Supplementary Materials: Dynamic Evidence Decoupling for Trusted Multi-view Learning

Anonymous Authors

We show the proof and visualization of the separation degree increase, more experiments about the semantic vagueness phenomenon, and hyper-parameter settings in this appendix.

A PROOF

Proposition A.1. *For a belief mass $\boldsymbol{b} = (b_1, b_2, \ldots, b_C)$ which has C classes, and each belief mass is non-negative. After the process of the separation degree increase operation, we obtain the result $\widetilde{\boldsymbol{b}}$, then the separation degree of $\widetilde{\boldsymbol{b}}$ is always bigger than or equal to $\boldsymbol{b}$:*

(1) $\sum_{c=1}^{C} \widetilde{b_c^{con}} = 0$,

$$SD(\widetilde{\boldsymbol{b}}) = SD(\boldsymbol{b}) = 0.$$

(2) $\sum_{c=1}^{C} \widetilde{b_c^{con}} > 0$,

$$\begin{aligned}
SD(\widetilde{\boldsymbol{b}}) &= \sum_{i=1}^{C} \sum_{i \neq j}^{C} \left| \frac{\sum_{c=1}^{C} b_c}{\sum_{c=1}^{C} \widetilde{b_c}} \widetilde{b_i} - \frac{\sum_{c=1}^{C} b_c}{\sum_{c=1}^{C} \widetilde{b_c}} \widetilde{b_j} \right| \\
&= \frac{\sum_{c=1}^{C} b_c}{\sum_{c=1}^{C} \widetilde{b_c}} \sum_{i=1}^{C} \sum_{i \neq j}^{C} |\widetilde{b_i} - \widetilde{b_j}| \\
&\geq \sum_{i=1}^{C} \sum_{i \neq j}^{C} |b_i - b_j| = SD(\boldsymbol{b}).
\end{aligned}$$

Proof. We will prove the above preposition in the following two cases:

(1) $\sum_{c=1}^{C} \widetilde{b_c^{con}} = 0$:

When all belief mass is equal to 0, we have

$$SD(\widetilde{\boldsymbol{b}}) = SD(\boldsymbol{b}) = 0.$$

(2) $\sum_{c=1}^{C} \widetilde{b_c^{con}} > 0$:

Swapping b_i and b_j at any two positions in $\boldsymbol{b}$ does not change the value of $SD(\boldsymbol{b})$, then we can assume that $b_1 \leq b_2 \leq \ldots \leq b_C$, therefore the proposition can be equivalently expressed as follows:

$$\begin{aligned}
& SD(\widetilde{\boldsymbol{b}}) \geq SD(\boldsymbol{b}) \\
\Leftrightarrow & \frac{\sum_{c=1}^{C} b_c}{\sum_{c=1}^{C} \widetilde{b_c}} \sum_{i=1}^{C} \sum_{i \neq j}^{C} |\widetilde{b_i} - \widetilde{b_j}| \geq \sum_{i=1}^{C} \sum_{i \neq j}^{C} |b_i - b_j| \\
\Leftrightarrow & \frac{\sum_{i=1}^{C} \sum_{i \neq j}^{C} |\widetilde{b_i} - \widetilde{b_j}|}{\sum_{c=1}^{C} \widetilde{b_c}} \geq \frac{\sum_{i=1}^{C} \sum_{i \neq j}^{C} |b_i - b_j|}{\sum_{c=1}^{C} b_c}.
\end{aligned}$$

We then convert this to the following formula and eliminate the absolute value symbol:

$$\frac{\sum_{1 \leq i \leq j \leq C} (\widetilde{b_j} - \widetilde{b_i})}{\sum_{i=1}^{C} \widetilde{b_i}} \geq \frac{\sum_{1 \leq i \leq j \leq C} (b_j - b_i)}{\sum_{i=1}^{C} b_i},$$

we also eliminate j for each term and keep only i:

$$\begin{aligned}
& \frac{\sum_{i=1}^{C} (2i - C - 1) \widetilde{b_i}}{\sum_{i=1}^{C} \widetilde{b_i}} \geq \frac{\sum_{i=1}^{C} (2i - C - 1) b_i}{\sum_{i=1}^{C} b_i} \\
\Leftrightarrow & -(C-1) \frac{\sum_{i=1}^{C} 2i\widetilde{b_i}}{\sum_{i=1}^{C} \widetilde{b_i}} \geq -(C-1) \frac{\sum_{i=1}^{C} 2ib_i}{\sum_{i=1}^{C} b_i} \\
\Leftrightarrow & \frac{\sum_{i=1}^{C} 2i\widetilde{b_i}}{\sum_{i=1}^{C} \widetilde{b_i}} \geq \frac{\sum_{i=1}^{C} 2ib_i}{\sum_{i=1}^{C} b_i}.
\end{aligned}$$

As $\sum_{i=1}^{C} \widetilde{\boldsymbol{b}_i^{con}} > 0$ and $\sum_{i=1}^{C} \boldsymbol{b}_i^{con} > 0$ accordingly, we can move the denominators on both sides of the formula to the other side separately:

$$\begin{aligned}
& \sum_{i=1}^{C} 2i\widetilde{b_i} \sum_{j=1}^{C} b_j \geq \sum_{i=1}^{C} 2ib_i \sum_{j=1}^{C} \widetilde{b_j} \\
\Leftrightarrow & \sum_{i=1}^{C} \sum_{j=1}^{C} 2i\widetilde{b_i} b_j \geq \sum_{i=1}^{C} \sum_{j=1}^{C} 2ib_i \widetilde{b_j} \\
\Leftrightarrow & \sum_{i=1}^{C} \sum_{j=1}^{C} i\widetilde{b_i} b_j \geq \sum_{i=1}^{C} \sum_{j=1}^{C} ib_i \widetilde{b_j}.
\end{aligned}$$

For the above formula, we can swap the sign i and j. Then we obtain the following formula:

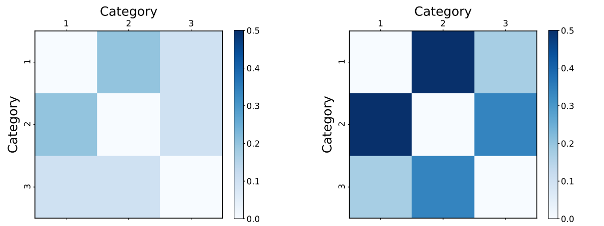
$$\begin{aligned}
& \sum_{i=1}^{C} \sum_{j=1}^{C} i\widetilde{b_i} b_j \geq \sum_{i=1}^{C} \sum_{j=1}^{C} j\widetilde{b_i} b_j \\
\Leftrightarrow & \sum_{i=1}^{C} \sum_{j=1}^{C} (i - j)\widetilde{b_i} b_j \geq 0.
\end{aligned}$$

For this formula, with the different relation of i and j, we can divide the left side into the sum of the two formulas:

$$\begin{aligned}
& \sum_{1 \leq i < j \leq C}^{C} (i-j)\widetilde{b_i} b_j + \sum_{1 \leq j < i \leq C}^{C} (i-j)\widetilde{b_i} b_j \geq 0 \\
\Leftrightarrow & \sum_{1 \leq i < j \leq C}^{C} (i-j)\widetilde{b_i} b_j + \sum_{1 \leq i < j \leq C}^{C} (j-i)\widetilde{b_j} b_i \geq 0 \\
\Leftrightarrow & \sum_{1 \leq i < j \leq C}^{C} (i-j)(\widetilde{b_i} b_j - b_i \widetilde{b_j}) \geq 0,
\end{aligned}$$

and the $\widetilde{b_i} = pow(b_i, \beta)$, where $\beta > 1$. Then when $1 \leq i < j \leq C, b_i \leq b_j$, we can obtain the following result:

$$\begin{aligned}
\widetilde{b_i} b_j - b_i \widetilde{b_j} &= (b_i)^{\beta} b_j - b_i (b_j)^{\beta} \\
&= ((b_i)^{\beta-1} - (b_j)^{\beta-1}) b_i b_j \leq 0
\end{aligned}$$



(a) Before (Separation Degree=0.80) **(b) After (Separation Degree=2.14)**

Figure 1: The visualization of separation degree before and after the separation degree increases operation.

Figure 2: Category examples of Colored-MNIST dataset

and we can obtain that

$$\sum_{1 \le i < j \le C}^{C} (i - j)(\widetilde{b}_i b_j - b_i \widetilde{b}_j) \ge 0,$$

this formula is equivalent to the proposition

$$SD(\widetilde{\boldsymbol{b}}) \ge SD(\boldsymbol{b}).$$

□

B SEPARATION DEGREE VISUALIZATION

To further understand the separation degree increase stage, we visualize the separation degree on the belief mass $\boldsymbol{b} = (0.2, 0.4, 0.3)$ which has three categories, and the parameter β is set to 3. As shown in figure 1, The left and right parts show the separation degree before and after the separation degree increase operation, respectively. The separation degree is increased from 0.80 to 2.04.

C SUPPLEMENTARY EXPERIMENTAL CONTENT

C.1 Introduction of Colored-MNIST Dataset

Colored-MNIST is a dataset of handwritten digits with RGB colored Backgrounds consisting of 3 colors (red, green, blue) for each number (‘0’ to ‘9’) with a total of 30 categories. We extract two types of features RGB and HOG which can distinguish the shape of digits and the color of the background, respectively. Therefore, this dataset is a typical semantic vagueness problem. The figure 2 shows the sample of the Colored-MNIST dataset.

C.2 Visualization of Evidence on Semantic Vagueness Phenomenon

To obtain a more intuitive understanding of how CCML tackles the semantic vagueness problem, we visualized the evidence of CCML on the training process of the Colored-MNIST dataset. This dataset consists of 30 categories, with categories 1 to 10 representing blue digits ‘0’ to ‘9’, categories 11 to 20 representing green digits ‘0’ to ‘9’, and categories 21 to 30 representing red digits ‘0’ to ‘9’. Here, we visualize samples of category 1 (digit ‘0’ with blue background). As shown in figure 3, we can see that view 1 can distinguish between different background colors, resulting in stronger evidence for categories 1 to 10 as it recognizes the blue background. On the other hand, view 2 can only differentiate shapes of digits based on grayscale, leading to higher evidence for specific categories (categories 1, 11, 21) compared to others in view 2. Finally, accurate classification results are obtained through the fusion of evidence by CCML, demonstrating CCML’s ability to effectively utilize complementary information across views and tackle the semantic vagueness phenomenon.

C.3 Different Method for Increase Separation Degree

To further explain the advantage of the method for increasing separation degree, we compared the results of the two methods on all datasets. Specifically, for the first method, we convert the separation degree to a negative value and add 1 to ensure that adding it to the overall loss function increases the separation degree:

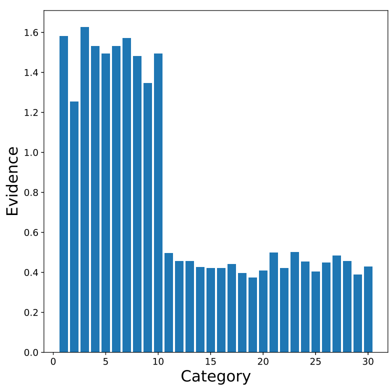
$$\mathcal{L}_{SD} = \sum_{i=1}^{C} \sum_{i \neq j}^{C} (1 - |b_i - b_j|). \tag{1}$$

We have conducted these two methods on all datasets and obtained the result, As shown in the table 1. We observe higher accuracy on all datasets by using our method instead of adding the negative separation degree to the loss function and variation by only reserving separation degree increase in the training process. The possible reason is that the first method can not improve the separation degree of the test set at the evidence level. When the trained model is used to obtain the evidence of the test sample, the separation degree of the consistent evidence of test samples cannot be improved, and the accuracy of the final result is affected. Our method can ensure that while improving the separation degree of consistent evidence of test samples, the total amount of overall evidence remains unchanged, and the final uncertainty remains unchanged to avoid the impact on the reliability of results.

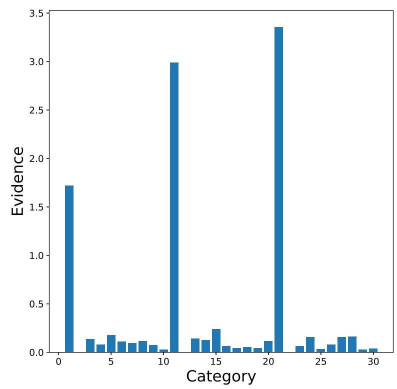
C.4 Hyper-parameters Detail

We observe that different datasets have different sensitivities to the hyper-parameters, appropriate parameter values can improve the overall performance of the model. Therefore, we set different hyper-parameters for each dataset based on the experiments. The table 2 shows the details of the hyper-parameters on different datasets, and the hyperparameter δ is set to 1.

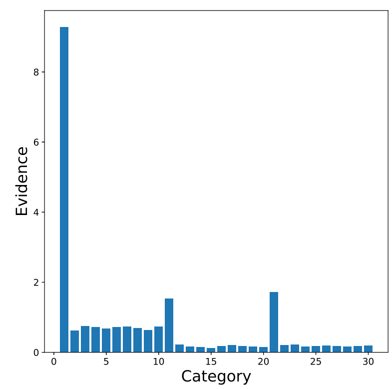
(a) Category Setting of Colored-MNIST Dataset



(b) View 1 (Identified the blue background)



(c) View 2 (Identified the digit '0')



(d) Final

Figure 3: The visualization of evidence on the instance of category 1 on the Colored-CNIST dataset.

Table 1: Different Methods for increasing separations degree

Method	Handwritten	Scene15	CUB	LandUse	PIE	Colored-MNIST
Add to Loss Function	98.55 ± 0.33	72.81 ± 1.81	91.24 ± 2.21	56.71 ± 1.74	91.03 ± 1.94	87.00 ± 2.06
Ours	99.28 ± 0.08	74.76 ± 0.85	94.58 ± 1.30	58.70 ± 1.75	94.56 ± 1.83	91.54 ± 1.48

Table 2: Hyperparameters of CCML

Hyperparameter	Handwritten	Scene15	CUB	LandUse	PIE	Colored-MNIST
η	1	0.5	0.3	0.5	1	1
γ	0.1	0.1	0.1	1	0.1	0.1
β	1.5	1.5	1.5	1.5	1	3