南京航空航天大學
Nanjing University of Aeronautics and Astronautics

模式分析与机器智能
工业和信息化部重点实验室
MIIT Key Laboratory of
Pattern Analysis & Machine Intelligence

Bridging the Gap between Model Explanations in Partially Annotated Multi-label Classification

CVPR 2023

Problem to be solved : Multi-Label Learning

	person	dog	bus	bicycle	apple	boat	laptop	couch
(a)	✓	✗	✓	✓	✗	✗	✗	✗
(b)	✓	✗	?	✓	?	✗	?	?
(c)	✓	?	?	?	?	?	?	?

(a) full annotations

(b) partial annotations

(c) single positive label

Problem to be solved : Multi-Label Learning

False Negative

	person	dog	bus	bicycle	apple	boat	laptop	couch
(a)	✓	✗	✓	✓	✗	✗	✗	✗
(b)	✓	✗	$\boxed{0}$	✓	?	✗	?	?
(c)	✓	?	?	?	?	?	?	?

(a) full annotations

(b) partial annotations

(c) single positive label

expensive to fully annotated

Related work : Partially annotated multi-label classification

Partial annotated multi-label classification

(1) treat all unobserved labels as missing labels

(2) treat all unobserved labels as negative labels

ROLE : from 'Multi-label learning from single positive label'

LL-R : reject false negative label;

LL-Ct : correct false negative label;

LL-Cp : correct false negative label to positive label;

CAM : class activation map

Method : partially annotated multi-label classification

Baseline approach :

Assuming unannotated labels as **N**egative labels (**AN**)

	(a)	(b)		(c)	
person	1	1	1	1	1
dog	0	0	0		0
bus	1		0		0
bicycle	1	1	1		0
apple	0		0		0
boat	0	0	0		0
laptop	0		0		0
couch	0		0		0

(a) full annotations

(b) partial annotations

(c) single positive label

false negative label

Disadvantage :

label noise (false negative label)

Method : partially annotated multi-label classification

Baseline approach :

Assuming unannotated labels as **N**egative labels (**AN**)

$$\mathcal{L}_{AN} = \frac{1}{C}\left[\sum_{i \in \mathcal{I}^p} \mathcal{L}_+ + \sum_{i \in \mathcal{I}^n \cup \mathcal{I}^\phi} \mathcal{L}_-\right] \quad (1)$$

$$\mathcal{L}_+ = -\log(\sigma(g_i)), \mathcal{L}_- = -\log(1-\sigma(g_i))$$

$\mathcal{I}^{tn}$ true negative

$\mathcal{I}^{fn}$ false negative

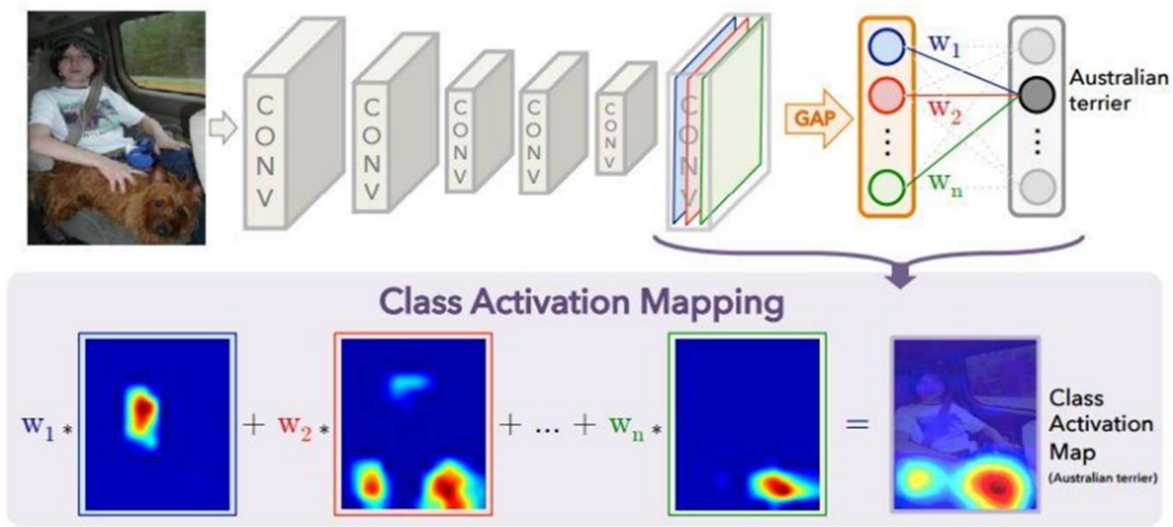
Analysis on model explanation

full annotations

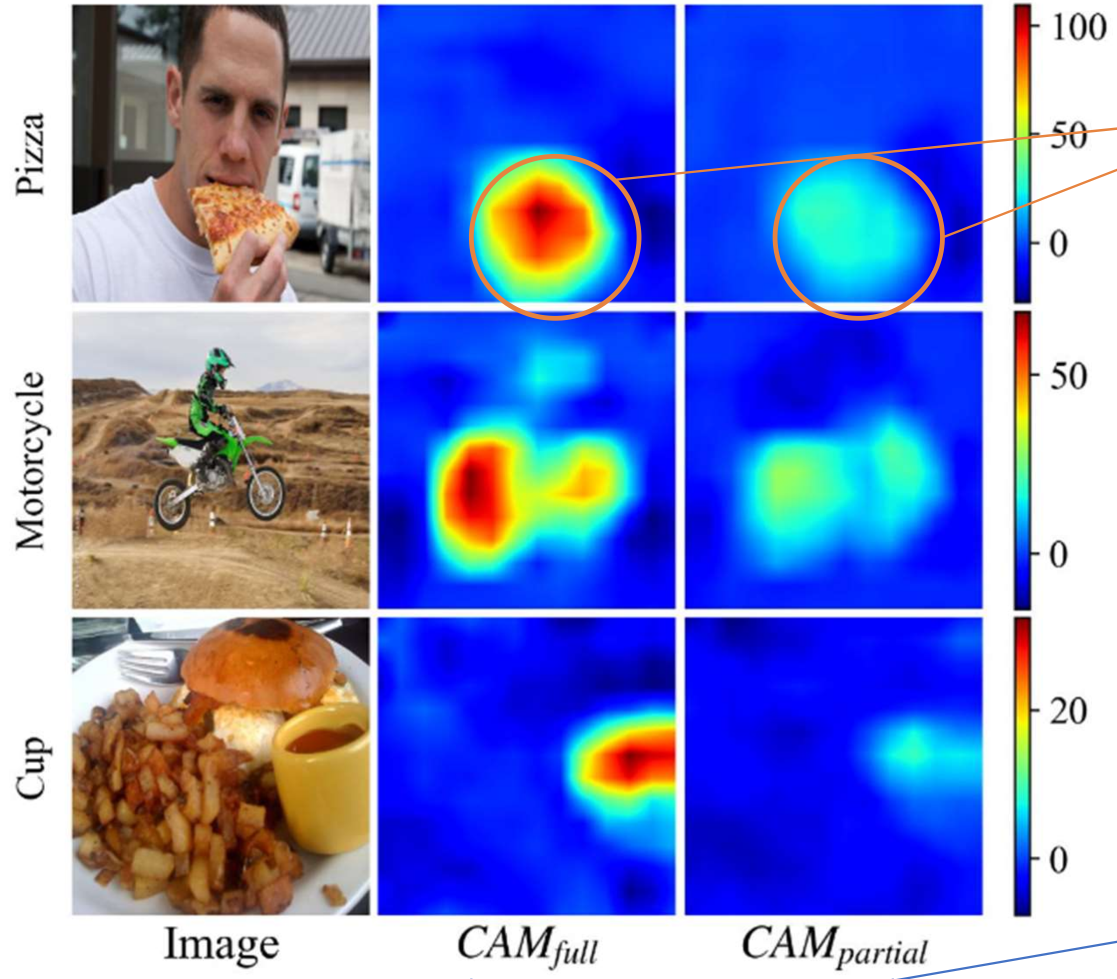
$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \rightarrow$ Model 1 (ResNet-50)

$\begin{bmatrix} 1 \\ 0 \\ \color{red}{0} \end{bmatrix} \rightarrow$ Model 2 (ResNet-50)

partial annotations



An interesting observation



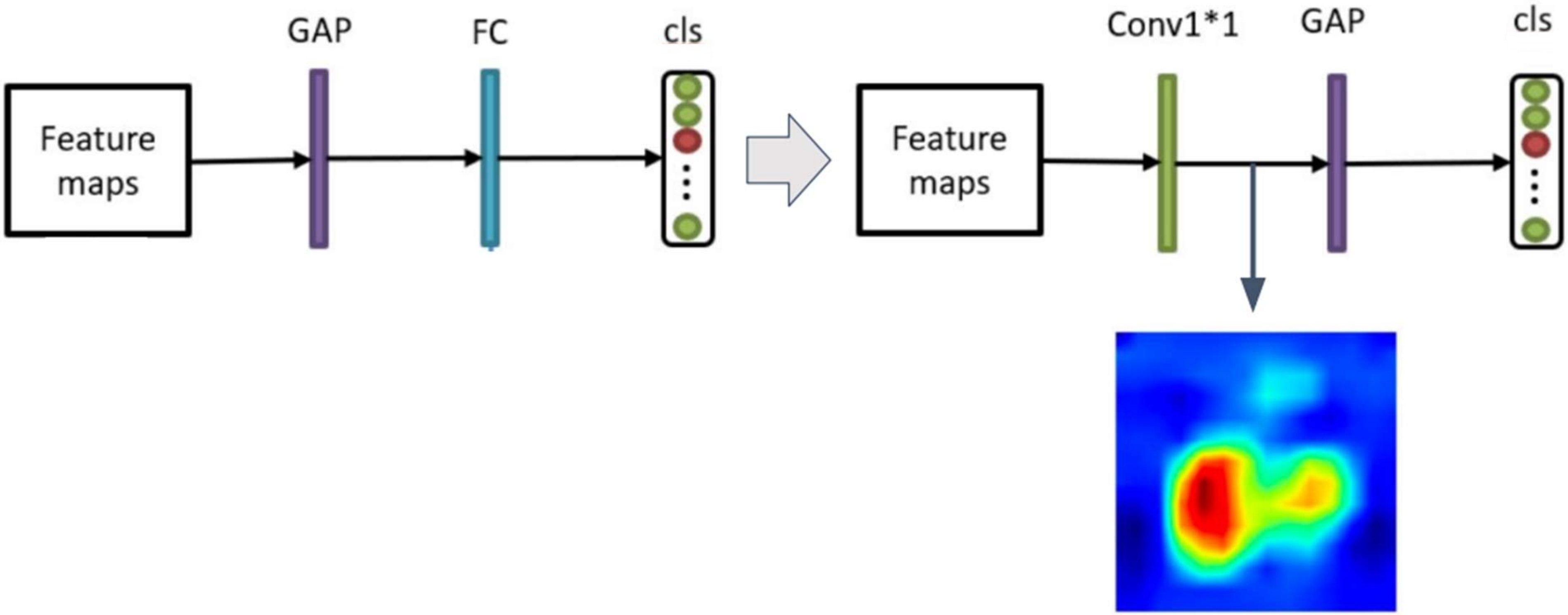
similar

From the view of model explanation, false negative labels don't effect the model's localization ability.

Bridge the gap !

Method : CAM

1) Modify CNN classificatio network architecture



Method : CAM

instead of post-processing, but forward pass

$$\boldsymbol{M}_c = \sum_{d=1}^{D} \boldsymbol{W}_{cd} \boldsymbol{F}_d \ , \quad (2)$$

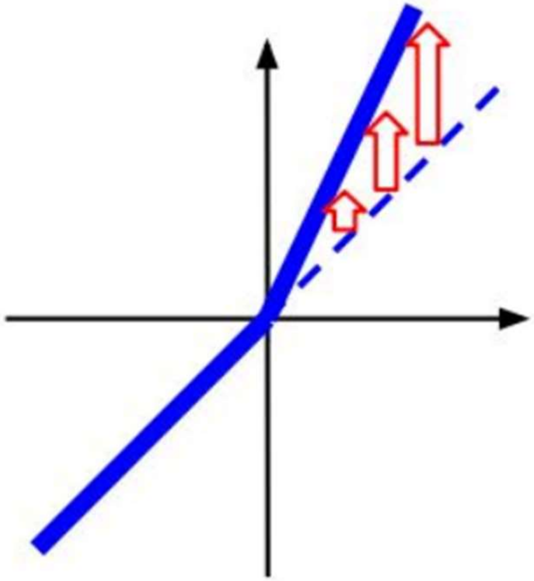
$$g_c = \frac{1}{HW} \sum_{i=1}^{H} \sum_{j=1}^{W} \boxed{(\boldsymbol{M}_c)_{ij}} . \quad (3)$$

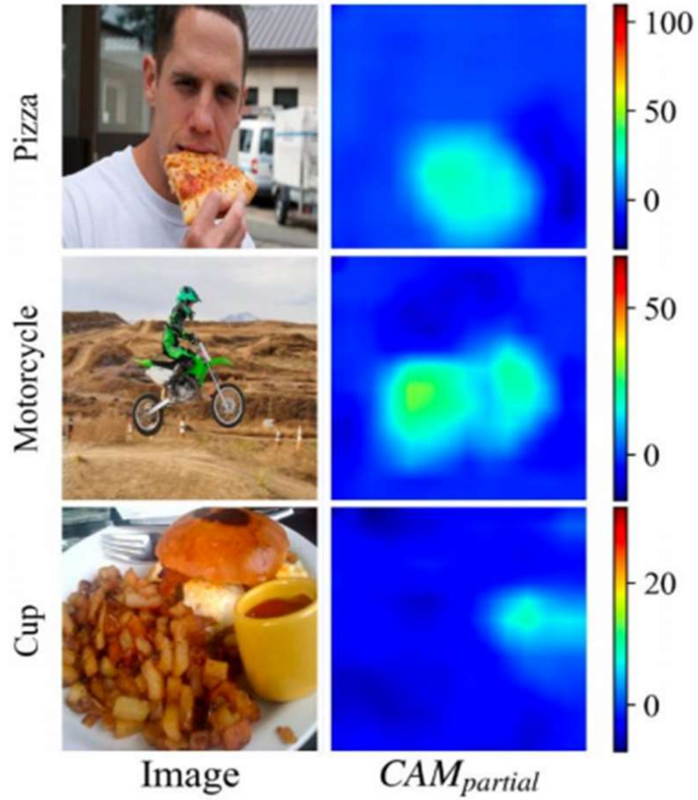
attribution score

Method : BoostLU

2) Apply BoostLU on CAM element-wisely

BoostLU(x) = max(x, αx)

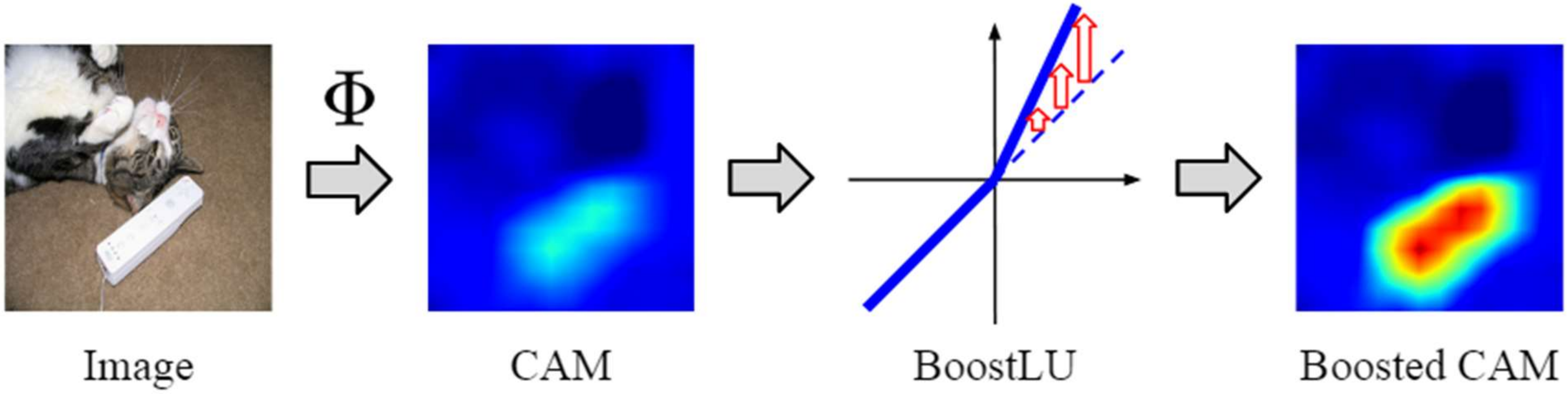




Method : BoostLU

2) Apply BoostLU on CAM element-wisely

BoostLU(x) = max(x, αx), set α=5



Experiments

Several scenarios for BoostLU application

i) Apply only in inference phase

-> Performance improves without additional training

BoostLU in inference	Performance	
	VOC	COCO
	86.10	64.58
✓	87.31	66.27

Experiments

Several scenarios for BoostLU application

ii) Apply also in training phase with large loss modification scheme

-> Performance improves further !

BoostLU in inference	BoostLU in training	LL-R in training	Performance VOC	Performance COCO
			86.10	64.58
✓			87.31	66.27
✓	✓	✓	**89.27**	**72.82**

Experiments : all ablation study

BoostLU in inference	BoostLU in training	LL-R in training	Performance VOC	Performance COCO
			86.10	64.58
✓			87.31	66.27
✓	✓		86.73	65.33
		✓	88.24	70.60
	✓	✓	87.18	68.45
✓		✓	88.90	70.87
✓	✓	✓	**89.27**	**72.82**

Ablation study on BoostLU and LL-R. We test seven combinations of using BoostLU and LL-R [21] on VOC and COCO datasets. Training a model with both LL-R and BoostLU and applying BoostLU during inference shows the best mAP.

Experiments : Single positive label

Methods	VOC	COCO	NUS	CUB
Full Label	89.42	76.78	52.08	30.90
AN	85.89	64.92	42.27	18.31
LS [30]	87.90	67.15	43.77	16.26
ASL [33]	87.76	68.78	46.93	18.81
ROLE [11]	87.77	67.04	41.63	13.66
ROLE + LI [11]	88.26	69.12	45.98	14.86
EM [50]	89.09	70.70	47.15	20.85
EM + APL [50]	89.19	70.87	47.59	**21.84**
LL-R [21]	88.27	70.70	48.76	19.56
+ BoostLU (Ours)	**89.29**	**72.89**	**49.59**	19.80
LL-Ct [21]	87.79	70.29	48.08	19.06
+ BoostLU (Ours)	88.61	71.78	48.37	19.25
LL-Cp [21]	87.44	70.27	47.92	19.21
+ BoostLU (Ours)	87.81	71.41	48.61	19.34

Designed by 4 multi-label classification datasets:

- CUB
- NUSWIDE
- MS COCO 2014
- PASCAL VOC

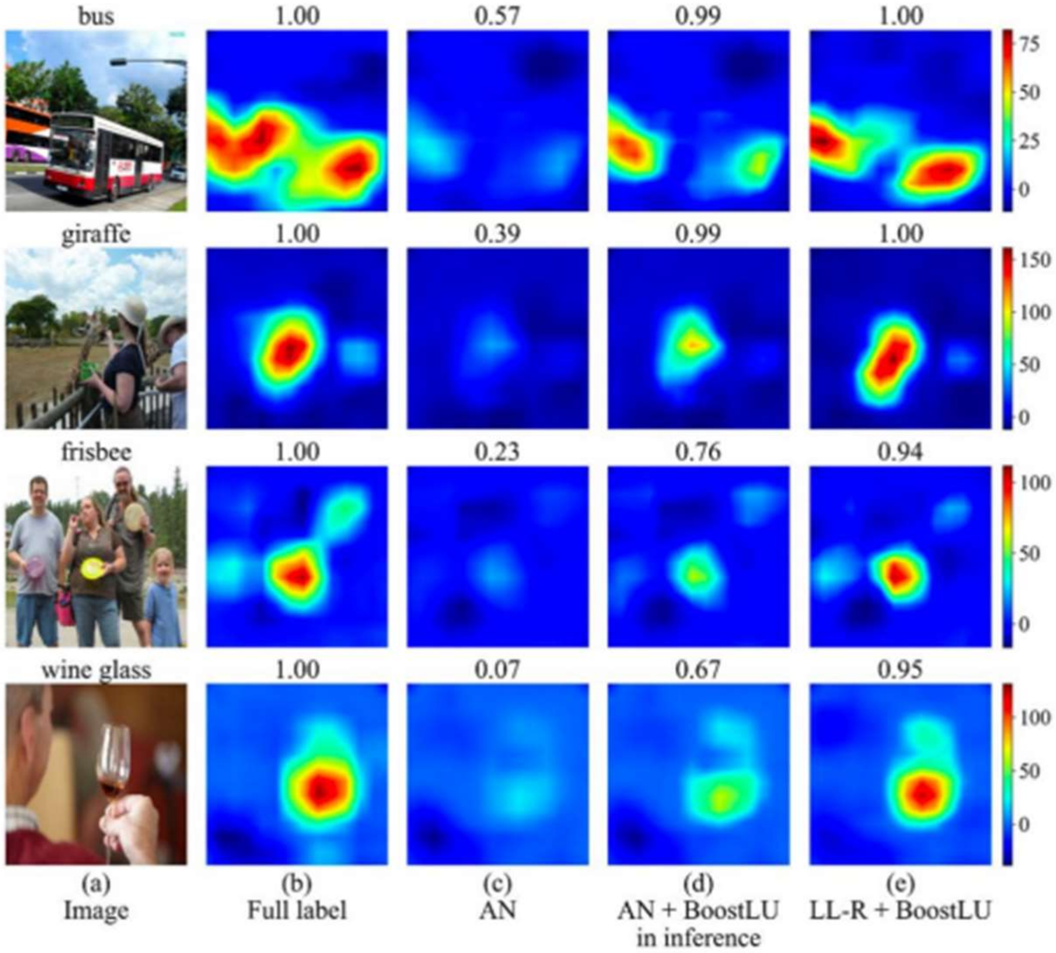
Experiments : Large-scale partial label (OpenImages V3 dataset)

Each group includes 1,000 classes without overlapping. Group 1 has the smallest annotations, and Group 5 has the most

Methods	Group 1	Group 2	Group 3	Group 4	Group 5	All Classes
CNN-RNN [39]	68.76	69.70	74.18	78.52	84.61	75.16
Curriculum Labeling [13]	70.37	71.32	76.23	80.54	86.81	77.05
IMCL [17]	70.95	72.59	77.64	81.83	87.34	78.07
P-ASL [2]	73.19	78.61	85.11	87.70	90.61	83.03
LL-R [21]	77.76	79.07	81.94	84.51	89.36	82.53
+ BoostLU (Ours)	79.28	80.81	83.32	85.63	90.27	83.86
LL-Ct [21]	77.76	79.18	81.97	84.46	89.51	82.58
+ BoostLU (Ours)	79.43	80.75	83.41	85.70	90.41	83.94
LL-Cp [21]	77.49	79.22	81.89	84.51	89.18	82.46
+ BoostLU (Ours)	79.53	81.04	83.40	85.85	90.39	**84.04**

Experiments

Qualitative results



南京航空航天大學
Nanjing University of Aeronautics and Astronautics

模式分析与机器智能
工业和信息化部重点实验室
MIIT Key Laboratory of
Pattern Analysis & Machine Intelligence

THANKS