Spiking neural network chip combines low latency and energy consumption with high inference accuracy

June 1 2021

Task	Method	Accuracy	$Energy_{au}/s$	Energy ratio	Error ratio	Efficiency
ECG-qtdb	Adaptive SRNN	*85.9%*	325.7	1x	1x	1x
	LIF SRNN	75.5%	179.9	.55 x	1.67x	.91x
	RELU SRNN	86.4%	5784.6	17.8x	.93x	16.5x
	LSTM	78.9%	20422.8	62.7x	1.43x	89.6x
	GRU	77.3%	15400	47.2x	1.54x	72.7x
	Vanilla RNN	74.8%	9597.6	29.5x	1.71x	50.5x
	Bidirectional $LSTM_{290}$	80.76%	563580	1729.9x	1.31x	2266.2x
SMNIST	Adaptive SRNN	*98.7%*	8250.3	1x	1x	1x
	RELU SRNN	98.99%	487623.8	59.1x	.74x	43.4x
PSMNIST	Adaptive SRNN	*94.32%*	7775.1	1x	1x	1x
	RELU SRNN	93.47%	487623.8	62.7x	1.1x	69.0x
SHD	Adaptive SRNN	*87.81%*	3515.7	1x	1x	1x
	RELU SRNN	88.93%	442097.2	125.8x	.91x	114.5x
	Bidirectional LSTM	87.2%	3468215.52	986.5x	1.05x	1035.8x
	CNN[27]	92.4%	–	–x	–x	–x
SSC	Adaptive SRNN	74.18%	10154.1	1x	1x	1x
	RELU SRNN	74.36%	2373918.	234.0x	.99x	231.7x
	LSTM[27]	73.1%	–	–x	–x	–x
	CNN[27]	77%	–	–x	–x	–x
SoLi	Adaptive SRNN	79.8%	13,804	1x	1x	1x
	Vanilla RNN	63.6%	4101324.8	297.1x	1.80x	524.8x
	GRU	79.20%	6580531.2	476.7x	1.03x	491.1x
	LSTM	79.99%	8360972.8	605.7x	0.99x	599.6x
	RELU SRNN	79.8%	3283950.9	237.9x	1.x	237.9x
TIMIT	Adaptive SRNN	66.13%	10626.8	1x	1x	1x
	RELU SRNN	–%	11861974	175.2x	–x	–x
GSC	Adaptive SRNN	92.12%	4120.3	1x	1x	1x
	RELU SRNN	–%	11861974	167.5x	–x	–x
	CNN[27]	92.4%	80600	19.6x	0.96x	18.8x

Imec's adaptive neuron-based SNN ('Adaptive SRNN') was evaluated against six other neural networks—using eight different data sets including Google's radar (SoLi) and Google's speech datasets. Credit: IMEC

In April 2020, imec introduced the world's first chip to process radar signals using a spiking recurrent neural network (SNN). Its flagship use-case? The creation of a smart, low-power multi-sensor perception system for drones that identifies obstacles in a matter of milliseconds.

Contrary to the artificial neural networks that are a key ingredient of today's robotics perception systems, SNNs mimic the way groups of biological neurons operate—firing electrical pulses sparsely over time, and, in the case of biological sensory neurons, only when the sensory input changes. It is an approach that comes with important benefits: at the time of the announcement, imec's SNN chip showed to consume up to a hundred times less power than traditional implementations while featuring a tenfold reduction in latency—enabling almost instantaneous decision-making.

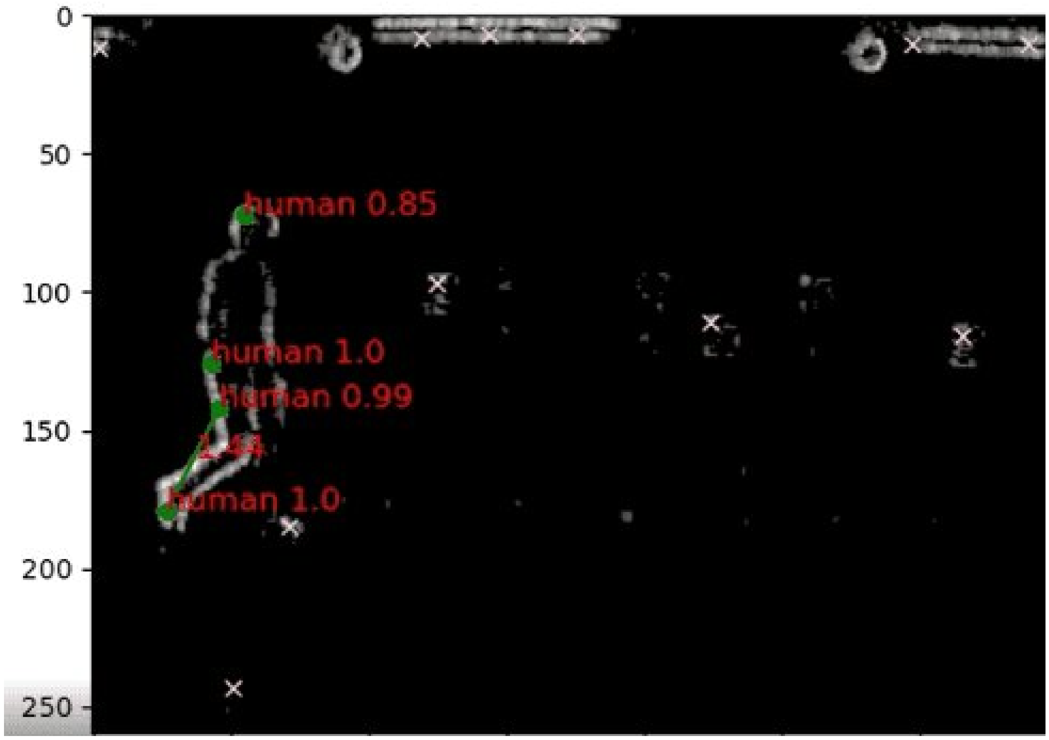
In the following article, Ilja Ocket—program manager of neuromorphic sensing at imec—provides more insights into some of imec's recent advances in this domain.

Optimizing and scaling up the original SNN chip

In the last year, imec has been optimizing and scaling up its original SNN chip—the details of which were recently published in *Frontiers in Neuroscience*—to host a variety of (IoT and autonomous robotics) use-cases. The chip builds on an entirely event-based digital architecture, and was implemented in low-cost 40nm CMOS technology. It supports event-driven processing and relies on local on-demand oscillators and a novel delay-cell to avoid the use of a global clock. Moreover, it does not exploit separate memory blocks; instead, memory and computation are co-localized in the IC area to avoid data access and energy overheads.

Imec's SNN ranks amongst the top performers in

terms of inference accuracy



Imec’s end-to-end spiking approach at work. Local feature detection forms the first layer for a more complex semantic build-up. Credit: IMEC

Meanwhile, research with the Dutch national research institute for mathematics and computer science (CWI), confirms that spiking neurons with adaptive thresholds can be trained to achieve top-notch inference accuracy. A comprehensive study conducted by imec and CWI aimed to benchmark SNNs using adaptive neurons against six other neural networks. To do so, eight different data sets were used—including Google's radar (SoLi) and Google's speech datasets. The study clearly pointed out that SNNs using neurons with adaptive thresholds can achieve a low energy consumption, but not at the expense of a decreased inference accuracy. On the contrary: for each of the major data sets used in the study, the imec SNN ranked amongst the top performers in terms of accuracy.

"SNN technology will find its way in a broad range of use-cases: from smart, self-learning Internet of Things (IoT) devices—such as wearables and brain-computer interface applications—to autonomous drones and robots. But each of those use-cases comes with its own set of requirements," says Ilja Ocket. "While spiking neural networks for IoT applications should excel at operating within a very small power budget, autonomous drones demand a low-latency SNN that allows them to avoid obstacles quickly and effectively."

"Addressing those requirements using a one-size-fits-all SNN architecture is extremely challenging. A delicate balance needs to be struck between energy consumption, latency, accuracy, cost (chip area) and scalability. For example, an SNN with the lowest possible energy consumption and latency typically results in an increased chip area—and vice versa. Finding that balance is one of imec's SNN focus areas."

Going forward: spiking all the way

Drones are being used in an increasing number of application domains. Still, to improve their level of autonomy and to have them operate in more challenging environments (such as bad weather conditions), their sensing capabilities require yet another boost. According to Ilja Ocket, an end-to-end spiking approach—based on fused neuromorphic radar and camera inputs—might offer a way out.

Ilja Ocket says, "This obviously makes for a highly energy-efficient and super low latency system. Today, however, in order to connect such cameras to the underlying AI, one still needs to translate their feed into frames—which significantly limits the efficiency gains. That is why we are investigating how the spiking concept can be implemented end-to-end: from the cameras and sensors to the AI engine. We are actually the first ones to do so, with very promising results so far. To that end, we are still looking for companies from across the drone industry—such as OEM drone builders—to join our program and experiment with this exciting technology."

More information: Jan Stuijt et al, μBrain: An Event-Driven and Fully Synthesizable Architecture for Spiking Neural Networks, *Frontiers in Neuroscience* (2021). DOI: 10.3389/fnins.2021.664208

Provided by IMEC