

## 1 A Implementation Details

2 We have developed a library based on PyTorch, which includes a native CUDA implementation.  
3 In order to promote transparency and reproducibility, we will provide the source code for our  
4 implementation on GitHub. However, to comply with the double-blind review process, we have  
5 omitted the link to the GitHub repository in this submission. Instead, you can find the complete  
6 source code within the provided zip file for this paper.

## 7 B Extended Tables: Experimental Results

8 In this section, we present extended tables that include additional experimental results for the UCI  
9 and MNIST MLPerf Tiny datasets. The tables provide average accuracies, standard deviations, and  
10 training times for the selected hyperparameters. All training times were measured using an Nvidia  
11 GTX-1080ti GPU.

12 **Hyperparameters**  $n$  refers to the addressing length of RAM neurons.  $K$  is the number of hash  
13 functions used.  $L$  is the filter length.  $H$  is the number of hidden neurons ( $N * C$  in the case for  
14 WNNs, where  $N$  is the number of RAMs within a discriminator and  $C$  is the number of classes),  $p$  is  
15 the dropout probability, and  $\tau$  is the softmax temperature.

Table 1: Extend results for SoonWISARD, BTHOWeN, Diff Logic Net, FINN, and DeepShift on the Ecoli Dataset.

Ecoli	Size	Acc.	Std	Time Train	Hyperparameters
SoonWiSARD	0.74KiB	81.4%	5.0%	4s	$n=4, K=2, L=9, H=84, p=0.5, \tau=1$
BTHOWeN	0.87KiB	87.5%	-	-	$n=10, K=2, L=128, H=10$
Diff Logic Net	0.87KiB	79.3%	3.6%	5s	$H=1792, p=0.0, \tau=1/0.1$
FINN	0.88KiB	68.9%	7.0%	3s	$H=447, p=0.2$
DeepShift	0.88KiB	43.6%	1.4%	7s	$H=64, p=0.2$

Table 2: Extend results for SoonWISARD, BTHOWeN, Diff Logic Net, FINN, and DeepShift on the Iris Dataset.

Iris	Size	Acc.	Std	Time Train	Hyperparameters
SoonWiSARD	0.23KiB	98.3%	6.0%	1s	$n=4, K=2, L=9, H=72, p=0.5, \tau=1$
BTHOWeN	0.28KiB	98.0%	-	-	$n=2, K=1, L=128, H=6$
Diff Logic Net	0.28KiB	98.3%	3.1%	1s	$H=575, p=0.0, \tau=1/0.1$
FINN	0.29KiB	69.2%	16.9%	1s	$H=287, p=0.2$
DeepShift	0.30KiB	33.3%	17.2%	1s	$H=32, p=0.2$

Table 3: Extend results for SoonWISARD, BTHOWeN, Diff Logic Net, FINN, and DeepShift on the Letter Dataset.

Letter	Size	Acc.	Std	Time Train	Hyperparameters
SoonWiSARD	76.23KiB	96.0%	0.8%	72s	$n=11, K=2, L=16, H=1496, p=0.5, \tau=1$
BTHOWeN	78.00KiB	90.0%	-	-	$n=20, K=4, L=2048, H=12$
Diff Logic Net	78.00KiB	90.9%	0.7%	85s	$H=159744, p=0.0, \tau=1/0.1$
FINN	78.00KiB	4.79%	0.3%	96s	$H=14859, p=0.2$
DeepShift	78.00KiB	19.2%	1.6%	68s	$H=2636, p=0.2$

Table 4: Extend results for SoonWiSARD, BTHOWeN, Diff Logic Net, FINN, and DeepShift on the SatImage Dataset.

SatImage	Size	Acc.	Std	Time Train	Hyperparameters
SoonWiSARD	8.54KiB	92.5%	0.8%	20s	n=8, K=2, L=9, H=1295, p=0.5, $\tau=1$
BTHOWeN	9.00KiB	88.0%	-	-	n=12, K=4, L=512, H=24
Diff Logic Net	9.25KiB	90.2%	1.1%	19s	H=18432, p=0.0, $\tau=1/0.1$
FINN	9.00KiB	30.8%	0.0%	32s	H=1714, p=0.2
DeepShift	9.50KiB	48.0%	1.8%	23s	H=303, p=0.2

Table 5: Extend results for SoonWiSARD, BTHOWeN, Diff Logic Net, FINN, and DeepShift on the Vehicle Dataset.

Vehicle	Size	Acc.	Std	Time Train	Hyperparameters
SoonWiSARD	2.25KiB	78.3%	4.3%	3s	n=6, K=2, L=32, H=144, p=0.5, $\tau=1$
BTHOWeN	2.25KiB	76.2%	-	-	n=16, K=3, L=256, H=18
Diff Logic Net	2.25KiB	73.5%	4.0%	3s	H=4608, p=0.0, $\tau=1/0.1$
FINN	2.30KiB	27.2%	2.8%	5s	H=801, p=0.2
DeepShift	2.50KiB	28.3%	0.7%	4s	H=128, p=0.2

Table 6: Extend results for SoonWiSARD, BTHOWeN, Diff Logic Net, FINN, and DeepShift on the Vowel Dataset.

Vowel	Size	Acc.	Std	Time Train	Hyperparameters
SoonWiSARD	3.43KiB	94.0%	2.9%	3s	n=6, K=4, L=64, H=44, p=0.5, $\tau=1$
BTHOWeN	3.44KiB	90.0%	-	-	n=15, K=4, L=256, H=10
Diff Logic Net	3.44KiB	74.9%	4.1%	3s	H=7045, p=0.0, $\tau=1/0.1$
FINN	3.50KiB	17.7%	3.0%	6s	H=1280, p=0.2
DeepShift	3.50KiB	8.4%	1.8%	4s	H=203, p=0.2

Table 7: Extend results for SoonWiSARD, BTHOWeN, Diff Logic Net, FINN, and DeepShift on the Wine Dataset.

Wine	Size	Acc.	Std	Time Train	Hyperparameters
SoonWiSARD	0.23KiB	93.3%	4.2%	1s	n=2, K=2, L=4, H=36, p=0.5, $\tau=1$
BTHOWeN	0.42KiB	98.3%	-	-	n=13, K=3, L=128, H=9
Diff Logic Net	0.42KiB	86.0%	4.7%	1s	H=864, p=0.0, $\tau=1/0.1$
FINN	0.42KiB	14.0%	2.0%	1s	H=203, p=0.2
DeepShift	0.42KiB	27.3%	15.6%	1s	H=20, p=0.2

Table 8: Extend results for SoonWiSARD, BTHOWeN, Diff Logic Net, FINN, and DeepShift on the MNIST Dataset.

MNIST	Size	Acc.	Std	Time Train	Hyperparameters
SoonWiSARD (small)	23KiB	97.2%	0.3%	8m	n=14, K=2, L=16 H=11776, p=0.5, $\tau=1/0.1$
SoonWiSARD (medium)	98KiB	98.0%	0.2%	8m	n=14, K=2, L=16 H=50176, p=0.5, $\tau=1/0.03$
SoonWiSARD (large)	186KiB	98.2%	0.2%	8m	n=14, K=2, L=16 H=95232, p=0.5, $\tau=1/0.03$
BTHOWeN (small)	35KiB	92.8%	0.2%	-	n=28, K=2, L=512, H=560
BTHOWeN (medium)	70KiB	93.4%	0.2%	-	n=28, K=2, L=1024, H=560
BTHOWeN (large)	210KiB	94.3%	0.2%	-	n=28, K=2, L=2048, H=840
Diff Logic Net (small)	23KiB	94.1%	0.5%	7m	H=45000, p=0.0, $\tau=1/0.1$
Diff Logic Net (medium)	98KiB	95.2%	0.3%	12m	H=20000, p=0.0, $\tau=1/0.03$
Diff Logic Net (large)	188KiB	95.3%	0.4%	20m	H=384000, p=0.0, $\tau=1/0.03$
FINN (small)	23KiB	94.6%	0.2%	12m	H=230, p=0.2
FINN (medium)	98KiB	97.0%	0.2%	12m	H=1010, p=0.2
FINN (large)	188KiB	97.7%	0.2%	12m	H=1940, p=0.2
DeepShift (small)	23KiB	96.8%	0.2%	8m	H=45, p=0.2
DeepShift (medium)	98KiB	98.1%	0.2%	8m	H=200, p=0.2
DeepShift (large)	188KiB	98.4%	0.2%	8m	H=385, p=0.2

Table 9: Extend results for SoonWiSARD, BTHOWeN, Diff Logic Net, FINN, and DeepShift on the CIFAR-10 Dataset.

CIFAR-10	Size	Acc.	Std	Time Train	Hyperparameters
SoonWiSARD (small)	20KiB	49.7%	1.3%	8m	n=6, K=2, L=4, H=40960, p=0.5, $\tau=1/0.03$
SoonWiSARD (medium)	240KiB	55.6%	1.1%	28m	n=6, K=2, L=4, H=491520, p=0.5, $\tau=1/0.01$
SoonWiSARD (large)	620KiB	57.3%	0.8%	76m	n=6, K=2, L=4, H=1269760, p=0.5, $\tau=1/0.01$
BTHOWeN (small)	32KiB	14.06%	1.2%	-	n=30, K=2, L=128, H=2050
BTHOWeN (medium)	192KiB	36.47%	0.6%	-	n=10, K=2, L=256, H=6140
BTHOWeN (large)	384KiB	37.72%	0.4%	-	n=10, K=2, L=512, H=6140
Diff Logic Net (small)	24KiB	48.2%	0.0%	8m	H=49000, p=0.0, $\tau=1/0.03$
Diff Logic Net (medium)	250KiB	50.6%	0.0%	32m	H=512000, p=0.0, $\tau=1/0.01$
Diff Logic Net (large)	625KiB	52.3%	0.0%	84m	H=1280000, p=0.0, $\tau=1/0.01$
FINN (small)	24KiB	40.0%	0.8%	16m	H=64, p=0.2
FINN (medium)	250KiB	46.5%	0.7%	16m	H=665, p=0.2
FINN (large)	625KiB	48.0%	0.6%	16m	H=1660, p=0.2
DeepShift (small)	24KiB	40.3%	1.9%	12m	H=13, p=0.2
DeepShift (medium)	250KiB	53.0%	1.2%	12m	H=133, p=0.2
DeepShift (large)	622KiB	54.1%	0.9%	12m	H=331, p=0.2

Table 10: Extend results for SoonWiSARD, BTHOWeN, Diff Logic Net, FINN, and DeepShift on the KWS Dataset.

KWS	Size	Acc.	Std	Time Train	Hyperparameters
SoonWiSARD (small)	23KiB	58.2%	1.6%	4m	n=6, K=2, L=4, H=47104, p=0.5, $\tau=1/0.03$
SoonWiSARD (medium)	46KiB	67.7%	1.1%	4m	n=6, K=2, L=4, H=94208, p=0.5, $\tau=1/0.01$
SoonWiSARD (large)	92KiB	69.6%	0.8%	4m	n=6, K=2, L=4, H=188416, p=0.5, $\tau=1/0.01$
BTHOWeN (small)	15KiB	39.6%	0.0%	-	n=6, K=2, L=64, H=1920
BTHOWeN (medium)	53KiB	44.7%	0.0%	-	n=6, K=2, L=64, H=6784
BTHOWeN (large)	183KiB	48.3%	0.0%	-	n=10, K=2, L=512, H=2928
Diff Logic Net (small)	23KiB	57.3%	2.0%	3m	H=48000, p=0.0, $\tau=1/0.03$
Diff Logic Net (medium)	46KiB	63.3%	1.0%	4m	H=95000, p=0.0, $\tau=1/0.03$
Diff Logic Net (large)	92KiB	63.7%	0.9%	8m	H=188000, p=0.0, $\tau=1/0.03$
FINN (small)	23KiB	47.2%	1.8%	8m	H=346, p=0.2
FINN (medium)	46KiB	53.3%	2.5%	8m	H=692, p=0.2
FINN (large)	92KiB	55.9%	1.4%	8m	H=1384, p=0.2
DeepShift (small)	23KiB	18.6%	5.5%	4m	H=69, p=0.2
DeepShift (medium)	46KiB	22.2%	2.5%	4m	H=137, p=0.2
DeepShift (large)	92KiB	24.9%	4.4%	4m	H=274, p=0.2

Table 11: Extend results for SoonWiSARD, BTHOWeN, Diff Logic Net, FINN, and DeepShift on the ToyADMOS Dataset.

ToyADMOS	Size	Acc.	Std	Time Train	Hyperparameters
SoonWiSARD (small)	7KiB	88.4%		48m	n=6, K=1, L=16, H=3840, p=0.5, $\tau=1/0.03$
SoonWiSARD (medium)	15KiB	89.3%		48m	n=6, K=1, L=16, H=7680, p=0.5, $\tau=1/0.03$
SoonWiSARD (large)	30KiB	90.5%		48m	n=6, K=1, L=16, H=15360, p=0.5, $\tau=1/0.03$
BTHOWeN (small)	8KiB	72.3%		-	n=6, K=2, L=64, H=1700
BTHOWeN (medium)	13KiB	73.3%		-	n=6, K=2, L=64, H=2780
BTHOWeN (large)	80KiB	74.1%		-	n=10, K=2, L=1024, H=640
Diff Logic Net (small)	7KiB	88.3%		52m	H=15000, p=0.0, $\tau=1/0.03$
Diff Logic Net (medium)	15KiB	89.2%		52m	H=31000, p=0.0, $\tau=1/0.03$
Diff Logic Net (large)	30KiB	89.4%		56m	H=62000, p=0.0, $\tau=1/0.03$
FINN (small)	7KiB	84.8%		64m	H=90, p=0.2
FINN (medium)	15KiB	85.9%		64m	H=190, p=0.2
FINN (large)	30KiB	86.6%		64m	H=380, p=0.2
DeepShift (small)	7KiB	57.8%		88m	H=18, p=0.2
DeepShift (medium)	15KiB	57.8%		88m	H=38, p=0.2
DeepShift (large)	30KiB	57.9%		88m	H=76, p=0.2

Table 12: Extend results for SoonWiSARD, BTHOWeN, Diff Logic Net, FINN, and DeepShift on the VWW Dataset.

VWW	Size	Acc.	Std	Time Train	Hyperparameters
SoonWiSARD (small)	12KiB	57.9%	1.5%	12m	n=6, K=1, L=16, H=6144, p=0.5, $\tau=1/0.03$
SoonWiSARD (medium)	120KiB	59.8%	0.9%	16m	n=6, K=1, L=16, H=61440, p=0.5, $\tau=1/0.03$
SoonWiSARD (large)	240KiB	60.6%	1.1%	20m	n=6, K=1, L=16, H=122880, p=0.5, $\tau=1/0.03$
BTHOWeN (small)	12KiB	57.3%	0.0%	-	n=6, K=2, L=64, H=1536
BTHOWeN (medium)	120KiB	58.4%	0.0%	-	n=10, K=2, L=1024, H=960
BTHOWeN (large)	240KiB	59.3%	0.0%	-	n=10, K=2, L=2048, H=960
Diff Logic Net (small)	20KiB	52.6%	1.6%	20m	H=40960, p=0.0, $\tau=1/0.03$
Diff Logic Net (medium)	120KiB	55.4%	1.4%	32m	H=245760, p=0.0, $\tau=1/0.03$
Diff Logic Net (large)	250KiB	55.7%	1.3%	56m	H=512000, p=0.0, $\tau=1/0.03$
FINN (small)	20KiB	51.7%	1.5%	24m	H=17, p=0.2
FINN (medium)	120KiB	52.1%	1.1%	24m	H=106, p=0.2
FINN (large)	250KiB	52.3%	1.2%	24m	H=222, p=0.2
DeepShift (small)	20KiB	52.9%	1.2%	16m	H=3, p=0.2
DeepShift (medium)	120KiB	53.8%	0.9%	16m	H=21, p=0.2
DeepShift (large)	250KiB	54.6%	0.8%	16m	H=44, p=0.2