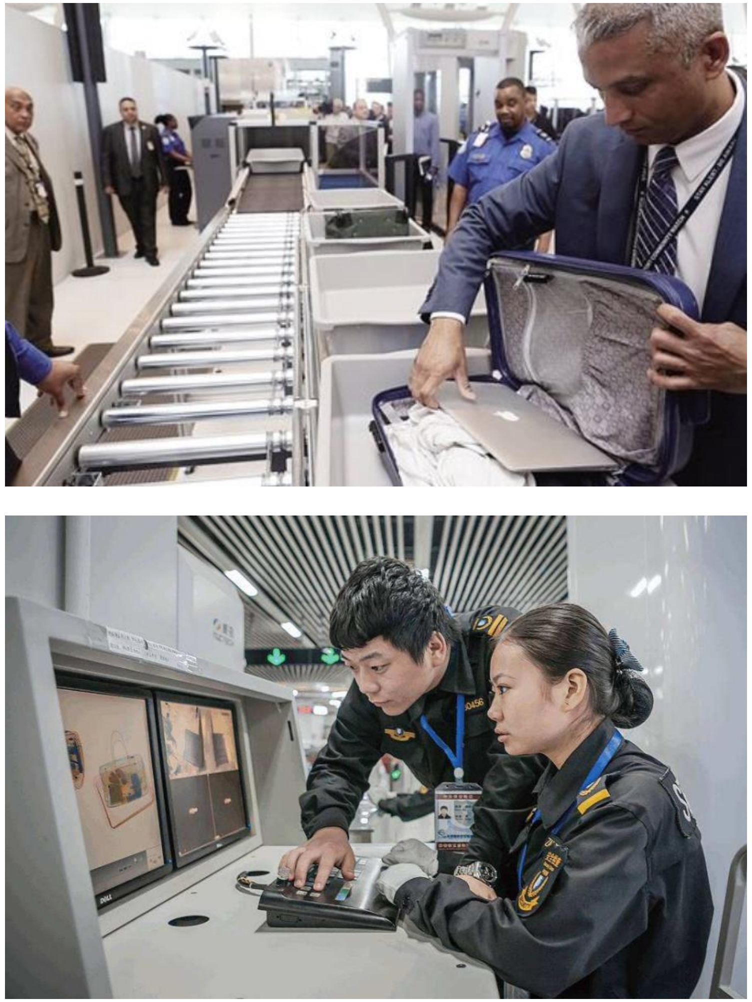
ICCV 2021 : Towards Real-world X-ray Security Inspection: A High-quality Benchmark and Lateral Inhibition Module for Prohibited Items Detection

Renshuai Tao[1,2] Yanlu Wei[1] Xiangjian Jiang[1] Hainan Li[1]
Haotong Qin[1] Jiakai Wang[1] Yuqing Ma[1] Libo Zhang[3] Xianglong Liu[1*]
[1]State Key Laboratory of Software Development Environment, Beihang University
[2]iFLYTEK Research [3]Institute of Software Chinese Academy of Sciences
{rstao, weiyanlu, hainan, qinhaotong, jk_buaa_scse, mayuqing, xlliu}@buaa.edu.cn
silencejiang12138@gmail.com, libo@iscas.ac.cn

陶仁帅
软件开发环境国家重点实验室
北京航空航天大学
科大讯飞AI研究院
个人主页：rstao95.github.io

Background

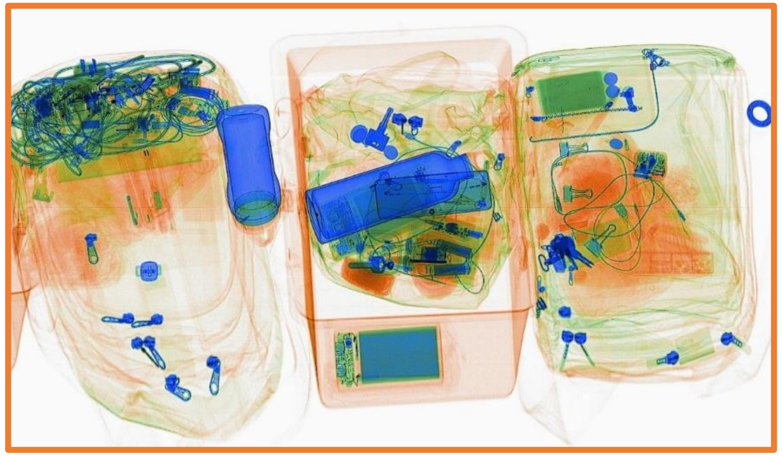
Deep learning applications



AI Security Inspection ?

Difficulties and Related Works

Difficulties:
(1) Various
(2) Small
(3) Randomly stacked
(4) Heavily overlapped

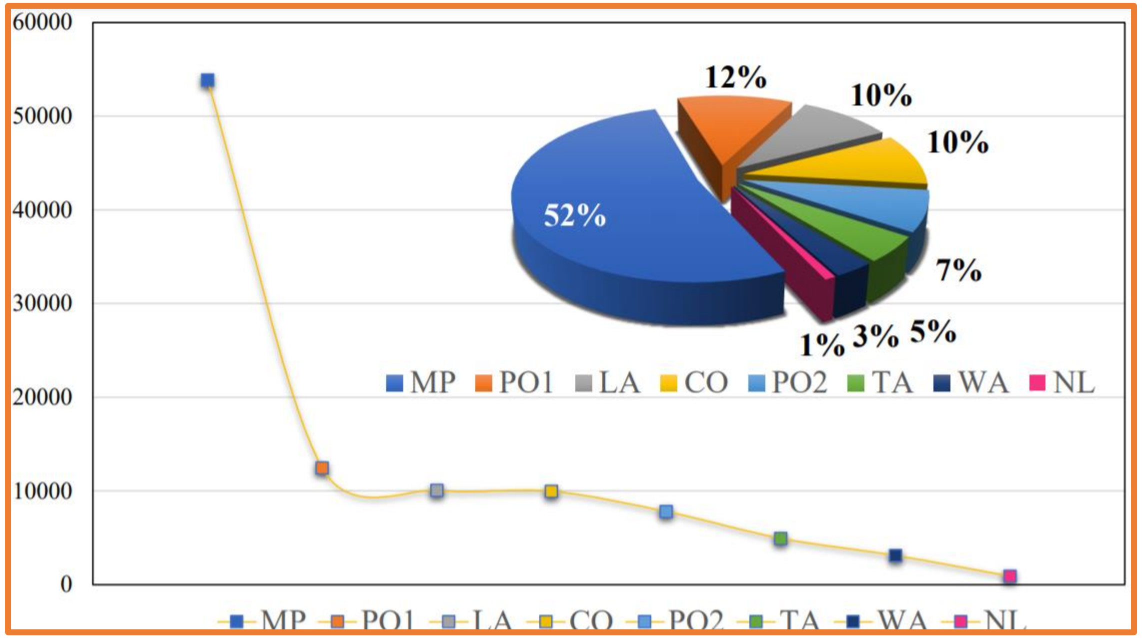


Related Works:

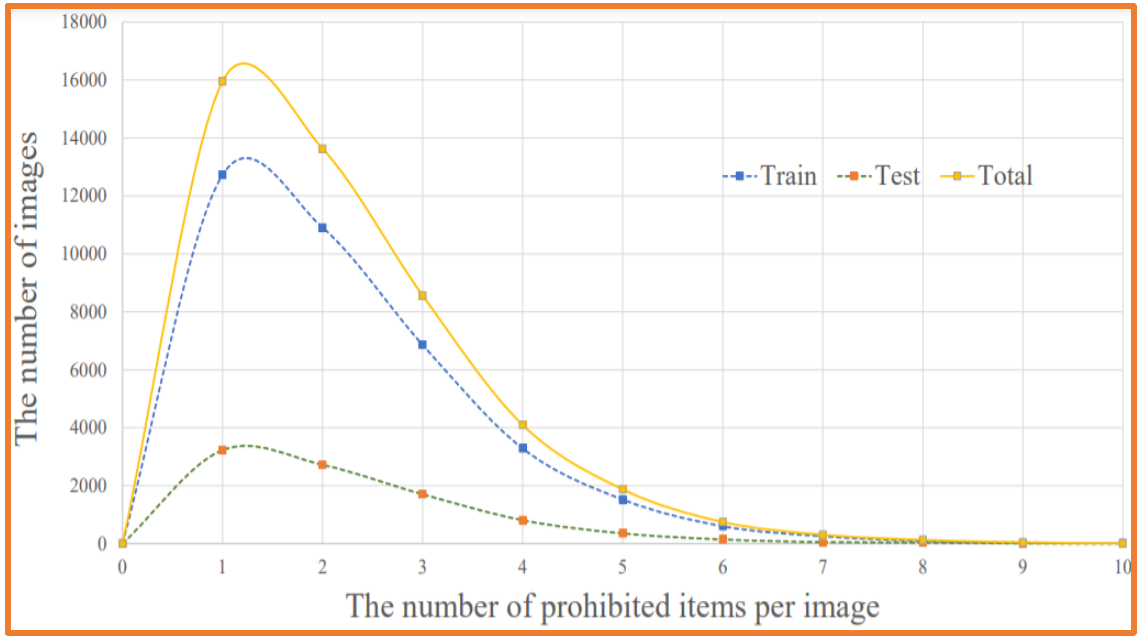
Dataset	Year	Category	N_p	Annotation			Color	Task	Data Source
				Bounding Box	Number	Professional			
GDXray [23]	2015	3	8,150	✓	8,150	✗	Gray-scale	Detection	Unknown
SIXray [25]	2019	6	8,929	✗	✗	✗	RGB	Classification	Subway Station
OPIXray [40]	2020	5	8,885	✓	8,885	✓	RGB	Detection	Artificial Synthesis
HiXray	**2021**	**8**	**45,364**	✓	**102,928**	✓	RGB	Detection	Airport

HiXray Dataset

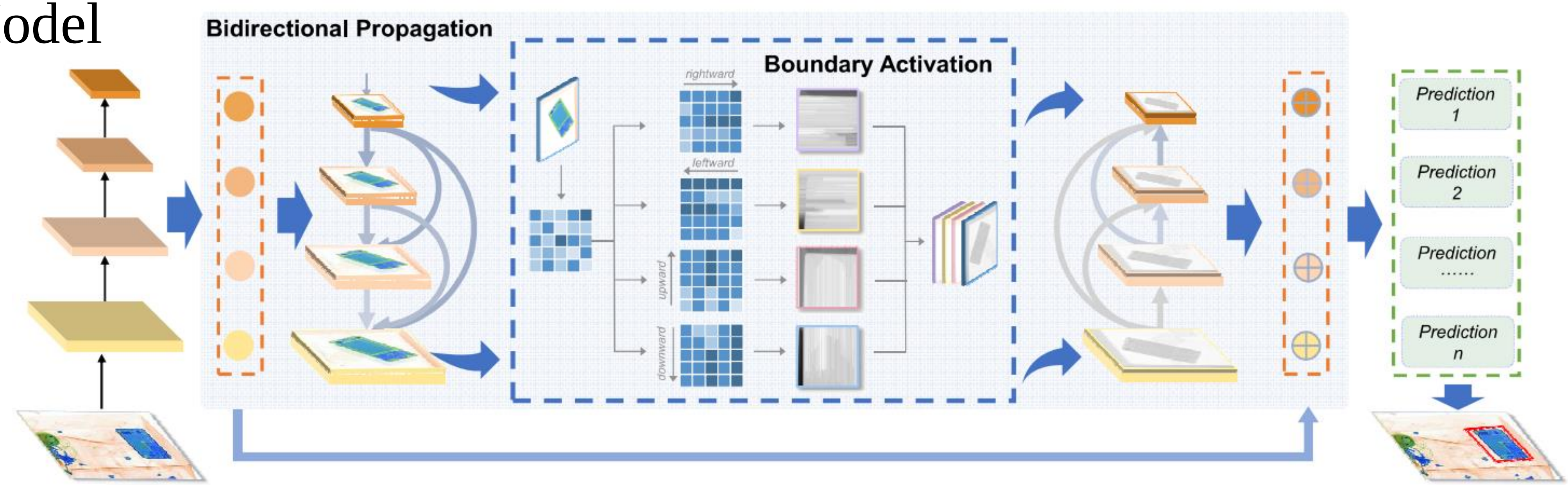
Category	PO1	PO2	WA	LA	MP	TA	CO	NL	Total
Training	9,919	6,216	2,471	8,046	43,204	3,921	7,969	706	82,452
Testing	2,502	1,572	621	1,996	10,631	997	1,980	177	20,476
Total	12,421	7,788	3,092	10,042	53,835	4,918	9,949	883	102,928



N_i	1	2	3	4	5	6	7	8	9	10
Training	12,726	10,905	6,860	3,286	1,521	602	254	91	35	11
Testing	3,227	2,722	1,705	810	354	145	54	41	8	2
Total	15,953	13,627	8,565	4,096	1,875	747	308	132	43	13



LIM Model



Algorithm 1: The Procedure of LIM.

Input: The feature map set
$\mathbf{F} = \{\mathcal{F}^1(\mathbf{x}), \cdots, \mathcal{F}^L(\mathbf{x})\}$.

Output: The refined feature map set
$\mathbf{C} = \{\mathbf{C}^1, \cdots, \mathbf{C}^L\}$.

for *all* $l = 1, 2, \ldots, L$ **do**
- Compute $\mathbf{A}^l$ based on Eq. (1);
- **for** *direction in fourDirections* **do**
  - **if** *direction is horizontal* **then**
    - // Avoid column loop for faster speed;
    - Rotate feature map;
  - **for** *row* $\leftarrow$ *1 to heightOfMap* **do**
    - Compute $\mathbf{B}^l_{ijc}$ based on Eq. (4);
- Generate $\mathbf{B}^l$ by concatenating all $\mathbf{B}^l_{ijc}$;
- Compute $\mathbf{C}^l_{\mathrm{t}}$ based on Eq. (2);
- Compute $\mathbf{C}^l$ based on Eq. (3);

Obtain the feature map set $\mathbf{C}$=$\{\mathbf{C}^1, \cdots, \mathbf{C}^L\}$.

$$\mathbf{A}^l = \mathcal{V}\left(\mathcal{F}^l(\mathbf{x})\right) + \sum_{m=1}^{L-l} \mathcal{U}^m\left(\mathbf{A}^{l+m}\right), \tag{1}$$

$$\mathbf{C}^l_{\mathrm{t}} = \mathcal{V}\left(\mathbf{B}^l\right) + \sum_{m=1}^{l-1} \mathcal{D}^m\left(\mathbf{C}^{l-m}_{\mathrm{t}}\right), \tag{2}$$

$$\mathbf{C}^l = \mathbf{C}^l_{\mathrm{t}} + \mathcal{F}^l(\mathbf{x}), \tag{3}$$

$$\mathbf{B}^l_{ijc} = \begin{cases} \mathbf{A}^l_{iWc} & \text{if } j = W, \\ \max\left\{\mathbf{A}^l_{ijc}, \mathbf{A}^l_{i(j+1)c}, \ldots, \mathbf{A}^l_{iWc}\right\} & \text{otherwise,} \end{cases} \tag{4}$$

Experiments

Comparing with detection methods:

Method	HiXray Dataset (**Ours**)									OPIXray Dataset [40]					
	AVG	PO1	PO2	WA	LA	MP	TA	CO	NL	**AVG**	FO	ST	SC	UT	MU
SSD [20]	71.4	87.3	81.0	83.0	97.6	93.5	92.2	36.1	0.01	70.9	76.9	35.0	93.4	65.9	83.3
SSD+DOAM [40]	72.1	88.6	82.9	83.6	97.5	94.1	92.1	38.2	0.01	74.0	81.4	41.5	95.1	68.2	**83.8**
SSD+LIM	**73.1**	**89.1**	**84.3**	**84.0**	**97.7**	**94.5**	**92.4**	**42.3**	**0.1**	**74.6**	**81.4**	**42.4**	**95.9**	**71.2**	82.1
FCOS [35]	75.7	88.6	86.4	86.8	89.9	88.9	88.9	63.0	13.3	82.0	86.4	68.5	90.2	78.4	86.6
FCOS+DOAM [40]	76.2	88.6	87.5	87.8	89.9	89.7	88.8	63.5	12.7	82.4	86.5	68.6	90.2	78.8	**87.7**
FCOS+LIM	**77.3**	**88.9**	**88.2**	**88.3**	**90.0**	**89.8**	**89.2**	**69.8**	**14.4**	**83.1**	**86.6**	**71.9**	**90.3**	**79.9**	86.8
YOLOv5 [14]	81.7	95.5	94.5	92.8	97.9	98.0	94.9	63.7	16.3	87.8	93.4	67.9	98.1	85.4	94.1
YOLOv5+DOAM [40]	82.2	95.9	94.7	93.7	98.1	98.1	95.8	65.0	16.1	88.0	93.3	69.3	97.9	84.4	**95.0**
YOLOv5+LIM	**83.2**	**96.1**	**95.1**	**93.9**	**98.2**	**98.3**	**96.4**	**65.8**	**21.3**	**90.6**	**94.8**	**77.6**	**98.2**	**88.9**	93.8

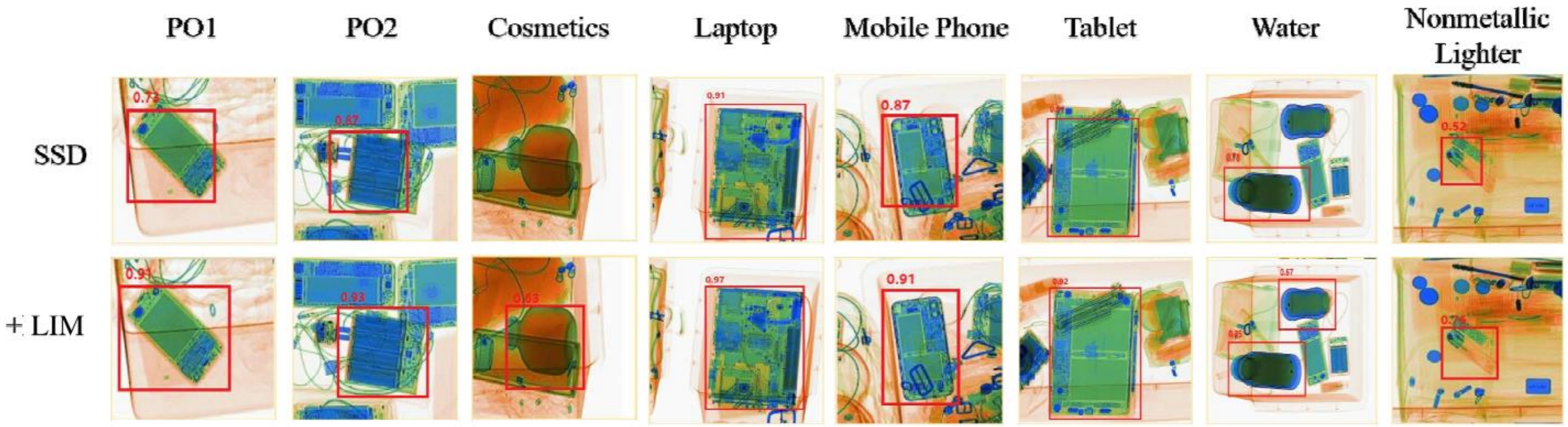
Comparing with Pyramid Networks:

Method	**AVG**	PO1	PO2	WA	LA	MP	TA	CO	NL
SSD [20]	71.4	87.3	81.0	83.0	97.6	93.5	92.2	36.1	0.01
+FPN [17]	72.0	87.4	81.5	83.2	97.9	93.9	92.2	40.3	0.02
+PANet [39]	72.0	88.3	83.2	82.8	**97.9**	93.8	**92.6**	37.3	0.01
+LIM	**73.1**	**89.1**	**84.3**	**84.0**	97.7	**94.5**	92.4	**42.3**	**0.1**

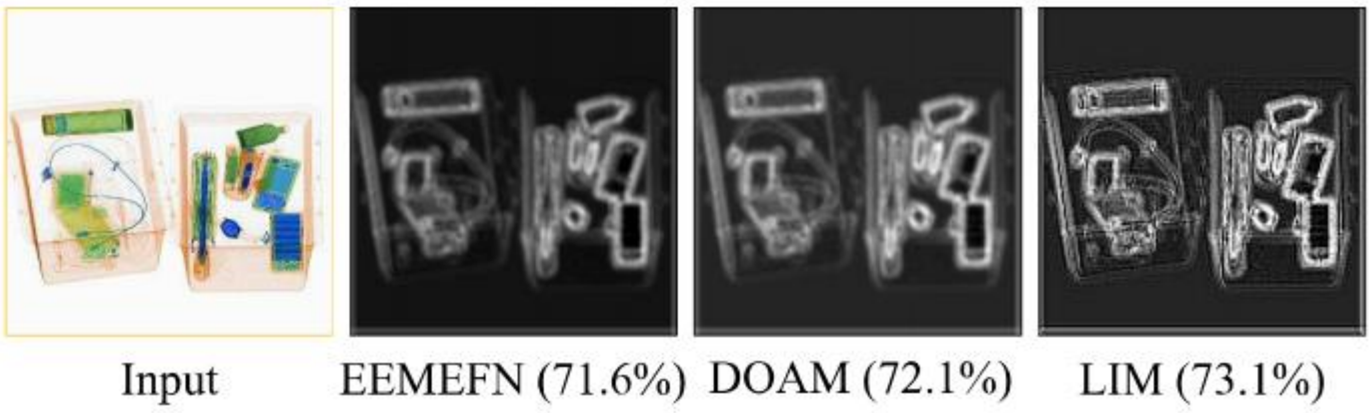
Ablation Studies:

Method	**AVG**	PO1	PO2	WA	LA	MP	TA	CO	NL
SSD [20]	71.4	87.3	81.0	83.0	97.6	93.5	92.2	36.1	0.01
+SP	72.1	87.9	82.3	83.8	97.9	92.4	92.6	38.8	0.63
+BP	72.6	88.1	83.4	83.9	**97.8**	93.8	**92.8**	40.3	0.03
+BP+BA	**73.1**	**89.1**	**84.3**	**84.1**	97.7	**94.5**	92.4	**42.3**	**0.1**

Visualization



Visualization of SSD and SSD+LIM



Visualization of the boundary aggregation

Thank you for listening!

陶仁帅
软件开发环境国家重点实验室 博士生
个人主页：rstao95.github.io

刘祥龙
教授、博士生导师、国家优青获得者
个人主页：sites.nlsde.buaa.edu.cn/~xlliu

For more information, welcome to click on my home page!