Within-Camera Multilayer Perceptron DVS Denoising

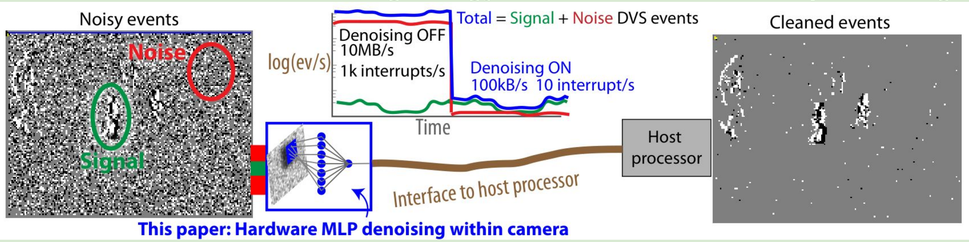
June 18, 2023 EVNTVIS https://tub-rip.github.io/eventvision2023/

Antonio Rios-Navarro[2,1], Shasha Guo[3,1], (Aba) G. Abarajithan[5], (Vijay) K. Vijayakumar[5], Alejandro Linares-Barranco[2], Thea Aarrestad[4], Ryan Kastner[5,1], and Tobi Delbruck[1]

[1] *Sensors Group, Institute of Neuroinformatics, Univ. of Zurich and ETH Zurich, Switzerland,* [2] *Robotic and Tech of Computers Group, SCORE lab, ETSI-EPS, Univ. of Seville, Spain*

[3] *College of Electronic Engineering, National University of Defense Technology (NUDT), China,* [4] *Inst. of Particle Physics and Astrophysics, ETH Zurich, Switzerland,* [5] *Univ. of California, San Diego, USA*

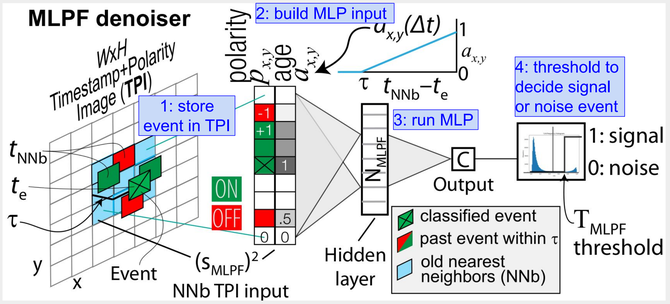
In-camera event denoising reduces the data rate of event cameras by filtering out noise at the source, **reducing energy for communication and processing**.

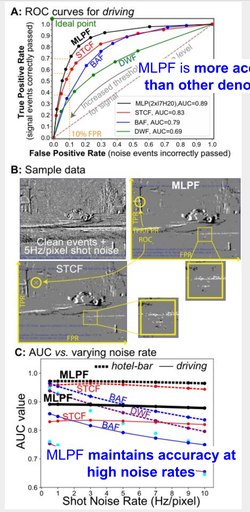


This paper proposes **the first digital logic implementations of a novel hardware neural network denoiser** and quantifies their resource cost, power, and latency.

Method

A lightweight **multilayer perceptron denoising filter (MLPF**) **provides state-of-the-art low-cost denoising accuracy**. It processes a small neighborhood of pixels from the timestamp image around each event to discriminate signal and noise events.





Hardware Implementation Results

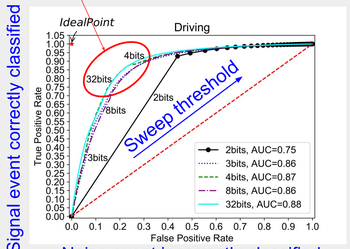
The hardware MLP **quantizes the weights and hidden unit activations to 4 bits** and has about **1k weights with about 40% sparsity**.

Model trained with QKERAS

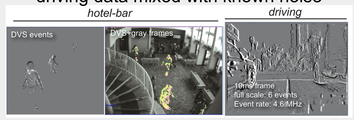
Table 1. Hardware MLPF specifications.

$W \times H$	DVS width×height	346×260[a]
S^2_{MLPF}	input patch	7^2px
N_{MLPF}	# hidden units	10
τ	age window	64ms[b]
	Quantization bits	(weights/activations)[c]
	Input units	4+1/4+1
	Hidden units	4+1/4
	Output unit	4+1/15+1
	Threshold T_{MLPF}	15+1
	Accumulators	16.6[d]
	Network	
	Num. weights+bias	1001
	Sparsity	40%[e]
	Accuracy (AUC)	0.87[f]

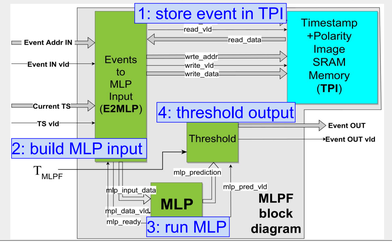
The 4-bit ***Area-Under-Curve*** **(AUC)** Receiver Operating Characteristic accuracy is **nearly indistinguishable from that of the floating point 32-bit network**.

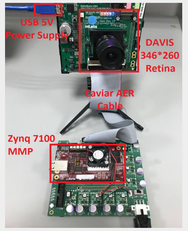


Measure accuracy using clean surveillance and driving data mixed with known noise

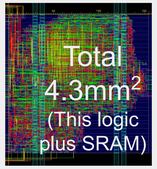


The hls4ml FPGA MLPF p**rocesses each event in only 10 clock cycles**. In FPGA, it uses 3.5k flip flops and 11.5k LUTs.





Our RTL **ASIC implementation** in 65nm digital technology for a 346x260 pixel camera



4 nJ per event at event rates up to 25MHz. 4 mW at 1 MHz

MLPF fits into small FPGA or small ASIC area and has 40 ns latency. Memory for the 2D timestamp image dominates cost.

Table 2. FOMs of the hardware MLPFs for Table 1 parameters.

	12 y old FPGA Xilinx Zynq XC7Z100		Modern $400 FPGA Xilinx Zynq U+ ZU3CG		Cheap digital process 65nm ASIC. Vdd=1.08V	
Max. Clock Freq.	100 MHz		236 MHz		833 MHz	
Latency	cycles	ns	cycles	ns	cycles	ns
E2MLP	7	70	7	∼30	30	36
MLP	3	30	3	∼13	3	3.6
Total	10	100	10	∼43	33	40

	Xlinx Zynq xc7z100			Xlinx Zynq U+ ZU3CG			65nm ASIC	
	LUT	FF	BRAM	LUT	FF	BRAM	Logic area	SRAM area
Resources	18.4k	3.9k	400[a]	24.6k	3.8k	400[a]	0.022mm²	4.3[b]mm²
Resource %	6%	0.72%	26%	34%	2%	92%	–	–
Power	Not relevant						40mW[c]	
Energy/event	Not relevant						4nJ[d]	

Comparison with SOA

Hardware denoisers generally seek to minimize memory, which leads to poor discrimination accuracy at high noise rates. **The MLPF has by far the best AUC accuracy. Its maximum denoising rate of 25 MHz is suitable for recent event cameras**.

Denoiser	*hotel-bar* AUC[a]	*driving* AUC[a]	Mem(#)[b]	Max. Event Rate[c] MHz
MLPF FPGA	**0.96**	**0.87**	$W \times H$	23
MLPF ASIC	"	"	$W \times H$	25
BAF [17]	0.89	0.79	$W \times H$	3.6
ONF [16]	0.01[e]	0.01[e]	$2 \times (W + H)$	3
HashHeat [18]	0.67	0.56[f]	**128**	100
IIRs [19]	NA	NA	$W \times H$[g]	**385**
LDSI [20]	NA	NA	$W \times H$	3?[d]

Our accurate and quick denoiser can be easily integrated into an event camera using an FPGA or as an ASIC in the camera chip or same package.

It could open new areas of always-on event camera application under scavenged and battery power.

This work was supported by the Swiss SNF projects SCIDVS, VIPS and PZ00P2 201594, and by the Spanish grant MINDROB.
We thank Dr. Min Liu and Prof. Shih-Chii Liu (INI), Prof. Mingu Kang (UCSD), and anonymous reviewers for their support and suggestions to improve the paper.

https://github.com/SensorsINI/dnd_hls