POLITECNICO DI TORINO

Dipartimento di Automatica e Informatica

PhD in Computer and Control Engineering

XXXII cycle

Supervisor

Prof. Andrea Calimera

Prof. Enrico Macii

Design Methodologies and CAD tools for Ultra-low Power IoT Systems and Applications

PhD Candidate: *Valentino Peluso*

1. Context

A scalable implementation of Internet-of-Things calls for smart sensors able to generate information from the collected data. The rise of deep learning, Convolutional Neural Networks (ConvNet) in particular, represents a new opportunity in the IoT. The processing of ConvNet on low-power sensor nodes is highly critical in terms of computation power: we need novel **hardware/software optimization** tools to overcome the technology limits posed by the twilight of the Moore's law.

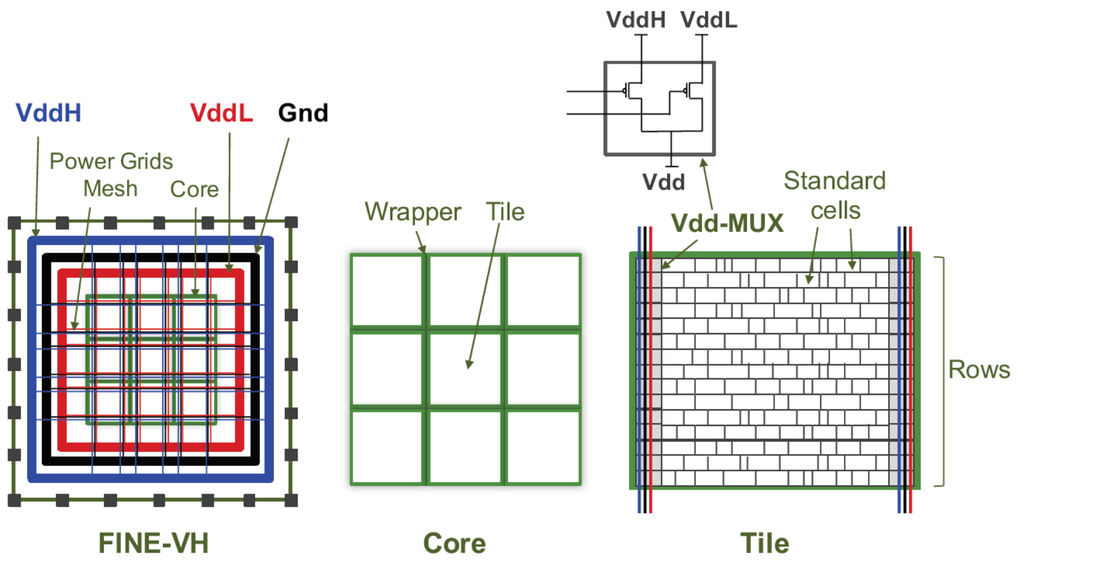
2. Hardware Optimization

DVFS is the de-facto standard for low energy digital circuits. It is based on the principle of lowering the supply voltage to the minimum threshold that satisfies the frequency constraint required by the actual workload. An ideal-DVFS deals with the availability of on-chip high resolution voltage regulators that can deliver the supply voltage with a fine step resolution, a design option that is too costly.

While previous research focused on alternative solutions that can achieve, or at least get close to, the efficiency of ideal-DVFS while using a discrete set of supply voltages, this work introduces Ultra-Fine Grain Vdd-Hopping (FINE-VH) [1][2], a practical methodology that brings DVFS beyond its theoretical limit.

FINE-VH leverages the working principle of Vdd-Hopping applied at ultra-fine granularity, i.e. within-the-core.

Experimental results demonstrate FINE-VH allows substantial power savings w.r.t. coarse-grain (i) ideal-DVFS, (ii) Vdd-Hopping, (iii) Vdd-Dithering.



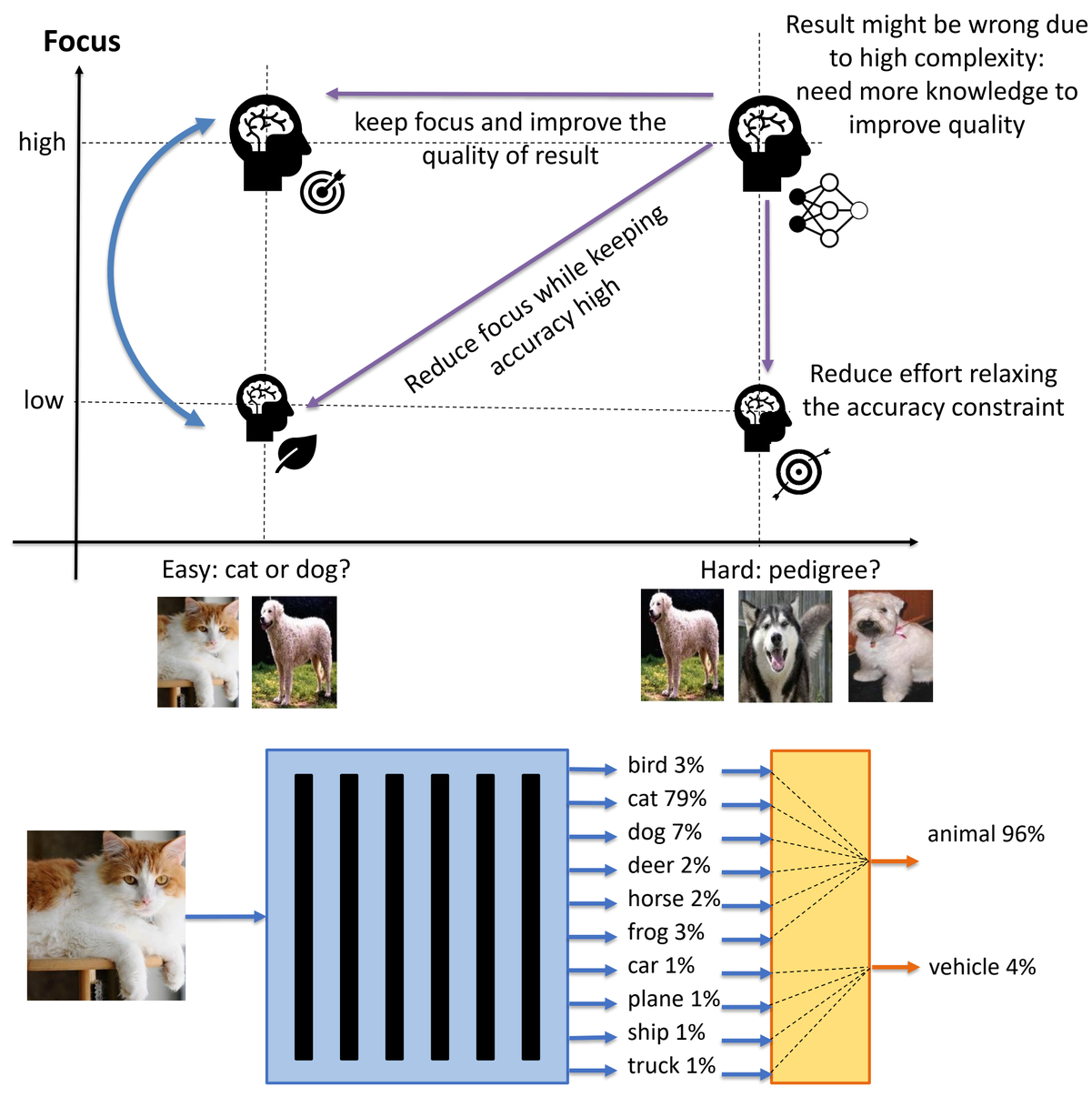
From 32.0% to 38.2% w.r.t. Ideal-DVFS

3. Software Optimization

Most of today's supervised ConvNets are designed, trained, and then used as static models. They are flat classifiers that expend equal effort no matter the surrounding context and the level of accuracy required by the application. The human brain, to which ConvNets are inspired, works differently. We, as humans, can categorize concepts using hierarchical semantic abstractions. Such organization allows reasoning process to move toward the proper abstraction level depending on the actual requirement. For instance, we first recognize the surrounding environment using high-level abstraction, e.g. indoor or outdoor. Then, only when/if required, we focus on more details and bring reasoning to the lower levels by pushing more effort. We might be interested in classifying not just the place we are, but also things around us, e.g. the type of vehicles in the street, car or truck, or even their brand and model.

The abstraction process is typically driven by some external trigger. This scalable mechanism is what allows our brain to adapt to the context and reach the effort-accuracy trade-off that minimizes energy waste.

We developed a practical strategy for the implementation of this brain-inspired paradigm with embedded ConvNets. [3].



4. References

1. V. Peluso et al. "Ultra-Fine Grain Vdd-Hopping for Energy-efficient Multi-processor SoCs", Proc. VLSI-SoC 2016.
2. V. Peluso et al. "Beyond Ideal DVFS Through Ultra-Fine Grain Vdd-Hopping", VLSI-SoC Book, Springer, 2017.
3. V. Peluso et al. "Scalable-effort ConvNets for Multilevel Classification", Proc. ICCAD 2018.