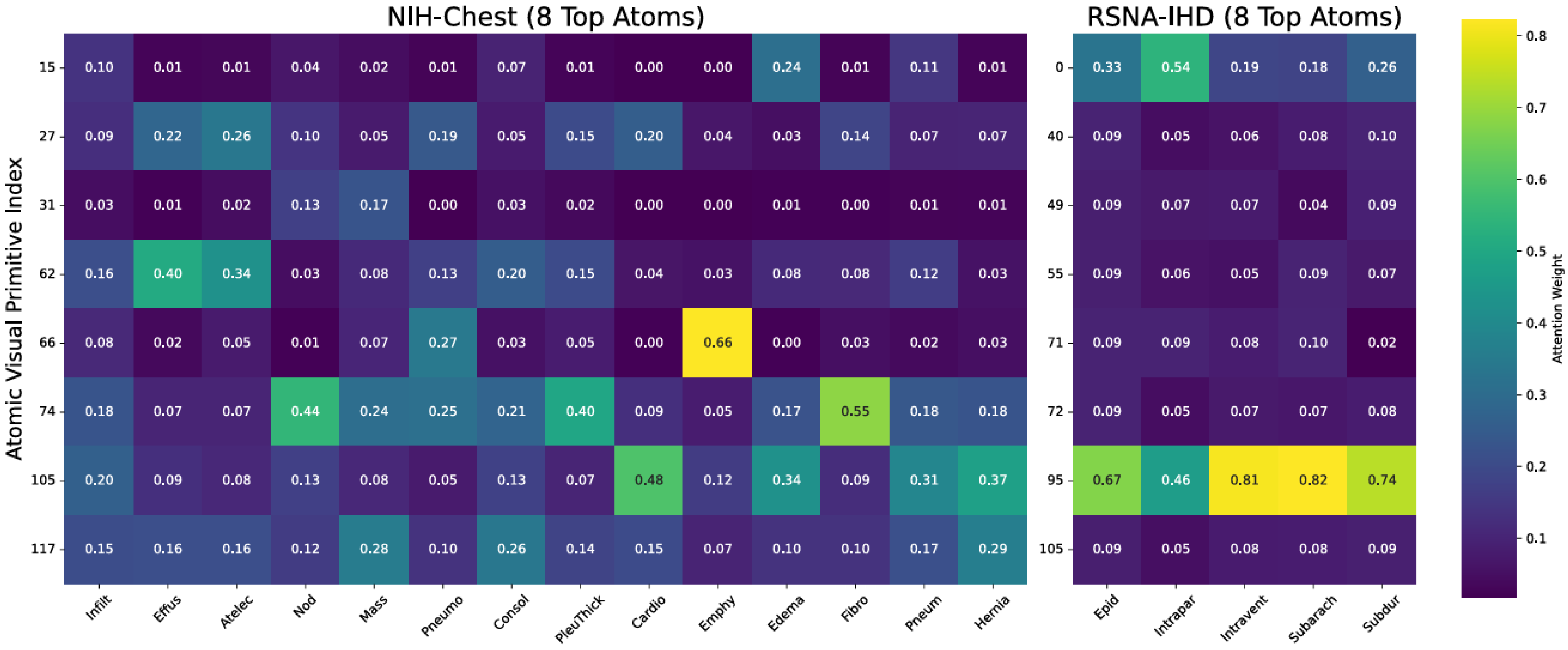
(see Tab. 14)

Diseases	Prevalence	Group	Methods				
			MultiCo(2025)	LDR(2024)	HydraViT(2025)	Ours	△
Adenopathy	1.19%	Middle	68.50	70.69	76.28	**78.55**	+2.27
Atelectasis	28.26%	Head	79.42	80.05	80.11	**80.19**	+0.08
Azygos Lobe	0.06%	Tail	69.77	73.28	69.72	**85.39**	+12.11
Calcification of the Aorta	1.46%	Middle	76.90	85.88	81.74	**87.95**	+2.07
Cardiomegaly	30.38%	Head	79.04	79.45	**79.52**	78.58	-0.94
Clavicle Fracture	0.06%	Tail	58.55	62.96	64.67	**76.23**	11.56
Consolidation	6.82%	Middle	74.90	76.62	**76.98**	76.60	-0.38
Edema	17.18%	Head	82.82	83.39	**84.15**	83.40	-0.75
Emphysema	1.36%	Middle	82.00	86.98	**88.60**	88.19	0.41
Enlarged Cardiomediastinum	11.18%	Head	55.34	55.49	**59.45**	58.61	-0.84
Fibrosis	0.41%	Tail	76.40	85.06	87.89	**89.21**	+1.32
Fissure	1.00%	Middle	59.03	67.64	68.02	**76.63**	+8.61
Fracture	4.34%	Middle	64.02	68.41	70.50	**74.05**	+3.55
Granuloma	0.97%	Tail	70.36	73.46	73.82	**76.16**	+2.34
Hernia	1.40%	Middle	74.88	**85.77**	84.76	85.57	-0.20
Hydropneumothorax	0.26%	Tail	74.78	90.67	93.14	**93.82**	+0.68
Infarction	0.29%	Tail	49.14	51.67	56.27	**61.04**	+4.77
Infiltration	3.72%	Middle	54.65	56.80	57.71	**58.10**	+0.39
Kyphosis	0.20%	Tail	78.13	87.99	86.07	**90.19**	+2.20
Lobar Atelectasis	0.06%	Tail	48.99	82.26	78.35	**85.03**	+2.77
Lung Lesion	0.92%	Tail	65.33	69.70	73.82	**76.36**	+2.54
Lung Opacity	31.69%	Head	74.07	74.83	75.64	**75.77**	+0.13
Mass	2.03%	Middle	70.21	73.62	75.34	**77.45**	+2.21
Nodule	2.63%	Middle	69.90	72.41	72.91	**74.91**	+2.00
Normal	12.28%	Head	79.27	78.98	**79.58**	79.42	-0.16
Pleural Effusion	29.27%	Head	89.42	89.89	**90.42**	89.92	-0.50
Pleural Other	0.20%	Tail	64.56	81.37	82.13	**83.49**	+1.36
Pleural Thickening	1.12%	Middle	71.28	77.81	79.55	**83.25**	+0.70
Pneumomediastinum	0.31%	Tail	56.93	73.61	80.06	**83.41**	+3.35
Pneumonia	16.84%	Head	60.35	61.92	**62.68**	62.49	-0.19
Pneumoperitoneum	0.24%	Tail	61.81	71.14	74.33	**79.91**	+5.58
Pneumothorax	6.31%	Middle	78.86	80.56	82.67	**82.70**	+0.03
Pulmonary Embolism	0.66%	Tail	49.56	51.52	58.94	**59.37**	+0.43
Pulmonary Hypertension	0.35%	Tail	58.81	75.02	76.71	**77.94**	+1.23
Rib Fracture	3.29%	Middle	62.23	67.68	70.88	**74.14**	+3.26
Round(ed) Atelectasis	0.05%	Tail	57.21	92.11	**92.59**	90.18	-2.41
Subcutaneous Emphysema	0.97%	Tail	90.49	94.16	95.59	**96.08**	+0.49
Support Devices	41.42%	Head	91.79	91.75	91.85	**92.53**	+0.68
Tortuous Aorta	1.11%	Middle	73.47	75.63	78.35	**79.68**	+1.33
Tuberculosis	0.72%	Tail	69.28	71.47	76.03	**76.60**	+0.57

"△" represents the difference in metrics between our method and the second-best method.

Experiment

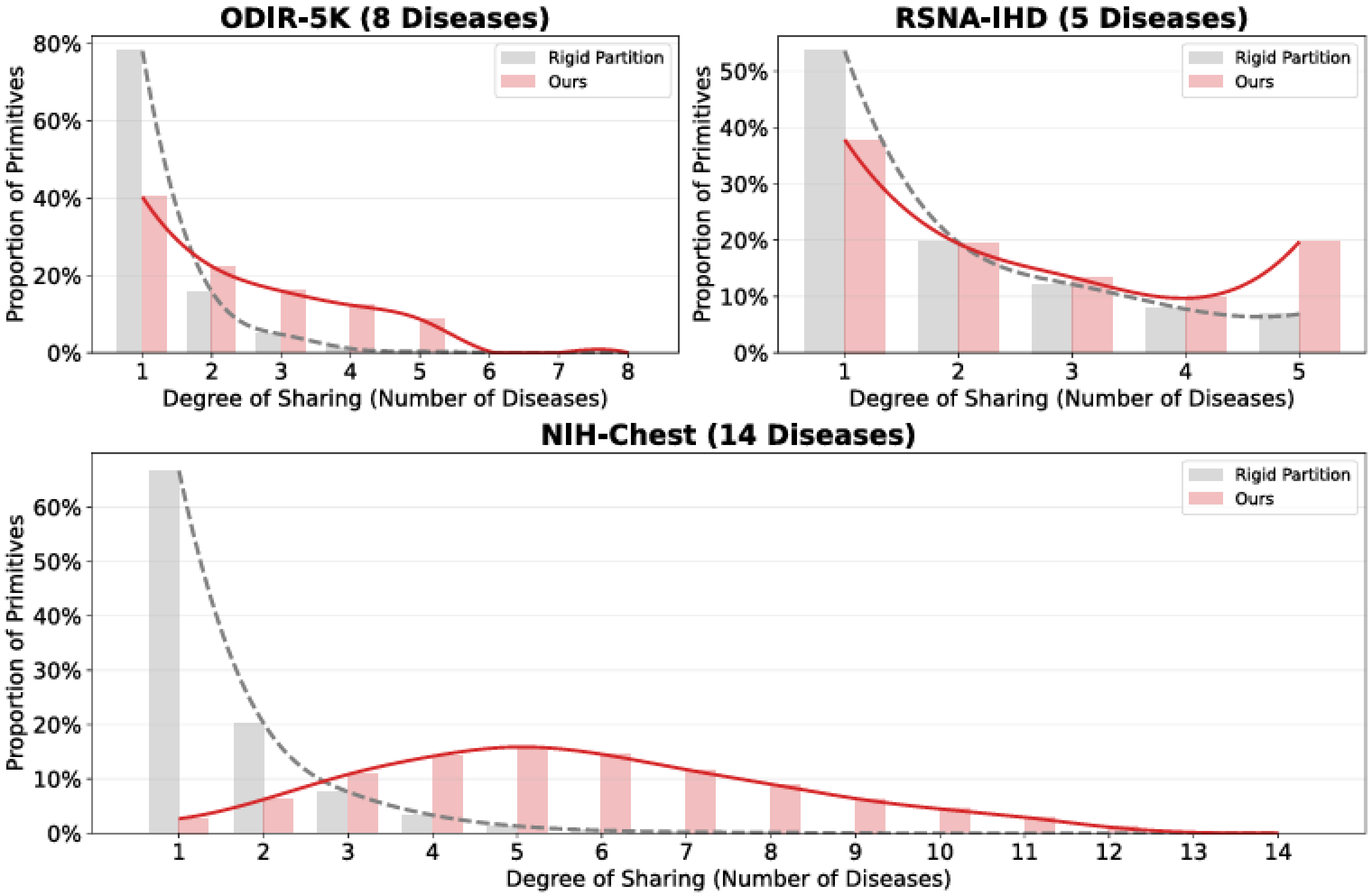
Effectiveness of polysemy modeling using Atomic Visual Space



Experiment

16-11 July · Seoul, South Korea

Effectiveness of polysemy modeling using Atomic Visual Space



Experiment

Effectiveness of Fuzzy membership using Gaussian kernel

Kernels	ODIR-5K				NIH-Chest				RSNA-IHD			
	mAP↑	F1↑	AUC↑	Avg.↑	mAP↑	F1↑	AUC↑	Avg.↑	mAP↑	F1↑	AUC↑	Avg.↑
Linear	69.94	89.64	91.82	83.80	51.94	89.85	87.23	76.34	89.15	92.16	96.64	92.65
Cosine	68.93	89.66	91.66	83.42	52.69	90.15	87.45	76.76	89.14	92.03	96.65	92.61
Laplacian	30.98	85.36	76.20	64.18	27.63	88.70	76.04	64.12	48.66	74.80	73.15	65.54
Polynomial	33.85	14.64	76.47	41.65	35.76	12.58	75.17	41.17	86.97	28.46	91.22	68.88
Ours	**72.61**	**90.91**	**92.97**	**85.50**	**58.95**	**90.44**	**87.81**	**79.07**	**89.72**	**92.97**	**96.90**	**93.20**

Experiment

Ablation Study
① Atomic Visual Space
② Comorbidity Topology

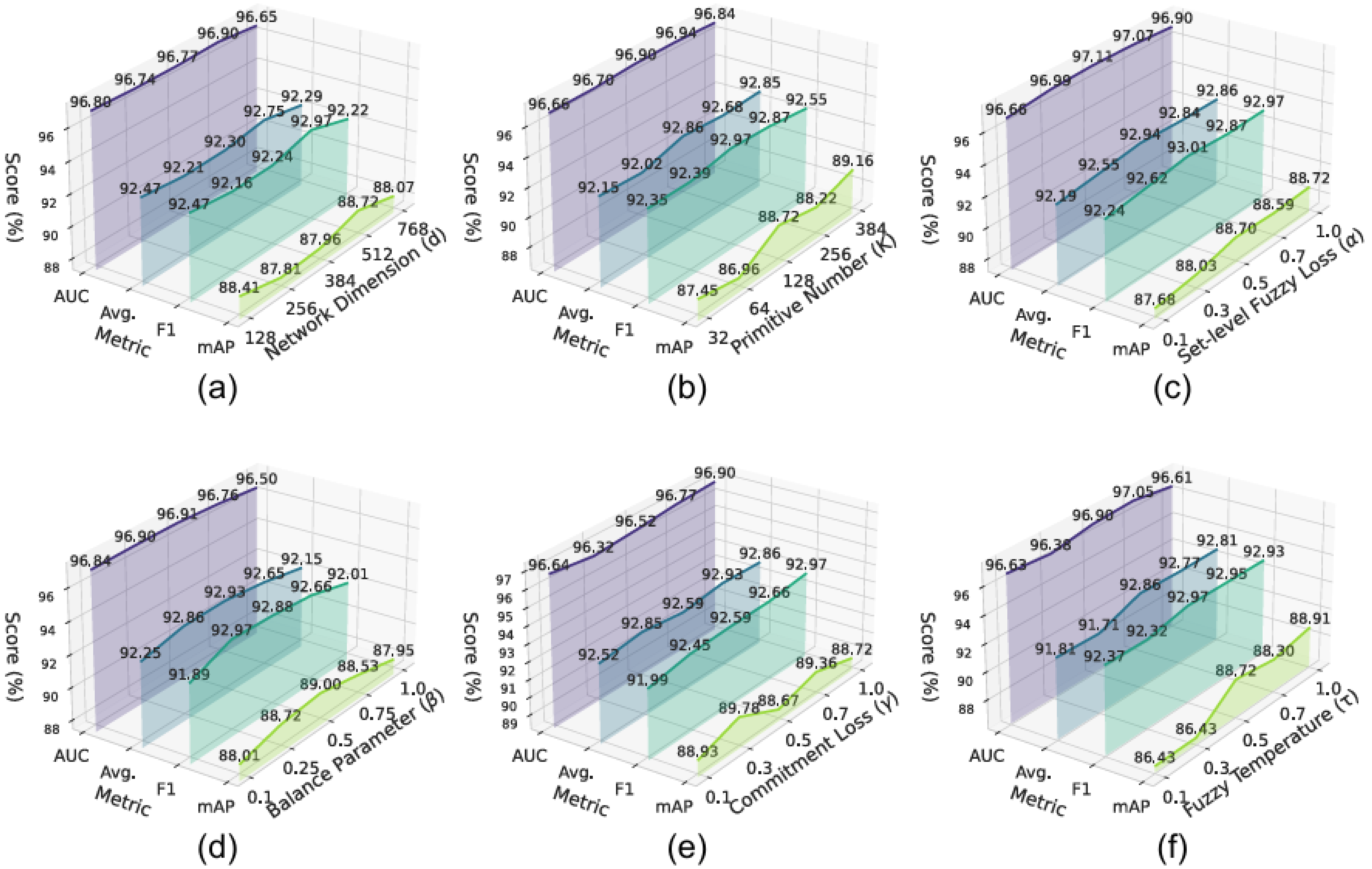
No.	①	②	mAP↑	F1↑	AUC↑	Avg.↑
#1	✗	✗	60.73	88.38	88.79	79.30
#2	✓	✗	70.31	90.00	92.62	84.31
#3	✗	✓	69.12	90.70	92.42	84.08
#4	✓	✓	**72.61**	**90.91**	**92.97**	**85.50**

Robustness on tong-tailed dataset

No.	Topo.		Objec.		Groups		
	w	w/o	$\mathcal{L}_{\text{fuzzy}}$	$\mathcal{L}_{\text{BCE}}$	Head	Middle	Tail
#1	✓	✗	✗	✓	**74.14**	71.13	67.16
#2	✗	✓	✓	✗	72.59	68.29	72.87
#3	✓	✗	✓	✗	73.83	**71.92**	**74.30**

Experiment

16-11 July · Seoul, South Korea



ICML
International Conference
On Machine Learning

16-11 July · Seoul, South Korea

Thanks For Your Attention!

Open Source

Please feel free to contact us!

cyy262511@gmail.com
yzhang@scu.edu.cn

For more experimental results, please follow our articles
The resources will be released in this GitHub repository