

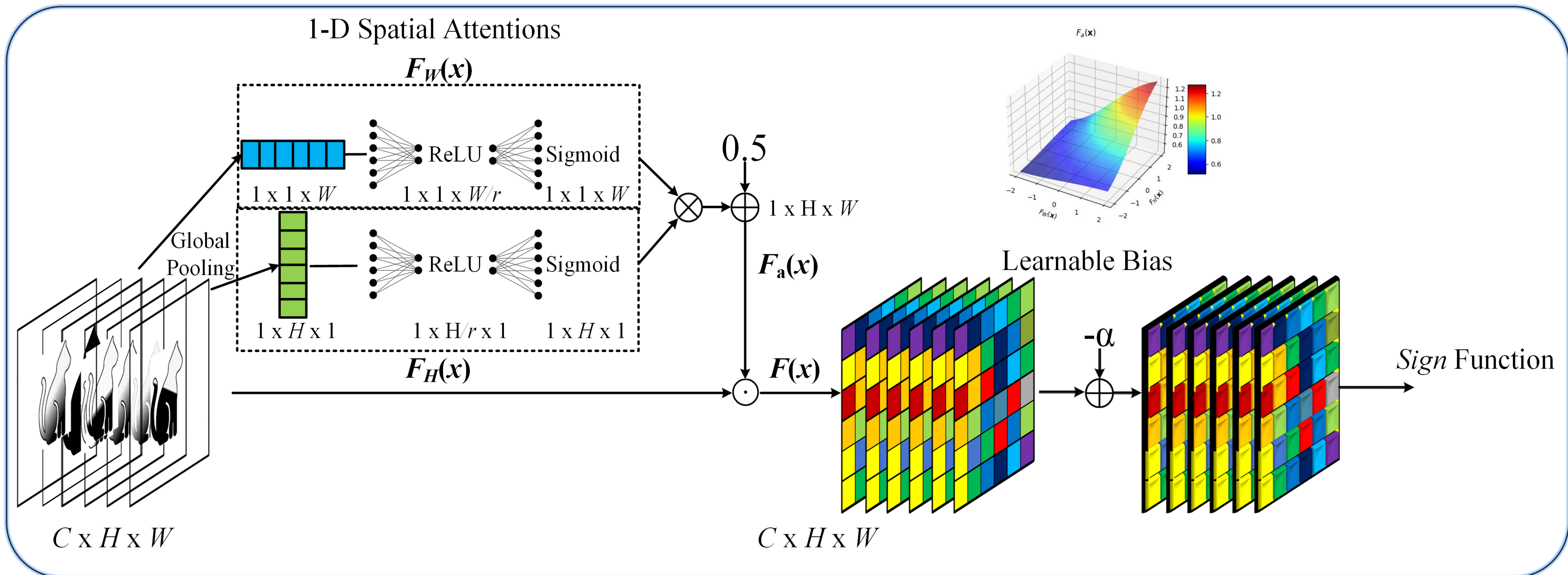
1-D SPATIAL ATTENTION IN BINARIZED CONVOLUTIONAL NEURAL NETWORKS

HyunJin Kim*, Jungwoo Shin*† Wansoo Kim* Alberto A. Del Barrio‡

* School of EEE, Dankook University, Republic of Korea

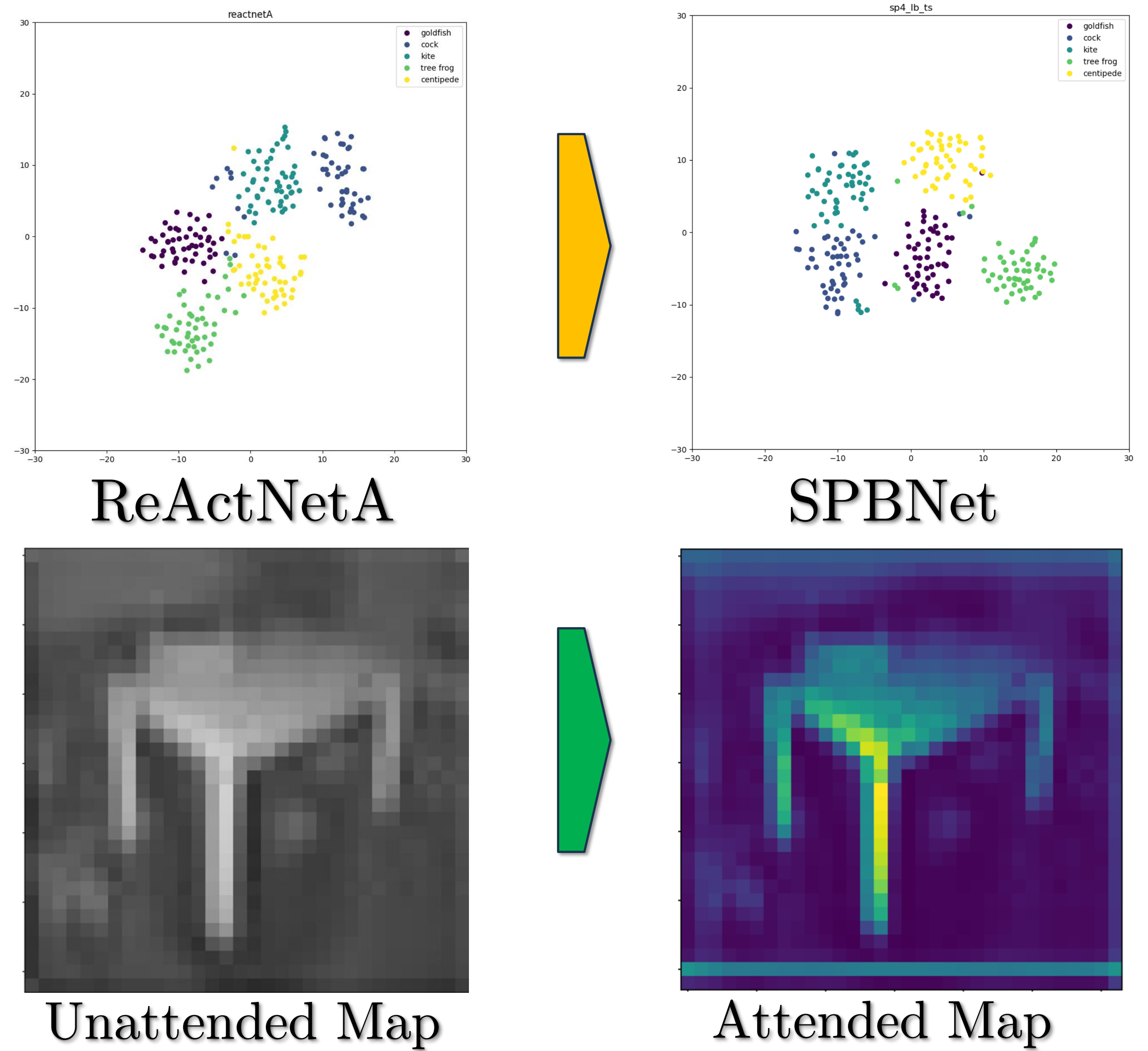
† VOOM AI LAB, LINE Plus Corporation

‡ Dept. of Computer Architecture and System Eng., Complutense University of Madrid, Madrid, ES



Contributions

- We propose SPBNet: 1-D spatial attention blocks for BCNNs.
- The proposed attention block has low-cost 1-D height-wise and 1-D width-wise convolutions, It has the attention bias to adjust the effects of attended features in $\times 0.5 - \times 1.5$.
- SPBNet shows that the biased 1-D spatial attention blocks can produce enhanced performance on both CIFAR-100 and ImageNet datasets by 2.7%@Top-1 and 1.5%@Top-1, respectively.



Results

- $sp_{r_{sp}}$: biased spatial attention with reduction ratio r_{sp} .
- ch_{r_c} : biased channel attention with reduction ratio r_c .
- lb : learnable bias for each channel.
- Without teacher-student training, the proposed structure shows comparable performance to the baseline model using teacher-student training.
- The cases only using the proposed 1-D spatial attention have good performance.

STRUCTURE	FLOPs ($\times 10^7$)	BOPS ($\times 10^9$)	MEM (MBITS)	OPS ($\times 10^7$)	TOP-1 (%)
RESNET18	57.5	-	359	57.5	75.6
XNOR-NET	1.94	1.09	18.6	3.65	71.5
REACTNET18	2.02	1.09	19.3	3.73	70.8
OUR EVALUATIONS					
ch_{16-lb}	2.17	1.09	23.8	3.88	72.6
lb	2.03	1.09	19.2	3.74	72.5
$sp_2, nobias-lb$	2.20	1.09	19.6	3.91	71.2
$sp_4, nobias-lb$	2.34	1.09	23.9	4.05	70.2
$ch_{16, nobias-lb}$	2.34	1.09	23.9	4.05	70.2
$sp_4-ch_{16}, residual-lb$	2.34	1.09	23.9	4.05	70.2
sp_2-ch_{16-lb}	2.34	1.09	24.0	4.05	73.0
sp_4-ch_{16-lb}	2.34	1.09	23.9	4.05	73.2
ch_{16-sp_2-lb}	2.34	1.09	24.0	4.05	72.4
ch_{16-sp_4-lb}	2.34	1.09	23.9	4.05	72.7
sp_2-lb	2.20	1.09	19.6	3.91	73.4
sp_4-lb	2.20	1.09	19.4	3.91	73.5

CIFAR-100

STRUCTURE	FLOPs ($\times 10^8$)	BOPS ($\times 10^9$)	MEM (MBITS)	OPS ($\times 10^8$)	TOP-1 (%)
RESNET18	37.5	-	374	37.5	69.8
XNOR-NET	2.88	3.35	33.6	3.40	51.2
REACTNETA	0.58	9.63	31.0	2.08	69.4
OUR EVALUATIONS					
sp_4-ch_{16-lb}	0.84	9.63	54.2	2.35	70.9
ch_{16-sp_4-lb}	0.84	9.63	54.2	2.35	70.8
sp_2-lb (SCRATCH)	0.73	9.63	32.8	2.23	69.3
sp_4-lb (SCRATCH)	0.73	9.63	32.8	2.23	69.4
sp_4-lb	0.73	9.63	32.8	2.23	70.7
MODEL	TOP-1(%)	MODEL	TOP-1(%)		
XNOR-NET	51.2	BI-REALNET	56.4		
CI-BCNN	56.7	XNOR-NET++	57.1		
MELIUSNETA	64.4	REAL-TO-BIN	65.4		
REACTNET18	65.9	REACTNETA	69.4		
HIGH-CAPACITY	70.0	SA-BNN	61.7		
RB-NETX2	70.1	POKEBNN-1.0X	73.4		
sp_4-ch_{16-lb}	70.9	sp_4-lb	70.7		

ImageNet