Supplementary material for the article "Fuzzy Inference Attention Module for Unsupervised Domain Adaptation".

Zhengshan Wang, Long Chen, *Senior Member, IEEE* and Fei-Yue Wang, *Fellow, IEEE*

S.I. Training time experiments of FIA

In order to prove that the proposed FIA will not significantly increase the training time of the model, we conducted separate tests to measure the running time of both the original model and the FIA-enhanced model. These experiments were conducted on an RTX3090 computing platform, with a batch size of 32, and we recorded the training time required for every thousand (k) iterations.

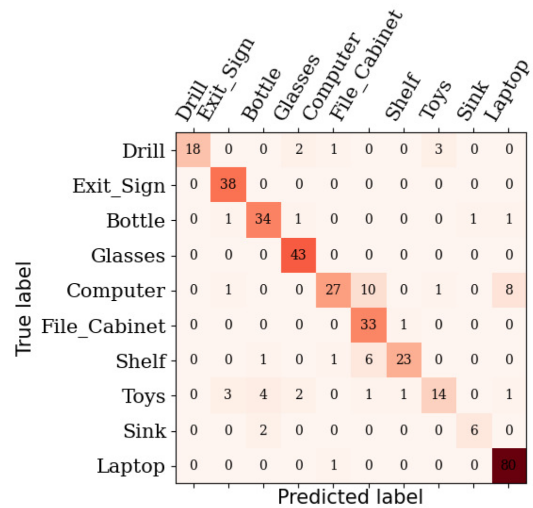
As shown in Table SI, the training times for DANN, CDAN, MDD, and MCC are 755s, 759s, 769s, and 890s, respectively. After incorporating FIA, the average training time increased by 13s, which corresponds to an average increase of 1.6%. This additional computational time is negligible compared to the original algorithm, providing evidence for the low overhead nature of our proposed module.

TABLE SI
Training time of different methods

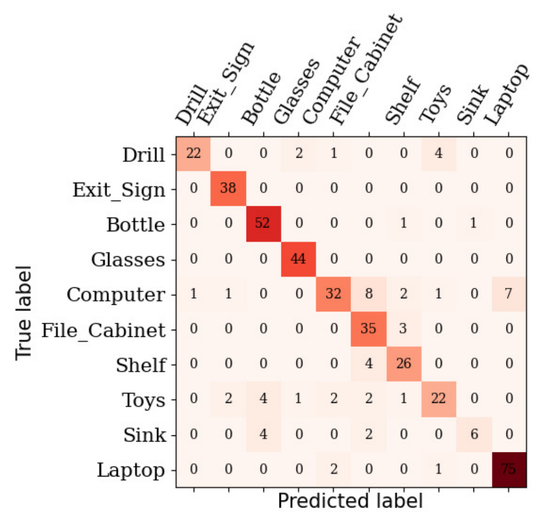
Methods	DANN	CDAN	MDD	MCC
Time(s)	755	758.8	769.1	890
Methods	DANN+f	CDAN+f	MDD+f	MCC+f
Time(s)	768	771.6	782	902

S.II. Confusion matrix visualization

In order to better demonstrate the effects of FIA, we have included a comparison of confusion matrix plots between original method and the FIA-enhanced method. We chose the Office Home dataset, with the Artistic (Ar) domain as the source domain and the Clipart (Cl) domain as the target domain. Specifically, we selected classes with labels 0 to 10 from the OfficeHome dataset, namely ('Drill', 'Exit Sign', 'Bottle', 'Glasses', 'Computer', 'File Cabinet', 'Shelf', 'Toys', 'Sink', 'Laptop'). We visualized the confusion matrix for both MCC and MCC+FIA. The Fig.S1 presents the comparison, clearly demonstrating the improved performance after incorporating the FIA module.



(a)

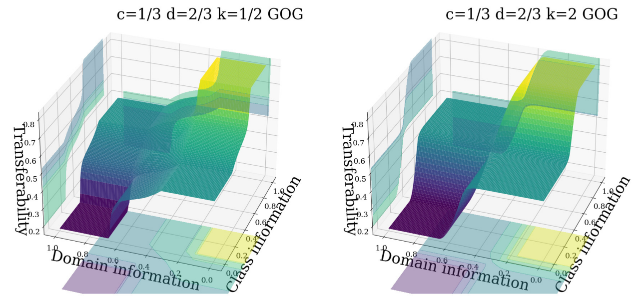


(b)

Fig. S1. Comparison between the confusion matrix plots for MCC and MCC+FIA.(a) MCC (b) MCC+f

S.III. Parameter Sensitivity Analysis of FIA

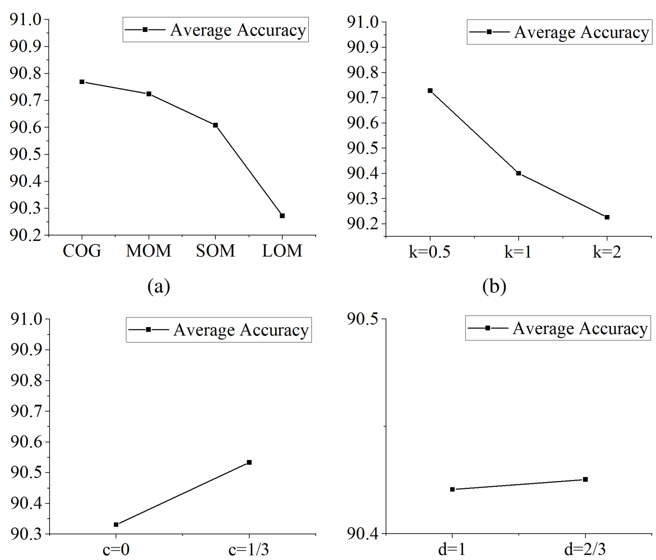
The FIA module is a dual-input, single-output system that incorporates expert prior knowledge. The input values are first fuzzified, which involves mapping the input values to fuzzy sets using membership functions. Then, fuzzy rules are applied to obtain the aggregated output. Finally, the aggregated output sets are defuzzified to obtain crisp output values. The parameters c, d, and k are utilized to regulate the membership functions and four defuzzification strategies are used to defuzzify the aggregated fuzzy sets. To better understand how the FIA handle the coupling relationship between the two types of information, we have created 3D fuzzy surface diagrams with class information and domain information on the x and y axes, respectively, and transferability on the z axis. Fig.S2 shows the 3D fuzzy surface with different settings.



(a) (b)

Fig. S2. 3D fuzzy surface diagrams of FIA under different settings

To evaluate the performance of FIA under various settings, experiments and comparisons are conducted between them. Experiment results of DANN+f under different settings are shown in Table SII. The leftmost column lists the four most commonly used defuzzification methods, and the next column lists the three parameters in the FIA module: k, c, and d. All results in the table are averaged by group and four line charts are drawn based on the averaged values. The line charts are presented in Fig.S3, which includes four subfigures demonstrating the influence of the defuzzification strategies, k, c, and d, respectively.



(a) (b) (c) (d)

Fig. S3. Comparison of FIA's performance with different settings on Office31 dataset. The vertical axis in this graph represents the accuracy. The horizontal axis represents the average accuracy of FIA under different settings. (a) (b) (c) and (d) demonstrate the influence of the defuzzification strategies, k, c, and d, respectively.

TABLE SII
PERFORMANCE (%) OF FIA WITH DIFFERENT SETTINGS ON OFFICE31 DATASET

defuzzy method	k	c	d	DANN +f (A→D)	DANN +f (A→W)	CDAN +f (A→D)	CDAN +f (A→W)	Avg
COG	1/2	0	1	86.7	89.4	95.0	93.8	91.2
COG	1/2	0	2/3	85.7	91.2	93.0	92.8	90.7
COG	1/2	1/3	1	85.1	90.1	93.0	93.0	90.3
COG	1/2	1/3	2/3	83.9	91.4	93.0	94.1	90.6
COG	1	0	1	85.1	94.1	94.8	89.3	90.8
COG	1	0	2/3	85.1	90.3	95	90.2	90.2
COG	1	1/3	1	85.5	90.1	93.2	91.9	90.2
COG	1	1/3	2/3	85.5	89.7	93.7	91.2	90.0
COG	2	0	1	85.5	90.8	94.0	93.0	90.8
COG	2	0	2/3	85.9	92.3	94.0	94.0	91.6
COG	2	1/3	1	84.7	91.7	95.2	92.8	91.1
COG	2	1/3	2/3	84.5	93.5	94.2	93.0	91.3
MOM	1/2	0	1	85.1	94.2	92.8	91.7	91.0
MOM	1/2	0	2/3	84.9	91.1	90.0	94.0	90.0
MOM	1/2	1/3	1	87.1	91.6	92.6	92.5	91.0
MOM	1/2	1/3	2/3	85.9	91.1	95.0	91.1	90.8
MOM	1	0	1	85.1	91.3	94.6	90.8	90.5
MOM	1	0	2/3	86.3	91.2	94.8	90.2	90.6
MOM	1	1/3	1	84.9	90.8	94.4	93.6	90.9
MOM	1	1/3	2/3	86.7	92.7	94.2	92.7	91.6
MOM	2	0	1	85.3	91.3	92.9	94.3	91.0
MOM	2	0	2/3	83.7	90.9	95.2	91.9	90.4
MOM	2	1/3	1	86.1	94.6	90.0	92.5	90.8
MOM	2	1/3	2/3	85.1	91.3	93.4	93.2	90.8
SOM	1/2	0	1	86.5	90.2	91.4	93.1	90.3
SOM	1/2	0	2/3	82.5	93.1	93.6	92.2	90.4
SOM	1/2	1/3	1	84.1	92.5	93.0	94.0	90.9
SOM	1/2	1/3	2/3	86.1	90.9	94.4	95.0	91.6
SOM	1	0	1	85.7	87.7	96.4	90.8	90.2
SOM	1	0	2/3	85.3	91.7	90.6	93.5	90.3
SOM	1	1/3	1	86.1	93.6	94.6	92.5	91.7
SOM	1	1/3	2/3	85.1	93.2	93.4	90.6	90.6
SOM	2	0	1	85.1	91.7	92.8	91.9	90.4
SOM	2	0	2/3	86.5	90.6	93.8	91.2	90.5
SOM	2	1/3	1	84.9	93.0	91.6	92.1	90.4
SOM	2	1/3	2/3	85.3	92.2	90.8	92.3	90.2
LOM	1/2	0	1	84.7	92.8	92.8	91.8	90.5
LOM	1/2	0	2/3	85.7	93.5	92.9	92.2	91.1
LOM	1/2	1/3	1	86.3	93.3	92.0	91.5	90.8
LOM	1/2	1/3	2/3	84.3	90.2	94.6	93.5	90.7
LOM	1	0	1	82.3	90.1	94.4	91.0	89.5
LOM	1	0	2/3	85.1	89.7	93.4	93.3	90.4
LOM	1	1/3	1	84.5	91.6	93.6	93.5	90.8
LOM	1	1/3	2/3	85.5	91.8	90.8	91.2	89.8
LOM	2	0	1	84.7	89.8	89.0	92.3	89.0
LOM	2	0	2/3	82.5	92.1	95.0	93.3	90.7
LOM	2	1/3	1	84.5	91.8	93.6	93.7	90.9
LOM	2	1/3	2/3	82.7	90.9	90.8	92.5	89.2