

FRAUNHOFER INSTITUTE FOR INTEGRATED CIRCUITS IIS

PRESS RELEASE

Energy-saving AI chip wins innovation contest

Erlangen: Organized by the German Federal Ministry of Education and Research (BMBF), the "Energy-efficient AI systems" pilot innovation contest posed the challenge: "Which team can produce a chip that detects atrial fibrillation in ECG data with at least 90 percent accuracy while consuming the least energy possible?" On March 11, 2021, Federal Minister of Education and Research Anja Karliczek presented prizes to the winners, which included the Fraunhofer Institute for Integrated Circuits IIS in collaboration with Friedrich-Alexander-Universität Erlangen-Nürnberg.

"The whole team is delighted," said the project leaders Dr. Marco Breiling from Fraunhofer IIS as well as Professor Dietmar Fey and Dr. Marc Reichenbach from Friedrich-Alexander-Universität Erlangen-Nürnberg. Breiling went on to explain that: "The prize money awarded for the follow-up project will provide further impetus for the development of neuromorphic hardware as part of our Next Generation Computing initiative. At our Erlangen and Dresden locations, we want to ensure that Germany plays a leading role in the practical applications of AI."

"Artificial intelligence still consumes too much energy, but it has enormous potential to bolster Germany's reputation as a hub for business and innovation and is increasingly finding its way into our everyday lives. With this in mind, the Federal Ministry of Education and Research is stepping in to provide funding and has called on higher education institutions and research establishments across the country to apply with their best ideas for developing an energy-saving AI chip. I'm delighted for the four winning teams selected today, who now have a unique opportunity to develop their projects further with funding of around a million euros each," explained the Federal Minister of Education and Research.

Putting the signal processing to sleep

The particularly energy-efficient processing of time-series signals is addressed in the Lo3-ML project, whose name stands for "Low-power low-memory low-cost ECG signal analysis using ML algorithms." This work involved developing a chip for AI calculations – or, more precisely, for deep learning inference – in order to analyze ECG signals with a view to determining whether a patient is healthy or suffering from atrial fibrillation.

Head of Corporate Communications

Thoralf Dietz | Phone +49 9131 776-1630 | thoralf.dietz@iis.fraunhofer.de | Fraunhofer Institute for Integrated Circuits IIS | Am Wolfsmantel 33 | 91058 Erlangen | www.iis.fraunhofer.de

PRESS RELEASE March 19, 2021 || Page 1 | 3

Editorial staff Claudia Wutz | Phone +49 9131 776-4071 | claudia.wutz@iis.fraunhofer.de | Fraunhofer Institute for Integrated Circuits IIS | www.iis.fraunhofer.de



FRAUNHOFER INSTITUTE FOR INTEGRATED CIRCUITS IIS

The key lies in putting part of the chip – specifically, the signal processing technology – "to sleep" while it isn't needed, resulting in energy savings of up to 95 percent. The circuit can also be used for other applications that involve processing time-series signals, such as for anomaly detection or predictive maintenance.

Several innovative concepts in the chip

By using special systolic arrays, the researchers were able to significantly reduce the energy required for control logic. The technology uses only the three weights -1, 0, and +1, as well as batch normalizations with powers of two, paving the way not only for a highly versatile algorithm but also for the utmost efficiency when it comes to implementing the neuron's calculations or normalization. Inside the chip is a form of nonvolatile memory known as resistive RAM (RRAM), along with ultra-low-power read/write circuits that store precisely these ternary values with a high level of efficiency. The non-volatile memory allows the signal processing to be put to sleep for long periods of time so that it consumes no energy. After all, with signal bandwidths typically in the (sub-)kHz range, data arrive at the chip much more slowly than they can be processed. These data are therefore collected in a section of the chip that is awake while the AI algorithm sleeps and then AI-processed rapidly once the algorithm wakes up – at which point all parameters are available immediately thanks to the RRAMs. This approach can achieve energy savings of up to 95 percent relative to an "always-on" system.

In addition to digital and analog circuit development and ASIC implementation, the project placed great importance on automating the design process with a view to optimizing the AI algorithm and the circuit jointly in terms of their energy consumption. These efforts also called for newly developed tools for hardware-aware training of neural networks, such as for the extreme quantization applied here. This process can also be applied to other problems and applications.

The project offers an impressive demonstration of the potential for optimization that can emerge from close collaborations between algorithm designers and digital or analog chip developers.

"Energy-efficient AI systems" pilot innovation contest

A prize also went to the Fraunhofer Institute for Industrial Mathematics ITWM and TU Kaiserslautern for a project entitled "Holistic approach to optimizing FPGA architectures for deep neural networks via AutoML – automated machine learning" (HALF).

Applications to take part in the contest were received from 27 teams from higher education institutions and research establishments. Of those, 11 teams were selected for the final round. Each of the four winning projects receives an exclusive opportunity to PRESS RELEASE March 19, 2021 || Page 2 | 3



FRAUNHOFER INSTITUTE FOR INTEGRATED CIRCUITS IIS

submit a research project to implement their idea with application partners with funding to the tune of 1 million euros.

PRESS RELEASE March 19, 2021 || Page 3 | 3

GEFÖRDERT VOM



für Bildung und Forschung

Funded by the German Federal Ministry of Education and Research

The Fraunhofer-Gesellschaft, headquartered in Germany, is the world's leading applied research organization. Its research activities are conducted by 75 institutes and research units at locations throughout Germany. The Fraunhofer-Gesellschaft employs a staff of around 29,000 people, who work with an annual research budget totaling 2.8 billion euros.

The Fraunhofer Institute for Integrated Circuits IIS, headquartered in Erlangen, Germany, conducts world-class research on microelectronic and IT system solutions and services. Today, it is the largest institute of the Fraunhofer-Gesellschaft. Research at Fraunhofer IIS revolves around two guiding topics: In the area of "Audio and Media Technologies," the institute has been shaping the digitalization of media for over 30 years. Fraunhofer IIS was instrumental in the development of mp3 and AAC and played a significant role in the digitalization of cinema. Current developments are opening up whole new worlds of sound and have applications in virtual reality, automotive sound systems, mobile telephony, streaming, and broadcasting.

In the context of "cognitive sensor technologies," the institute researches technologies with applications in sensor technology, data transmission technology, data analysis methods, and the exploitation of data as part of data-driven services and their accompanying business models. This adds a cognitive component to the function of the conventional "smart" sensor.

More than 1,100 employees conduct contract research for industry, the service sector, and public authorities. Founded in 1985, Fraunhofer IIS now has 16 locations in 12 cities: Erlangen (headquarters), Nuremberg, Fürth, and Dresden as well as Ilmenau, Munich, Bamberg, Waischenfeld, Coburg, Würzburg, Deggendorf, and Passau. The annual budget of 167.9 million euros is mainly financed by contract research projects. In addition to this, 29 percent of the budget is subsidized by federal and state funds. For more information, visit: www.iis.fraunhofer.de