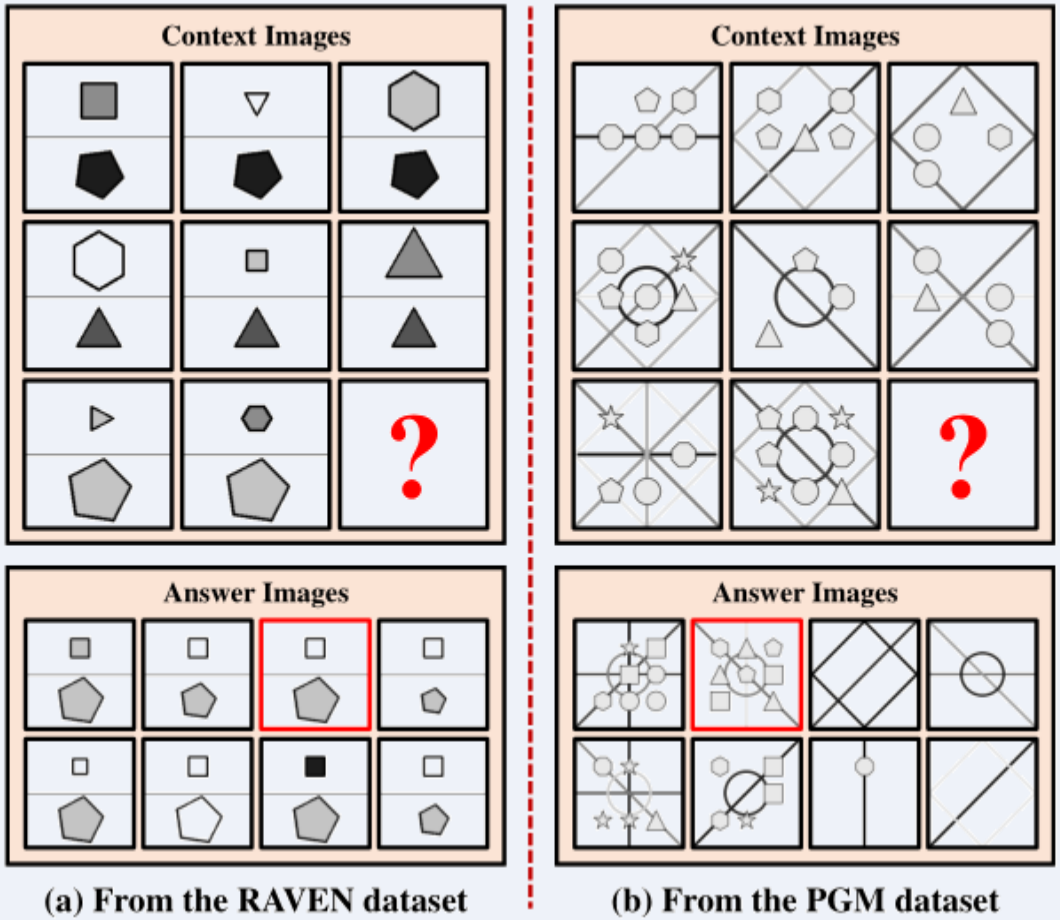
Neural Prediction Errors enable Analogical Visual Reasoning in Human Standard Intelligence Tests

Lingxiao Yang, Hongzhi You, Zonglei Zhen, Da-Hui Wang
Xiaohong Wan, Xiaohua Xie, Ru-Yuan Zhang

TOC

1 Raven's Progressive Matrix (RPM)

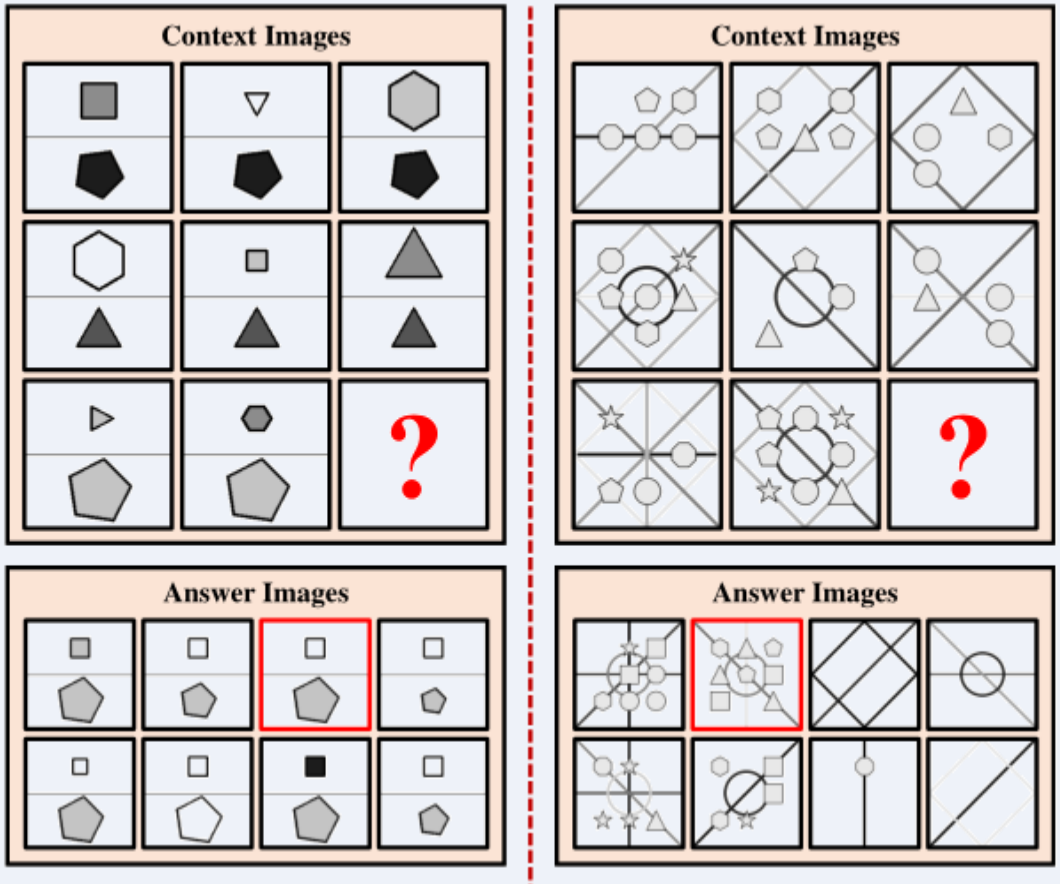
Task: Given eight context images. Choosing the correct one (highlighted in red) from eight answer images to fill in the missing one (denoted by ?), making three rows or three columns with similar patterns.



(a) From the RAVEN dataset (b) From the PGM dataset

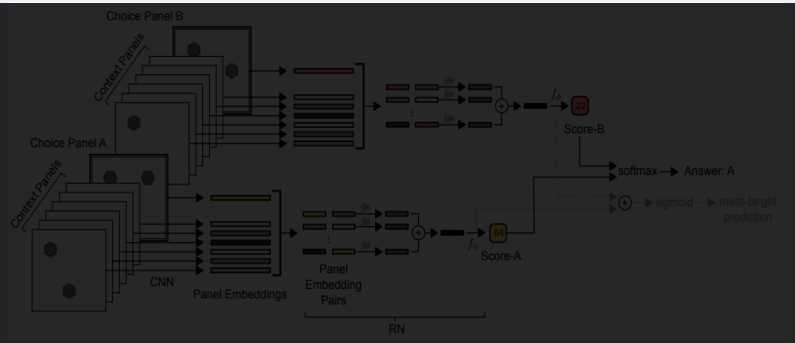
1 Raven's Progressive Matrix (RPM)

Problem: Not only **recognize objects** within each image, but also need to discover different **relationships between objects,** and then infer the right image for each RPM question.



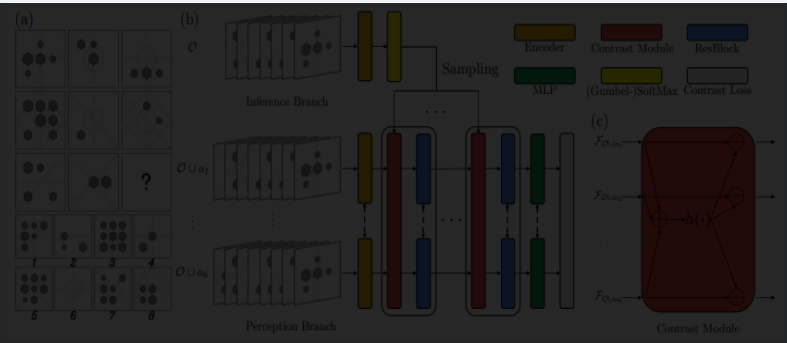
(a) From the RAVEN dataset (b) From the PGM dataset

2 Related Works – 4 representative works



WReN: Pairwise Relationship

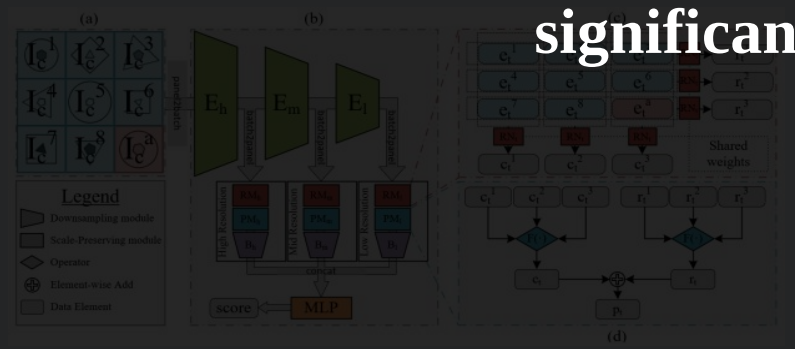
[David G.T. Barrett et al, 2018, *International Conference on Machine Learning*]



CoPINet: Contrasting Answers

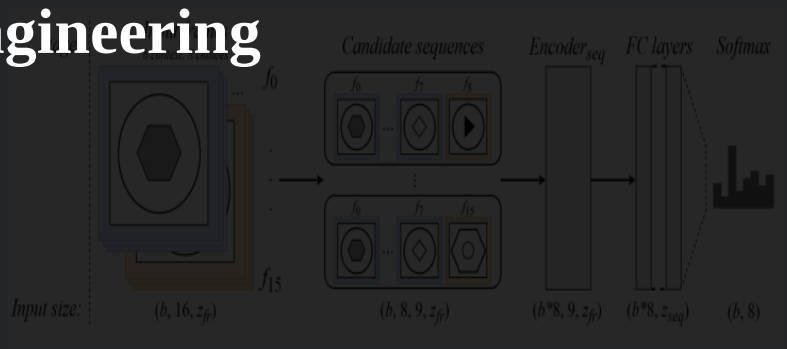
[Chi Zhang et al, 2019, *Neural Information Processing Systems (NeurIPS)*]

The choices of these network design require significant engineering



MRNet: Multi-Scale Relationship

[Yaniv Benny et al, 2021, *Computer Vision and Pattern Recognition Conference*]



Rel-Base: FC over images

[Steven Spratley et al, 2020, *European Conference on Computer Vision*]

2 Related Works – Analysis of existing works

Few studies show impressive performance on all datasets and different generalization cases

Method	WReN	LEN	CoPINet	SRAN	DCNet	MLRN	SCL	MXNet	Rel-Base	MRNet
PGM-N	62.6	68.1	56.4	71.3	68.6	98.0	88.9	66.7	85.5	94.5
RVN-O	16.8	72.9	91.4	54.3†	93.6	12.3†	91.6	83.9	91.7	**96.6**
RVN-F	30.3	51.0	50.6	72.9†	56.1†	29.5†	90.1†	35.1†	93.5†	88.4
I-RVN	23.8	41.4	46.1	60.8	47.2†	12.3†	95.0	26.8†	91.1†	83.5†
Avg	33.4	58.4	61.1	64.8	66.4	38.0	91.4	53.1	90.5	90.8

Similar relationship discovery method (different impl.)

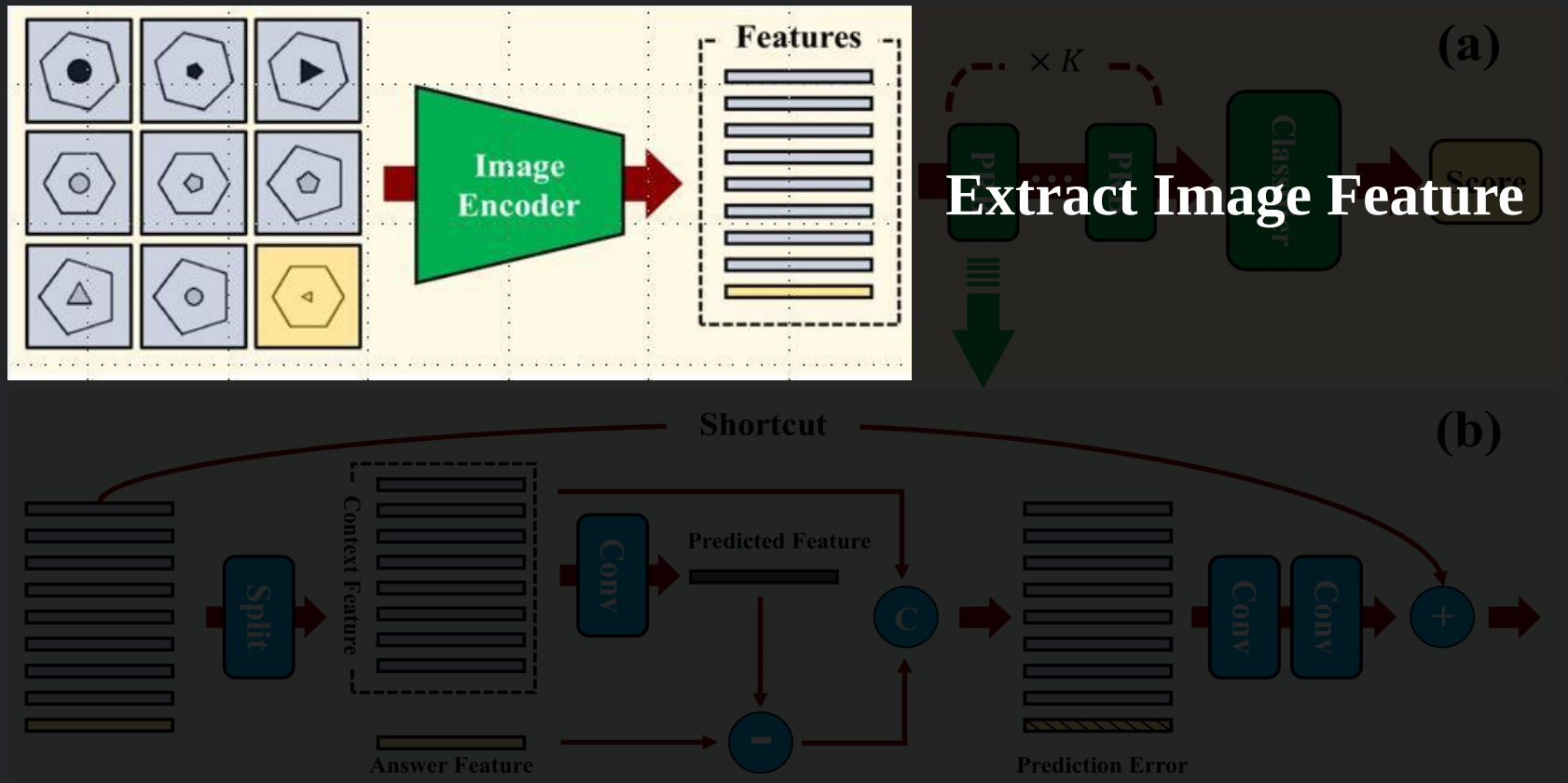
CoPINet -> DCNet (Answer contrasting)
Rel-Base -> SCL (FC over images)
MRNet -> MLRN (Multi-scale)
… …

3 Proposed Method – Motivations

- **Human brain can do well in inductive and transductive tasks. A lot of studies on brain have been proposed in last decades.**

- **Exploring a different method for discovering relationship that can generalize well across datasets and tasks.**

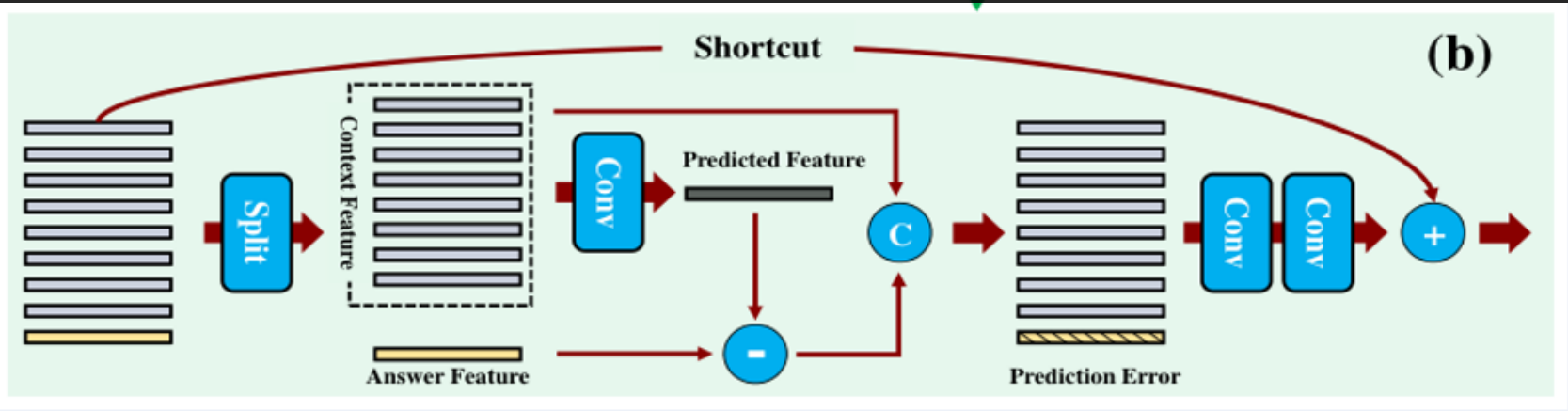
3 Proposed Method – PredRNet



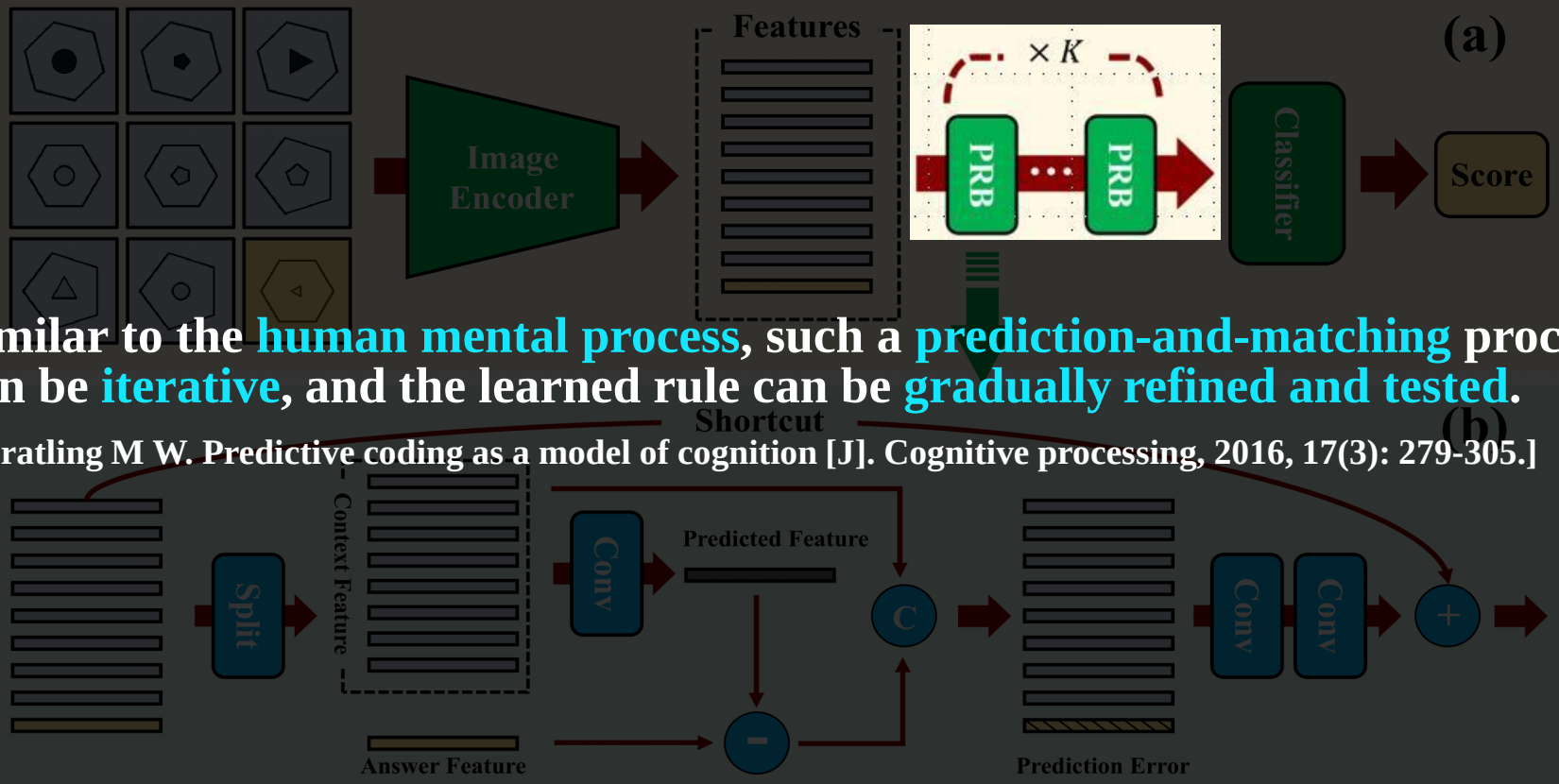
3 Proposed Method – PredRNet

Our brain constructs an internal model to approximate the operations of the external environment. This internal model generates predictions about what the observed sensory evidence should be, and the brain uses prediction errors to update the belief held in the internal model.

[Friston K, Kiebel S. Predictive coding under the free-energy principle[J]. Philosophical transactions of the Royal Society B: Biological Sciences, 2009, 364(1521): 1211-1221.]



3 Proposed Method – PredRNet



Similar to the human mental process, such a prediction-and-matching process can be iterative, and the learned rule can be gradually refined and tested.

[Spratling M W. Predictive coding as a model of cognition [J]. Cognitive processing, 2016, 17(3): 279-305.]

4 Experiments

Table 1: Recognition accuracy (%) on PGM Neutral (PGM-N), original RAVEN (RVN-O), RAVEN-FAIR (RVN-F), and Impartial-RAVEN (I-RVN). For all RAVENs, accuracy is obtained by averaging across all seven configurations. † indicates the performance was not reported in their original paper, and is obtained by running their published codes. The best and the second best results on each dataset are highlighted by **bold** and underline, respectively. Our PredRNet obtains the state-of-the-art average (Avg) performance on all the four compared datasets.

Method	WReN	LEN	CoPINet	SRAN	DCNet	MLRN	SCL	MXNet	Rel-Base	MRNet	STSN	PredRNet
PGM-N	62.6	68.1	56.4	71.3	68.6	98.0	88.9	66.7	85.5	94.5	**98.2**	97.4
RVN-O	16.8	72.9	91.4	54.3†	93.6	12.3†	91.6	83.9	91.7	**96.6**	89.7†	95.8
RVN-F	30.3	51.0	50.6	72.9†	56.1†	29.5†	90.1†	35.1†	93.5†	88.4	95.4†	**97.1**
I-RVN	23.8	41.4	46.1	60.8	47.2†	12.3†	95.0	26.8†	91.1†	83.5†	95.7	**96.5**
Avg	33.4	58.4	61.1	64.8	66.4	38.0	91.4	53.1	90.5	90.8	94.8	**96.7**

(a) Recognition accuracy (%) on all regimes of PGM (1 *Neutral* and 7 OOD subsets, Ntr: Neutral, Int: Interpolation, Ext: Extrapolation, H.O: Held-Out, P: Pairs, TP: TriplePairs, LT: LineType, SC: ShapeColor). The best and the second best results are highlighted using **bold** and underline. Our PredRNet obtains competitive results without using any extra supervision signals.

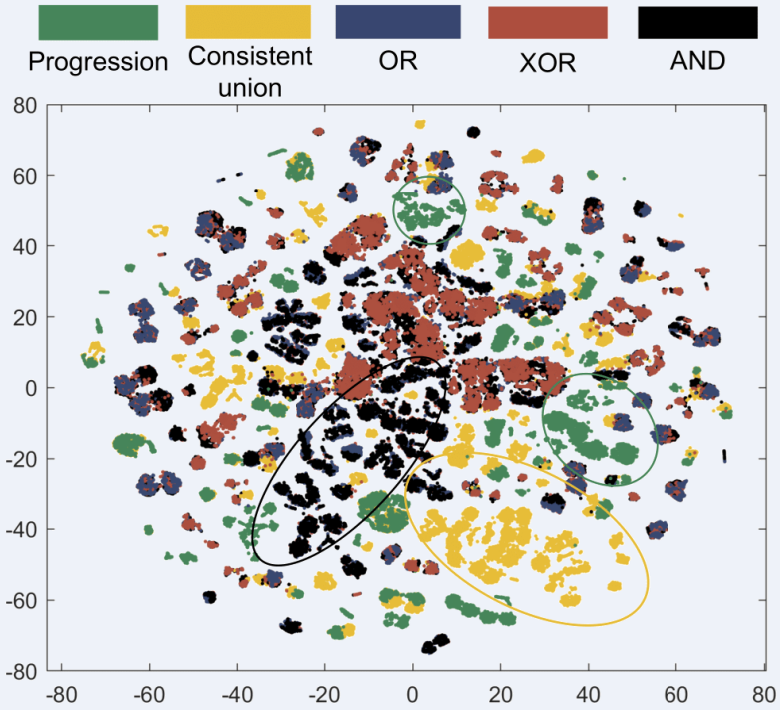
Method	Ntr	Int	Ext	H.O.P	H.O.TP	H.O.T	H.O.LT	H.O.SC	Avg
WReN	62.6	64.4	17.2	27.2	41.9	19.0	14.4	12.5	32.4
MXGNet	66.7	65.4	18.9	33.6	43.3	19.9	16.7	16.6	35.1
MRNet	93.4	68.1	19.2	38.4	55.3	**25.9**	**30.1**	**16.9**	43.4
PredRNet	**97.4**	**70.5**	**19.7**	**63.4**	**67.8**	23.4	27.3	13.1	**47.1**

(b) Recognition accuracy (%) on all configurations of CLEVR-Matrices (Mondal et al., 2022). The best and the second best results are highlighted using **bold** and underline. Our PredRNet obtains competitive results.

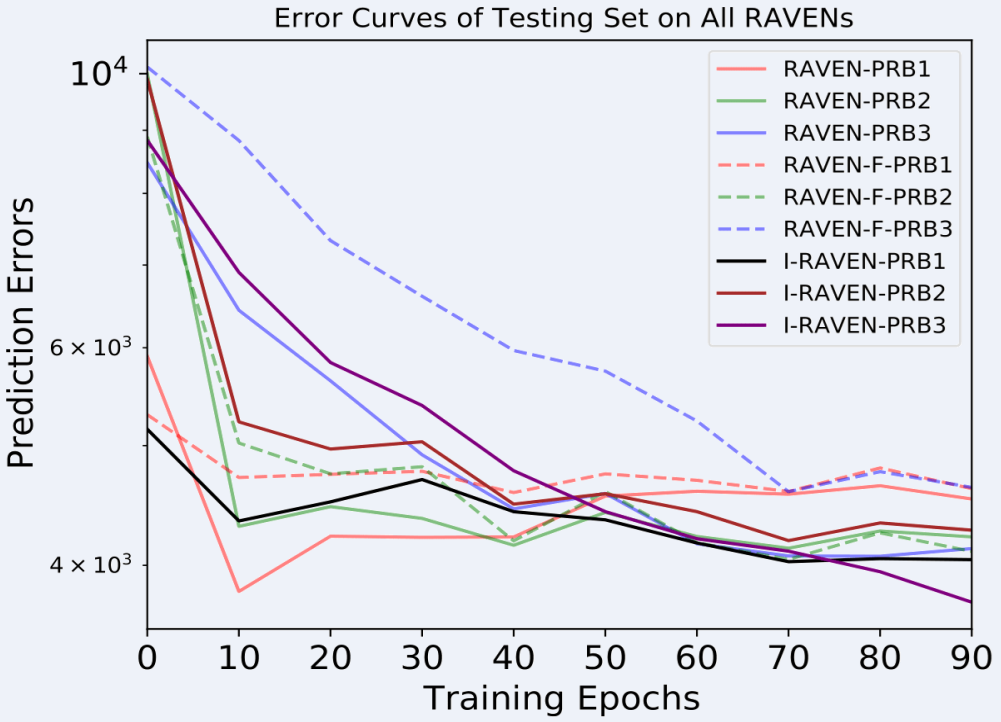
Method	Logic	Location	Count	Avg
MLRN	47.4	21.4	23.6	30.8
SCL	80.9	65.8	64.9	70.5
STSN	99.2	**100.0**	99.6	99.6
PredRNet	**100.0**	99.5	**99.9**	**99.8**

4 Experiments

t-SNE of prediction errors



Prediction errors during training



5 Conclusion

Highlights

- We, inspired from the well-known concept in the neuroscience, proposed a prediction error based reasoning network – PredRNet
- Our PredRNet obtains remarkable performance in a variety of datasets and tasks

Limitations

- May not fully explore the power of error computation
- Human reasoning process and our model are not strictly identical (semi-, un- supervised vs supervised learning)

THANKS

https://github.com/ZjjConan/AVR-PredRNet