Hyper-opinion Evidential Deep Learning for Out-of-Distribution Detection

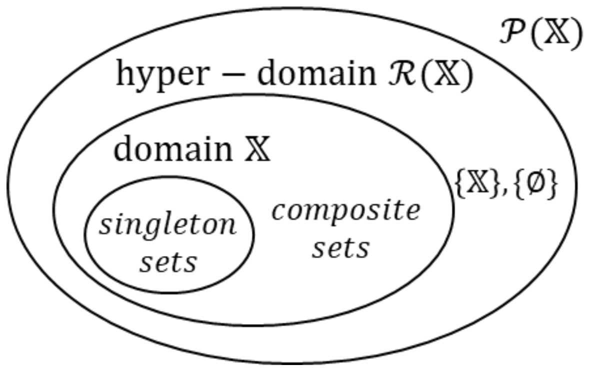
Jingen Qu · Yufei Chen · Xiaodong Yue · Wei Fu · Qiguang Huang

Hyper-opinion Evidential Learning(under review)

MedAI Lab
Tongji University

Subjective Logic

Subjective Logic (SL) is a theory of uncertain reasoning based on probability theory and belief theory in a **domain** $\mathbb{X}$, which represents the set of exclusive possible states of a variable situation, such as class labels. It introduces the concepts of belief mass and uncertainty mass to describe the degree of belief and uncertainty about an event.



For example:

$$u + \sum_{k=1}^{K} b_k = 1$$

And the projected probability distribution derived from the opinion in SL corresponds to the expected probability distribution derived from a Dirichlet distribution as:

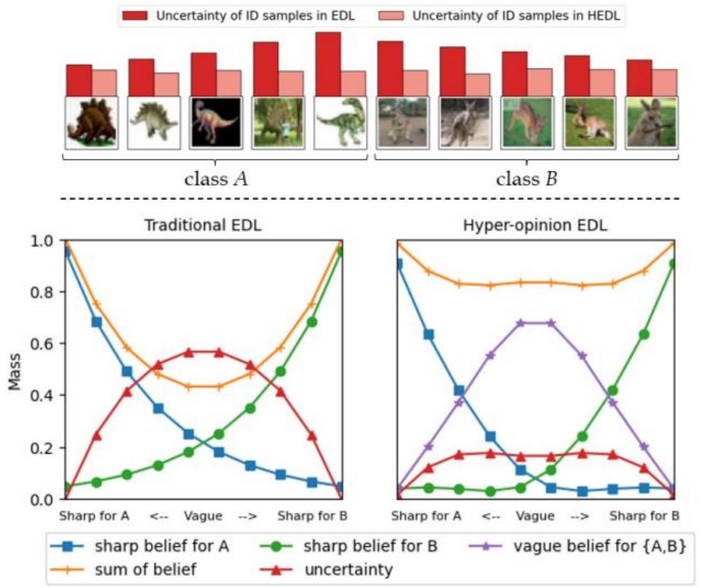
$$\boldsymbol{\omega} = (\boldsymbol{b}, u, \boldsymbol{a}) \leftrightarrow Dir(\boldsymbol{P} \,|\boldsymbol{\alpha})$$

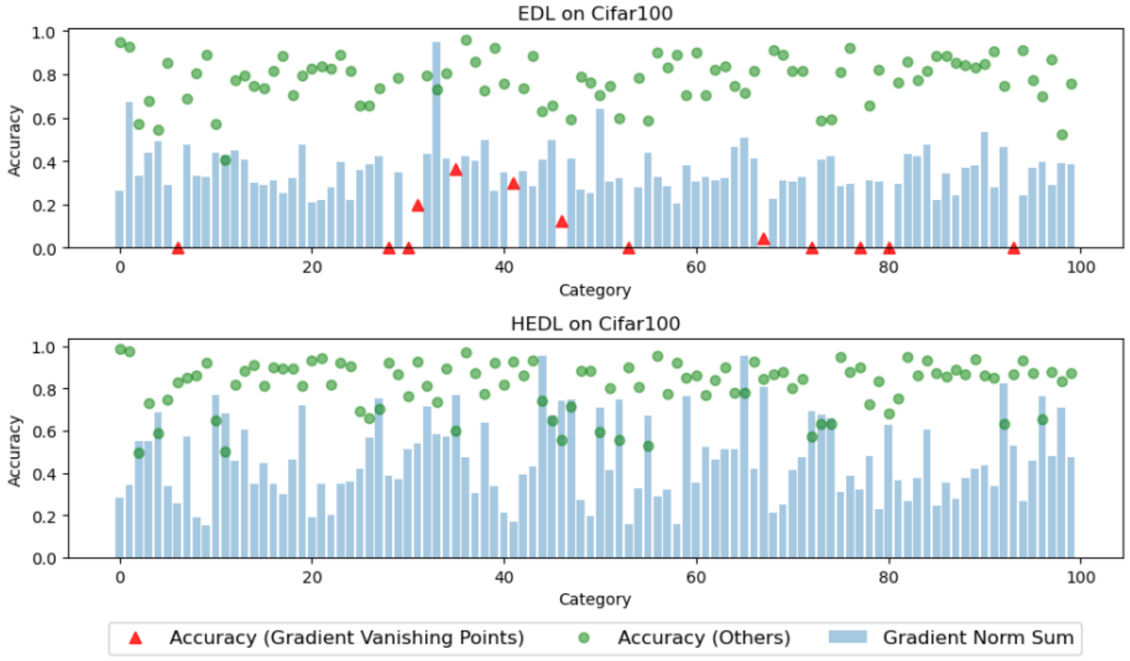
And α can be calculated through the evidence outputted by neural network. Therefore the neural network becomes uncertainty-aware.

Hyper-opinion Evidential Learning

- Existing methods apply multinomial-opinion in SL, which only contains the belief mass for singletons and ignore the belief mass for composite sets that contain multiple singletons.

- **We apply hyper-opinion in SL. It allows us to consider both belief mass assigned to singletons and composite sets.**

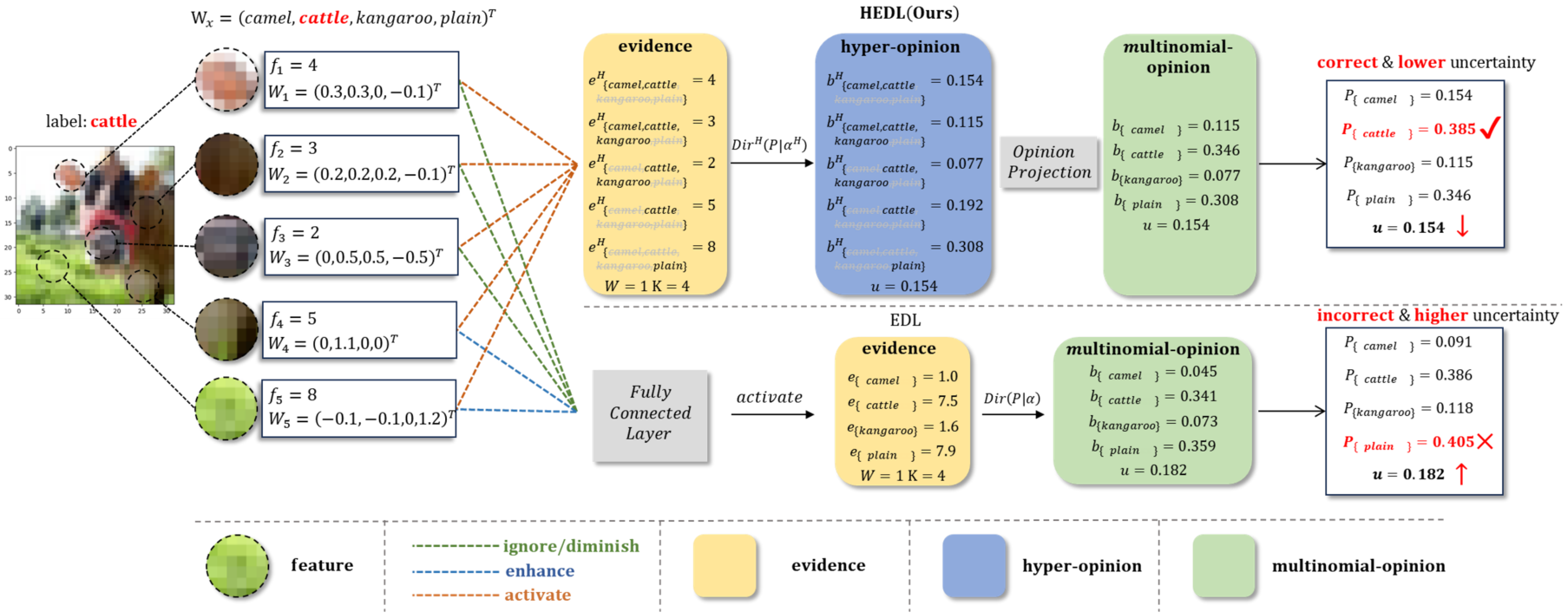




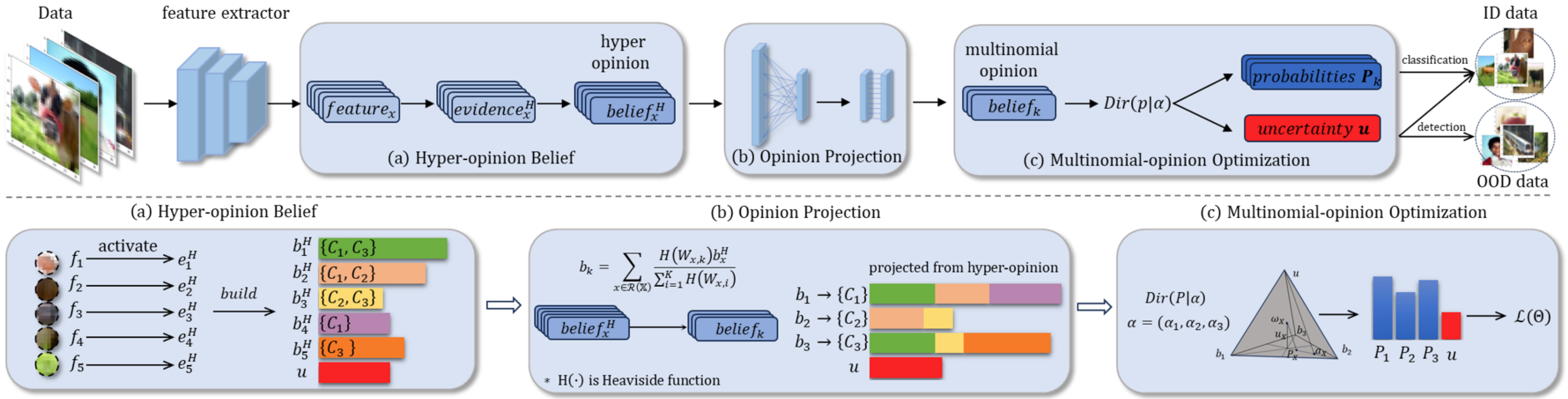
- In addition, the parameters of fully-connected layer in EDL models are facing vanishing gradient problem when number of category in datasets rises

- **Hyper-opinion Evidential Deep Learning (HEDL) projects hyper-opinion to multinomial-opinion, mitigating the vanishing gradient problem, while preserving computational efficiency.**

Hyper-opinion Evidential Learning



Hyper-opinion Evidential Learning



Our method models the evidence in hyper-domain $\mathcal{R}(\mathbb{X})$ with hyper-opinion, which provides a belief mass b_x^H, $x \in \mathcal{R}(\mathbb{X})$ representing the belief degree of set x. Along with $\boldsymbol{a}^H$ and u, the three compose a hyper-opinion:

$$b^H : \mathcal{R}(\mathbb{X}) \rightarrow [0, 1]$$

$$u + \sum_{x \in \mathcal{R}(\mathbb{X})} b_x^H = 1.$$

$$\omega^H = \left(\boldsymbol{b}^H, u, \boldsymbol{a}^H\right) \leftrightarrow Dir^H\left(\boldsymbol{P} \,|\, \boldsymbol{\alpha}^H\right)$$

We build Dirichlet hyper distribution on the features the neural network extracted.

Hyper-opinion Evidential Learning

MedAI Lab
Tongji University

CIFAR-10

Method	SVHN FPR95↓	SVHN AUPR↑	SVHN AUROC↑	Textures FPR95↓	Textures AUPR↑	Textures AUROC↑	Place365 FPR95↓	Place365 AUPR↑	Place365 AUROC↑	Average FPR95↓	Average AUPR↑	Average AUROC↑	ID data Acc.↑
MSP[16]	51.87	78.19	90.88	59.89	91.28	88.72	57.64	70.24	89.03	56.47	79.90	89.54	95.06
ODIN[29]	67.92	42.13	73.32	51.10	82.25	80.70	50.51	50.27	82.55	56.51	58.22	78.86	95.06
openGAN[24]	99.39	33.90	53.56	98.24	61.48	42.22	99.44	19.55	36.58	99.02	38.31	44.12	95.06
GradNorm[21]	91.65	78.89	53.91	98.09	48.05	52.07	92.46	86.63	60.50	94.07	71.19	55.49	95.06
VIM[60]	14.41	93.76	97.22	20.78	97.36	96.06	47.52	72.83	90.08	27.57	87.98	94.46	95.06
KNN[55]	33.32	92.31	95.13	46.01	95.93	92.77	43.78	80.15	91.82	41.04	89.47	93.23	95.06
DICE[53]	67.78	73.19	86.43	67.48	85.38	80.14	56.06	57.52	84.43	63.78	72.03	83.66	95.06
RankFeat[52]	64.49	80.33	68.15	59.71	55.39	73.46	43.70	94.66	85.99	55.97	76.79	75.87	95.06
ASH[8]	83.64	89.06	73.46	84.59	72.85	77.45	77.89	94.04	79.89	82.04	85.32	76.93	95.06
SHE[64]	62.74	94.46	86.38	84.60	77.28	81.57	76.36	94.88	82.89	74.57	88.87	83.61	95.06
GEN[32]	28.14	96.37	91.97	40.74	84.71	90.14	47.03	96.67	89.46	38.64	92.58	90.52	95.06
MCDropout[12]	44.58	85.03	92.67	56.60	91.74	88.83	56.20	67.20	88.43	52.47	81.32	89.98	94.95
G-ODIN[19]	8.42	96.63	98.41	23.32	96.03	94.51	39.80	75.49	91.10	23.84	89.39	94.67	94.70
CSI[56]	17.56	**97.75**	95.18	28.95	82.99	90.71	34.76	**96.38**	89.56	27.09	92.37	91.82	91.16
MOS[20]	90.85	70.55	51.09	85.56	90.89	52.91	71.74	78.67	74.15	82.71	80.03	59.38	94.83
VOS[9]	29.92	83.73	93.82	37.38	92.72	91.26	45.37	63.93	88.73	37.55	80.13	91.27	**95.82**
LogitNorm[61]	**5.30**	97.70	**98.86**	30.94	96.32	94.30	31.17	88.11	94.76	22.47	94.04	95.97	94.30
EDL[48]	11.56	88.60	93.92	19.95	99.07	95.70	19.36	93.15	**96.54**	16.96	93.61	95.39	95.72
RED[43]	65.75	29.85	61.30	86.49	71.56	28.06	72.37	19.83	51.16	74.87	40.41	46.84	95.80
HEDL(Ours)	8.43	94.09	96.86	**19.15**	**99.19**	**96.23**	**19.08**	90.14	95.71	**15.55**	**94.47**	**96.27**	95.66

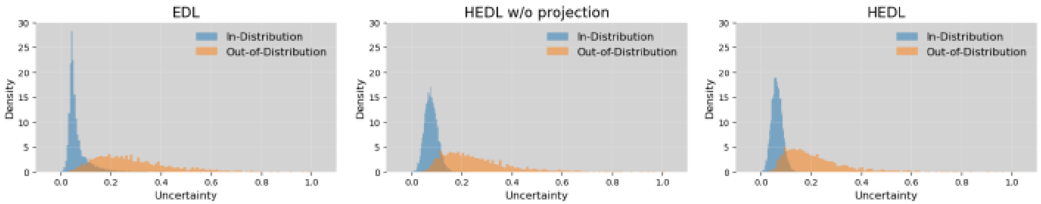
CIFAR-100

Method	SVHN FPR95↓	SVHN AUPR↑	SVHN AUROC↑	Textures FPR95↓	Textures AUPR↑	Textures AUROC↑	Place365 FPR95↓	Place365 AUPR↑	Place365 AUROC↑	Average FPR95↓	Average AUPR↑	Average AUROC↑	ID data Acc.↑
MSP[16]	83.69	60.76	76.04	83.83	85.24	76.93	81.24	62.39	79.44	82.91	69.46	77.47	77.25
ODIN[29]	89.76	52.36	71.08	78.37	86.67	79.39	81.27	60.85	79.83	83.13	66.62	76.77	77.25
openGAN[24]	83.96	60.85	78.68	86.31	80.18	73.53	88.37	38.87	70.15	86.21	59.96	74.12	77.25
GradNorm[21]	69.90	89.45	76.95	92.51	56.77	64.58	95.32	88.78	69.69	85.91	78.33	70.41	77.25
VIM[60]	82.79	72.82	81.20	**55.90**	92.15	87.41	83.85	56.24	75.76	74.18	73.74	81.46	77.25
KNN[55]	74.27	71.46	82.21	66.40	89.44	83.81	78.74	57.47	79.10	73.13	72.79	81.71	77.25
DICE[53]	79.93	65.95	79.97	80.53	85.41	77.70	80.75	62.76	80.18	80.40	71.37	79.28	77.25
RankFeat[52]	58.49	83.40	72.14	66.87	52.42	69.40	77.42	83.74	63.82	67.59	73.19	68.45	77.25
ASH[8]	46.00	**92.97**	85.60	61.27	68.97	80.72	62.95	91.48	78.76	56.74	84.47	81.69	77.25
SHE[64]	59.15	90.85	80.97	73.29	60.87	73.64	65.24	90.31	76.30	65.89	80.68	76.97	77.25
GEN[32]	55.45	90.36	81.41	61.23	64.52	78.74	**56.25**	**91.90**	80.28	57.64	82.26	80.14	77.25
MCDropout[12]	71.63	67.44	81.31	80.16	86.01	77.93	79.52	61.34	79.20	77.11	71.60	79.48	75.83
G-ODIN[19]	71.62	79.80	86.13	58.01	93.01	**88.35**	78.67	55.45	78.15	69.44	76.09	84.21	74.46
CSI[56]	67.21	91.76	80.24	90.51	51.46	62.22	69.41	88.16	70.99	75.71	77.13	71.15	61.60
MOS[20]	90.58	74.48	59.42	96.32	89.60	46.69	92.64	71.87	60.95	93.18	78.64	55.69	76.98
VOS[9]	98.62	56.36	68.99	94.54	76.20	68.33	97.81	43.20	68.21	96.99	58.59	68.51	77.20
LogitNorm[61]	79.16	75.57	83.03	87.06	79.08	71.53	80.20	63.10	79.84	82.14	72.58	78.13	76.34
EDL[48]	93.05	75.48	81.39	95.48	93.80	71.60	99.30	68.57	76.55	95.94	79.28	76.51	71.40
RED[43]	90.09	62.75	76.41	56.01	96.25	85.29	68.11	64.75	84.46	71.40	74.58	82.05	80.36
HEDL(Ours)	**39.56**	89.22	**93.46**	61.97	**96.85**	85.98	63.89	81.14	**89.32**	**55.14**	**89.07**	**89.59**	**80.40**

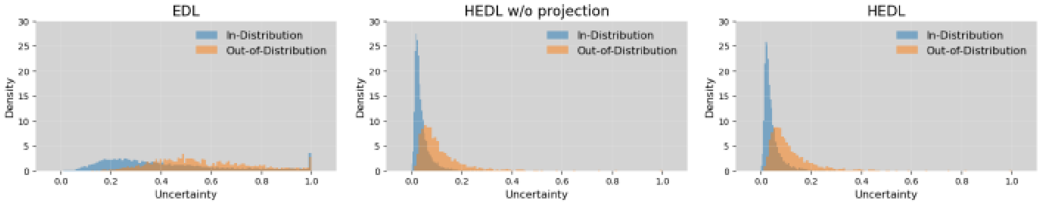
Multinomial-opinion	Hyper-opinion	Opinion-projection	Flower-102 Average OOD performance FPR95↓	AUPR↑	AUROC↑	Flower-102 ID data Acc.↑	CUB-200-2011 Average OOD performance FPR95↓	AUPR↑	AUROC↑	CUB-200-2011 ID data Acc.↑
-	-	-	14.86	95.94	97.42	83.75	30.29	91.18	94.35	**75.82**
✓	-	-	100.00	66.95	67.23	66.84	98.03	71.80	75.27	59.87
✓	✓	-	11.90	95.83	97.61	81.40	9.32	91.57	97.82	52.30
✓	✓	✓	**3.98**	**98.73**	**99.07**	**84.13**	**3.82**	**97.80**	**98.91**	74.62

Hyper-opinion Evidential Learning

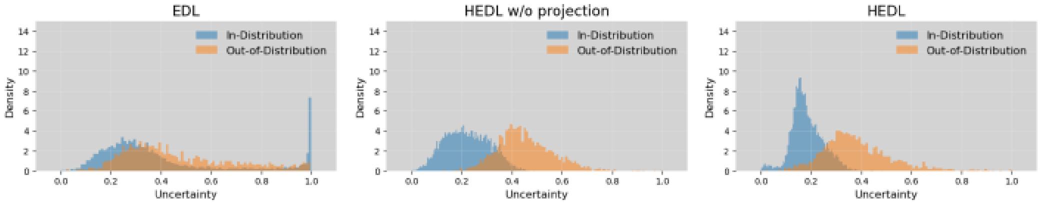
Method	Flower-102				CUB-200-2011			
	Average OOD performance			ID data	Average OOD performance			ID data
	FPR95↓	AUPR↑	AUROC↑	Acc.↑	FPR95↓	AUPR↑	AUROC↑	Acc.↑
MSP[16]	14.86	95.94	97.42	83.75	30.29	91.18	94.35	75.82
ODIN[29]	4.36	97.63	98.22	83.75	21.92	89.92	96.22	75.82
VIM[60]	6.34	96.70	97.94	83.75	6.71	97.27	98.26	75.82
GradNorm[21]	5.38	97.11	98.81	83.75	32.08	97.68	95.22	75.82
KNN[55]	18.45	88.83	95.30	83.75	14.35	88.63	97.40	75.82
DICE[53]	4.64	97.62	98.95	83.75	25.82	88.83	96.00	75.82
RankFeat[52]	96.57	76.62	60.98	83.75	74.68	83.38	71.09	75.82
ASH[8]	5.16	97.54	98.84	83.75	15.82	92.75	97.07	75.82
SHE[64]	11.69	93.96	97.79	83.75	22.94	96.14	96.18	75.82
GEN[32]	5.25	97.55	98.85	83.75	15.88	92.74	97.06	75.82
MCDropout[12]	14.77	96.22	97.41	83.98	42.46	87.08	91.76	75.83
G-ODIN[19]	56.92	69.88	82.12	24.30	29.51	85.13	93.85	66.74
VOS[9]	39.17	84.52	90.11	78.08	35.98	83.93	89.86	75.92
LogitNorm[61]	41.07	80.34	85.65	77.41	22.69	91.69	95.99	74.84
EDL[48]	100.00	66.95	67.23	66.84	98.03	71.80	75.27	59.87
RED[43]	95.87	80.10	76.45	**84.63**	36.01	94.58	94.89	**76.30**
HEDL(Ours)	**3.98**	**98.73**	**99.07**	84.13	**3.82**	**97.80**	**98.91**	74.62



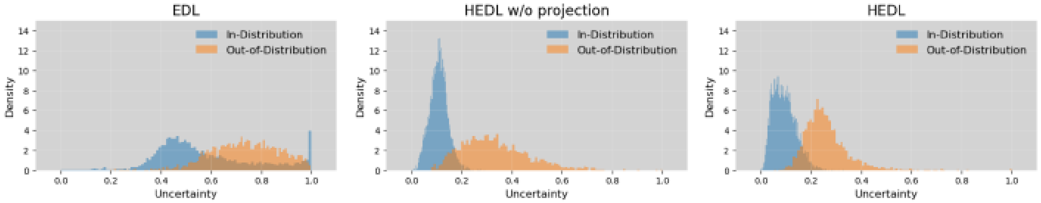
(a) CIFAR-10, the overlap between ID and OOD is 20%, 23%, and 18% for EDL, HEDL w/o projection, and HEDL, respectively.



(b) CIFAR-100, the overlap between ID and OOD is 62%, 45%, and 41% for EDL, HEDL w/o projection, and HEDL, respectively.



(c) Flower-102, the overlap between ID and OOD is 71%, 26%, and 29% for EDL, HEDL w/o projection, and HEDL, respectively.



(d) CUB-200-2011, the overlap between ID and OOD is 50%, 20%, and 17% for EDL, HEDL w/o projection, and HEDL, respectively.

Thanks!