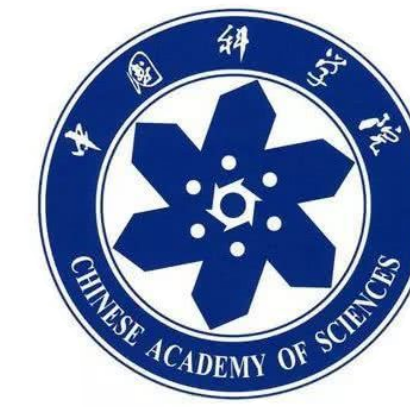


# MADiMa 2019 Mixed-dish Recognition with Contextual Relation Networks

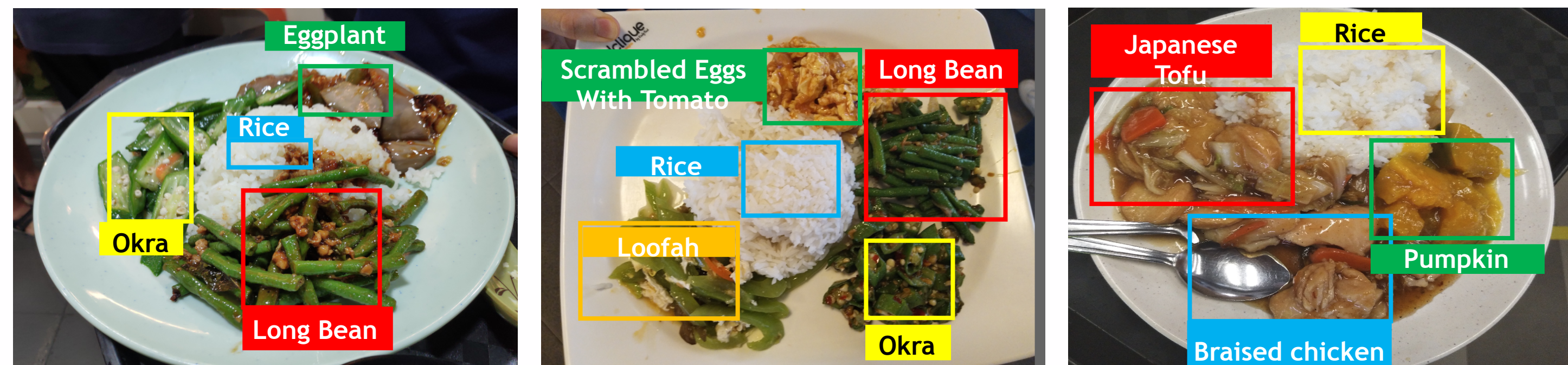


Lixi Deng, Jingjing Chen, Qianru Sun, Xiangnan He, Sheng Tang, Zhaoyan Ming, Yongdong Zhang, Tat-Seng Chua



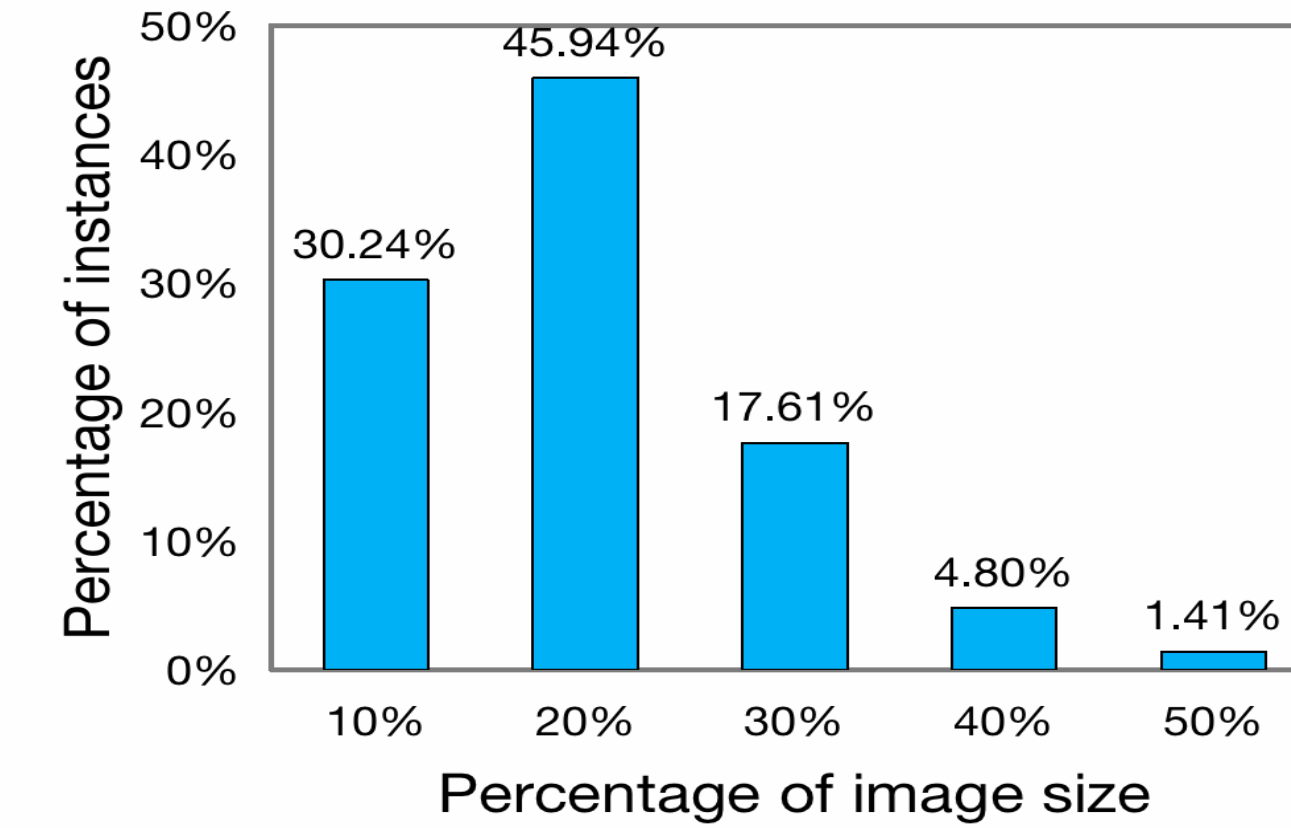
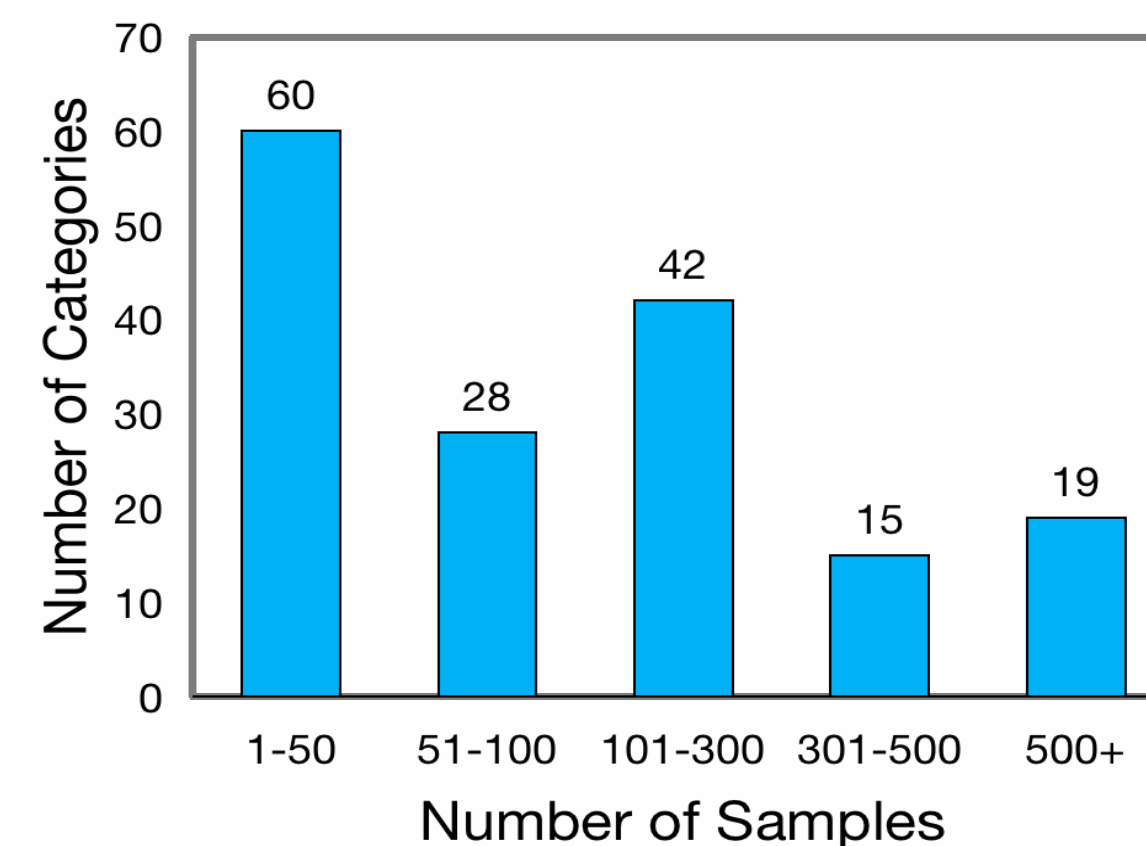
## Objective

>> Identify each of the dish presented in one plate.

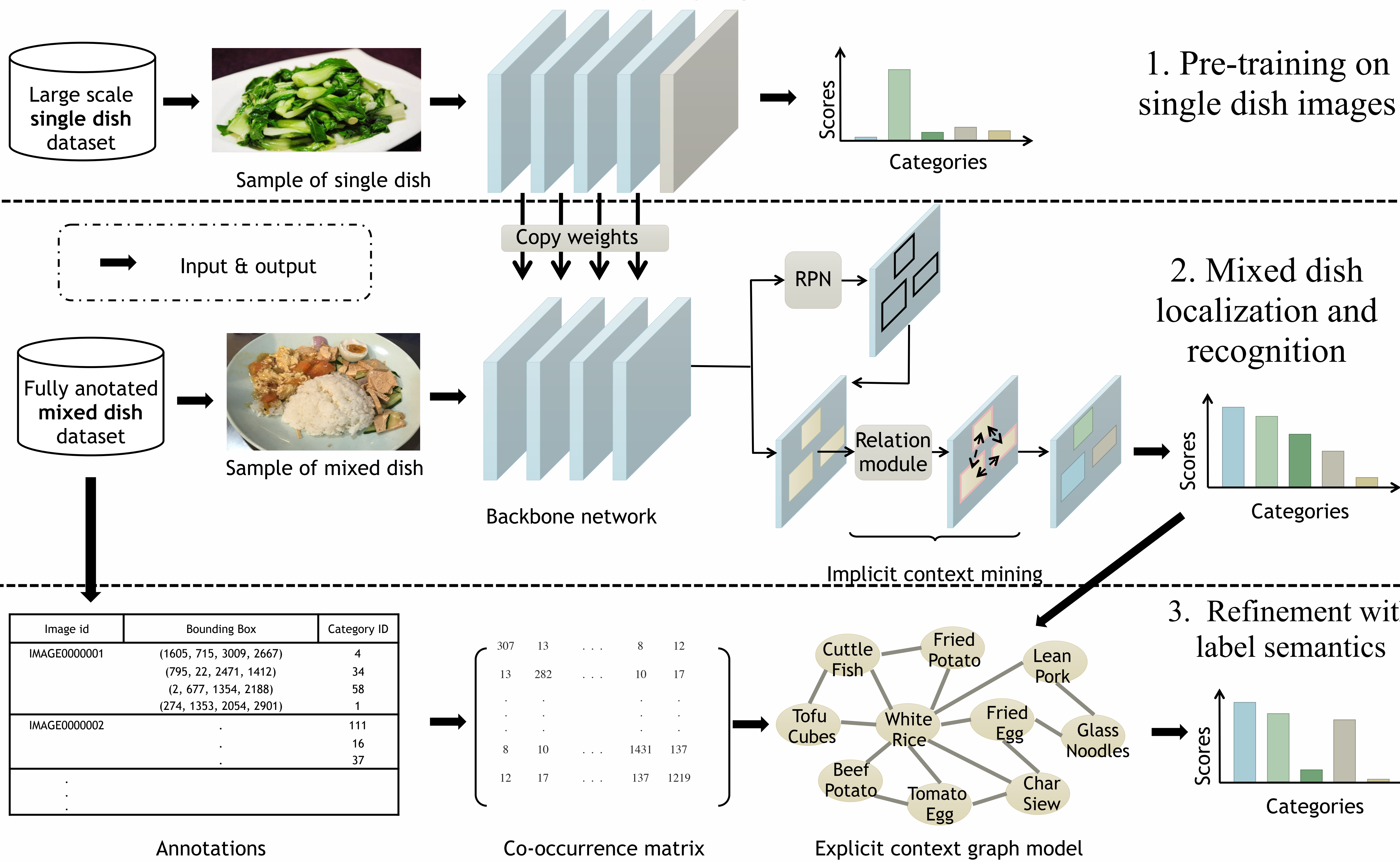


## Mixed-dish Dataset

>> 9,254 Images from 6 canteens; 39,668 Bounding boxes in total.



## Framework

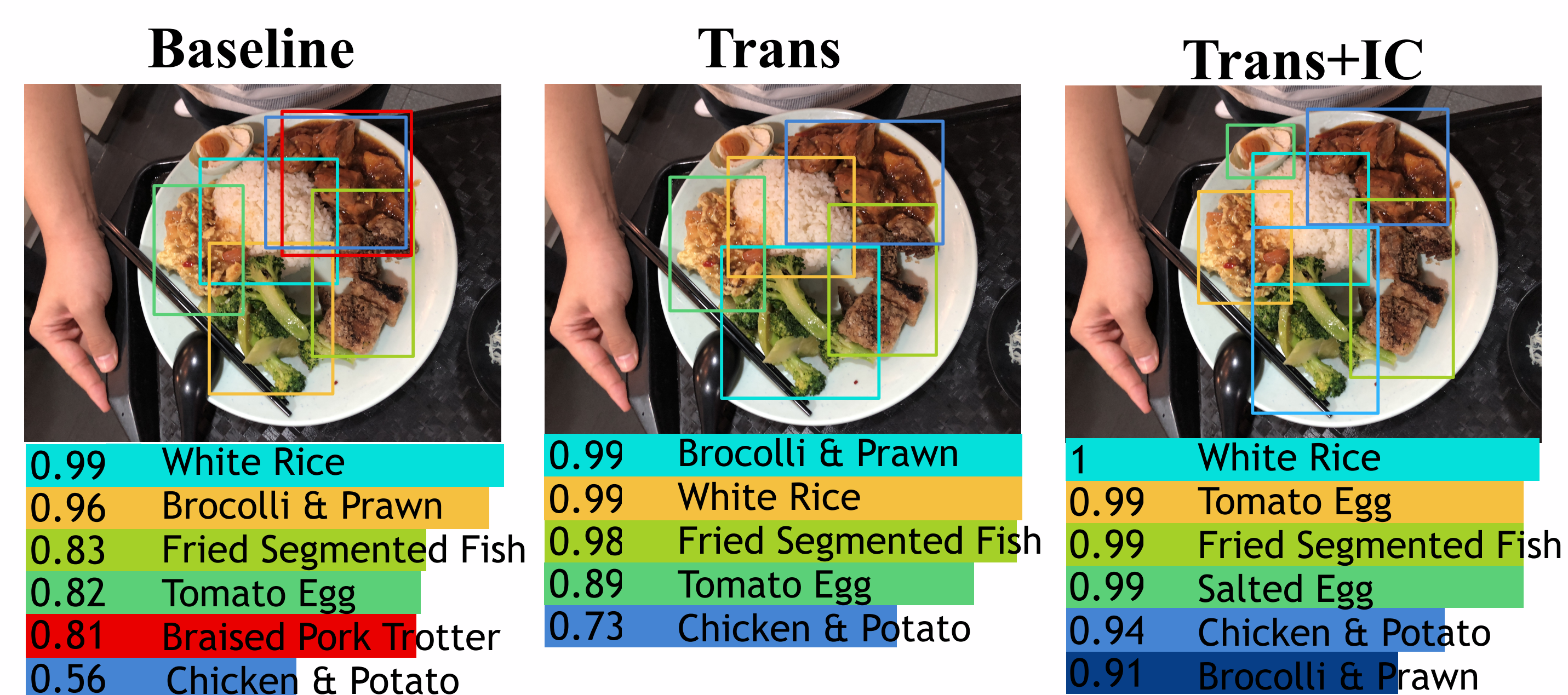


## Results

>> Comparison with State-of-the-arts

	Precision (%)	Recall (%)	F1 (%)
ResNet-50	44.70	44.79	44.74
ResNet-50*	49.79	49.87	49.82
Region-wise	70.92	70.85	70.88
Faster R-CNN	86.53	82.73	84.59
CR-Nets	87.74	89.12	88.42

>> Case study



>> Ablation study

	Trans?	IC?	EC?	Cross-domain		Self-domain	
				mAP (%)	F1 (%)	mAP (%)	F1 (%)
(a)				40.63	49.4	73.49	84.59
(b)	✓			44.96	51.05	74.17	84.83
(c)		✓		44.39	51.47	76.32	86.24
(d)			✓	-	50.86	-	85.56
(e)	✓	✓		47.57	52.58	76.89	86.89
(f)		✓	✓	-	52.37	-	87.63
(g)	✓		✓	-	52.22	-	85.59
(h)	✓	✓	✓	-	53.55	-	88.42

>> Conclusion

1. With knowledge transfer, it attains higher performance of dish detection.
2. Implicit context is quite effective in improving the dish detection performance.
3. Explicit context also improves the dish recognition performances.