

**Table 6: Performance on NUS-WIDE dataset for teacher and student models.**

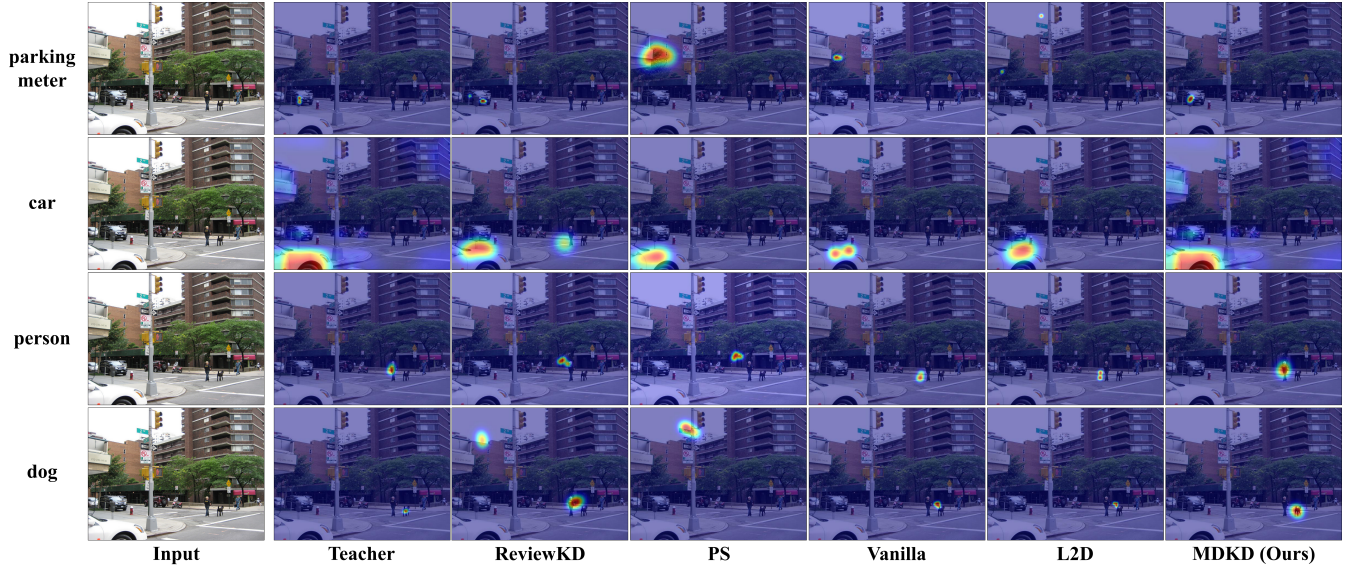
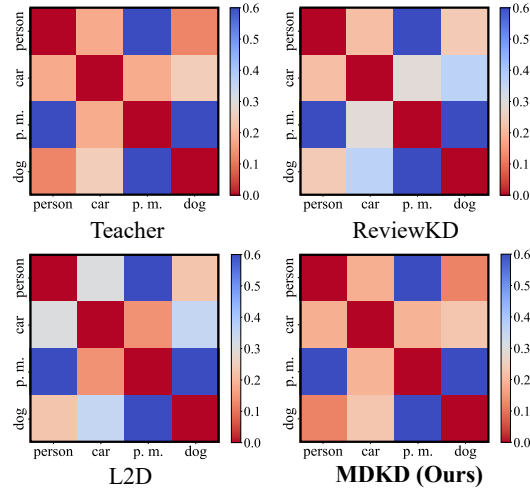
Teacher	ResNet-101			Swin-T			WRN-101		
Student	ResNet-34			MobileNet v2			WRN-50		
Metrics	mAP	OF1	CF1	mAP	OF1	CF1	mAP	OF1	CF1
Teacher	55.15	75.22	61.25	59.15	77.98	65.44	57.54	76.98	63.87
Student	53.81	74.64	59.08	54.65	75.61	61.98	56.69	76.02	62.48
RKD [17]	53.96	75.02	59.45	54.40	75.26	61.77	56.89	76.56	62.59
PKT [18]	54.14	74.84	59.98	55.59	75.15	62.01	57.02	76.45	62.72
ReviewKD [2]	54.08	74.78	60.15	55.06	75.98	62.48	57.45	76.18	62.88
MSE [25]	54.33	75.15	60.25	54.98	75.82	61.88	57.15	76.48	62.11
PS [22]	54.96	74.69	59.15	55.32	75.17	62.82	56.98	76.89	63.41
L2D [26]	55.03	75.33	59.96	56.38	76.21	63.04	57.22	77.33	63.05
FitNet [1]	52.47	75.04	58.48	54.90	76.00	61.53	56.51	76.68	62.17
<b>MDKD</b>	<b>55.08</b>	<b>76.15</b>	<b>62.58</b>	<b>58.46</b>	<b>76.69</b>	<b>63.44</b>	<b>58.05</b>	<b>78.44</b>	<b>64.09</b>

**Table 7: Comparison between the teacher model and the proposed MDKD on the third-order label pair dependencies, which are measured by the mean square difference and square difference, respectively (lower values indicating higher correlation). We present the scores among four labels: *parking meter* (p. m. for short), *car*, *person*, and *dog*. The input image is shown in Figure 1.**

Label pair dependencies	Batch-level				Instance-level			
	Teacher	ReviewKD	L2D	MDKD	Teacher	ReviewKD	L2D	MDKD
person-car-p. m.	0.179	0.251	0.222	0.187	0.003	0.007	0.006	0.004
person-car-dog	0.211	0.289	0.327	0.206	0.005	0.009	0.020	0.006
person-p. m.-car	0.432	0.446	0.425	0.441	0.005	0.012	0.012	0.006
person-p. m.-dog	0.714	0.631	0.672	0.722	0.009	0.020	0.017	0.010
person-dog-car	0.183	0.302	0.287	0.174	0.006	0.009	0.024	0.006
person-dog-p. m.	0.432	0.449	0.429	0.436	0.007	0.017	0.015	0.007
car-person-p. m.	0.431	0.408	0.507	0.436	0.005	0.010	0.009	0.005
car-person-dog	0.149	0.225	0.265	0.151	0.003	0.007	0.008	0.002
car-p. m.-Dog	0.462	0.475	0.387	0.472	0.007	0.017	0.013	0.009
car-dog-p. m.	0.494	0.513	0.492	0.491	0.009	0.018	0.027	0.011
p. m.-person-dog	0.403	0.420	0.468	0.404	0.005	0.011	0.013	0.005
p. m.-car-dog	0.212	0.328	0.244	0.210	0.006	0.010	0.022	0.007

**Table 8: Comparison between the teacher model and the proposed MDKD on the fourth-order label pair dependencies, which are measured by the mean square difference and square difference, respectively (lower values indicating higher correlation). We present the scores among four labels: *parking meter* (p. m. for short), *car*, *person*, and *dog*. The input image is shown in Figure 1.**

Label pair dependencies	Batch-level				Instance-level			
	Teacher	ReviewKD	L2D	MDKD	Teacher	ReviewKD	L2D	MDKD
person-car-p. m.-dog	0.367	0.387	0.359	0.376	0.005	0.013	0.010	0.007
person-car-dog-p.m.	0.389	0.413	0.429	0.388	0.007	0.014	0.019	0.008
person-p. m.-car-dog	0.370	0.419	0.400	0.370	0.007	0.012	0.020	0.007
person-p. m.-dog-car	0.557	0.543	0.565	0.557	0.009	0.017	0.023	0.010
person-dog-car-p. m.	0.182	0.298	0.238	0.180	0.005	0.009	0.019	0.006
person-dog-p. m.-car	0.348	0.396	0.333	0.354	0.005	0.014	0.013	0.007
car-person-p. m.-dog	0.535	0.492	0.550	0.542	0.007	0.015	0.012	0.007
car-person-dog-p. m.	0.347	0.370	0.388	0.352	0.005	0.013	0.011	0.006
car-p. m.-person-dog	0.328	0.377	0.358	0.333	0.005	0.011	0.012	0.005
car-dog-person-p. m.	0.350	0.402	0.429	0.346	0.007	0.011	0.021	0.006
p. m.-person-car-dog	0.369	0.394	0.455	0.367	0.006	0.011	0.018	0.006
p. m.-car-person-dog	0.160	0.247	0.223	0.164	0.003	0.008	0.008	0.003

Figure 7: Attention maps for *parking meter*, *car*, *person*, and *dog* by different methods.Figure 8: The batch-level second-order label pair dependencies of different methods, measured by mean square difference (lower values indicating higher correlation). The scores among four labels: *parking meter* (p. m. for short), *car*, *person*, and *dog* are presented. The input image is shown in Figure 1.