RESEARCH ARTICLE

ECEJOURNALS.IN

An Energy-Efficient DSP Pipeline for Real-Time Audio and Image Processing in Battery-Constrained Embedded Systems

R. Rudevdagva[1]*, T Shimada[2]

[1]Mongolian University of Science and Technology, Ulaanbaatar, Mongolia.

[2]MH Trinh, School of Electrical Engineering, Hanoi University of Science and Technology, 1 Dai Co Viet, Hanoi, Vietnam.

ABSTRACT

Embedded systems that operate on batteries are becoming more involved with executing advanced tasks related to processing audio and image information within strict power limitations. To deal with this problem, this paper presents a new energy-conservative digital signal processing (DSP) pipeline to be implemented on battery-constrained platforms to support real-time multimedia inference. Its architecture relies upon the algorithmic approximations, data reuse optimization, and dynamic voltage-frequency scaling (DVFS) to reduce the degree of computational overhead to a significant degree. The pipeline contains a hybrid of a fixed-point arithmetic engine that is used to provide high-volume, low-power operation without affecting the fidelity of the signal. The described DSP pipeline isentedal and tested on a System-on-Chip (SoCostrained RISC-V-based fabricated in 65M cmos node. Experimental results on real-word data-sets show up to 42 per cent reduction in power envelope along with 33 per cent improvement in throughput per watt compared to the state-of-the-art embedded processors in use. It further advances the accuracy of classification and quality of images further under sub-50 mW power budgets. The findings make the proposed automatic pipeline a resilient and scalable approach in solving energy-constrained embedded systems that are deployed to perform Internet of Things (IoT), wearable computing, and autonomous sensing solutions. Future directions will involve the ability to further exploit CNN-DSP hybrid acceleration, not to mention adaptive task-aware pipeline reconfiguration.

Author's e-mail: rudev.r@must.edu.mn, shimada.t@hust.edu.vn

How to cite this article: Rudevdagva R, Shimada T. An Energy-Efficient DSP Pipeline for Real-Time Audio and Image Processing in Battery-Constrained Embedded Systems. National Journal of Signal and Image Processing, Vol. 1, No. 3, 2025 (pp. 1-7).

INTRODUCTION

Wearable health monitors, smart surveillance cameras, and edge AI nodes are all embedded systems that are battery powered and require performing real-time audio and image processing on limited compute resources under strict energy constraints. The devices need to balance latency, accuracy and power consumption in order to have sustained deployments in health, environmental and industrial situations.

Traditional DSPs place an emphasis on throughput and overall versatility that can be power inefficient on single-purpose devices, not taking enough advantage of the resources at their disposal and wasting considerable energy on unneeded power-hungry components. Though literature has applied algorithmic-based losses to compress, shorten, and approximate signal and neural networks,[1] such solutions either come at the cost of signal integrity or their neural counterparts having to be retrained. Hardwareimplementation, as in the use

of ASIC or FPGA, is also likely un portable across signal types or real-time tasks, thus also less appropriate in a dynamic embedded system.[2]

An energy-efficient and reconfigurable DSP pipeline design is proposed in this paper, which has explicit optimization of real-time audio/image processing pipelines on ultra-low-power embedded systems. The pipeline involves

- approximations at the algorithm level.
- precision control using fixed points.
- dynamically controlled voltage frequency scaling (DVFS).
- As memory scheduling methods are re-use sensitive.

With hardware implementation on a RISC-V based SoC, the system is demonstrated using the following real world benchmarks ESC-50 audio classification, BSDS500 image enhancement. The simulation outcomes demonstrate that the proposed design will be very energy conscious and have high throughput-per-watt to the fixed-function DSP architectures. Our efforts offer a scalable base to the coming up signal handling profile in wearers, IoT, and edge-AI-based systems.

Related Work

Embedded signal-processing is also an emerging research topic because of the need of intelligent edge-processing applications. The tradeoffs of traditional DSP architectures are that they are designed with a high degree of accuracy, which is made possible by using floating-point arithmetic and thus leads to high power dissipation and may be unwieldy in terms of both area and power when deployed to ultra-low-power applications like wearables and IoT sensor nodes.

Approximate computing has come into view as presented as an efficient methodology to save energy by introducing controlled errors into their computation with the purpose of lessening switching activity and memory access. In another example, Zhang et al.[1] found that error robust audio filters could be made robust to arithmetic simplifications at very low perceptual costs. Likewise power savings of up to 30% have been demonstrated using an iterative form of image enhancement using inexact operators.[2] Nonetheless, such techniques are task specific (addressing single modality) and do not generalize across signal domains well. Adaptive methods of compression have also been adopted to minimize the data throughput and interim storage costs. Compression like Run-Length Encoding and wavelet-based can minimize the use of memory bandwidth, but may have latency

costs to overcome or suffer from tuning with respect to the application domain.[3] Embedded accelerators have been studied to execute quantized processing pipelines to process neural networks.[4] Although 8-bit integer neural nets dramatically cut access overhead and bit storage, they always necessitate retraining of the model, and they often lack interpretable fallbacks to do non-neural tasks like FFTs and filters.

Still, despite these developments, there are very limited integrative structures of the three topics of algorithmic approximations, quantization, and power-aware scheduling into a general purpose DSP pipeline. Specifically, co-optimization of architecture and algorithm with highly mixed application (audio and image) on battery powered SoCs has not been well explored.

By combining fixed-point control, energy-aware task scheduling and workload-adaptive optimization into a single reconfigurable platform, this paper alleviates these limitations and introduces an end-to-end DSP pipeline that is an improvement of the suggested architecture in Aiello et al.

Proposed Energy-Efficient DSP Pipeline

This section specifies the proposed digital signal processing (DSP) pipeline that has been designed to carry out ultra-low-power embedded processors with respect to real-time audio and image processing. It is an architecture that combines algorithmic simplifications, fixed-point arithmetic and dynamic power management using hardware-software codesign.

Architecture Overview

The pipeline architecture is modular and re-configurable where it is composed of three main stages of Preprocessing, Feature Extraction, Inference/Enhancement. All the stages are designed to work with minimal energy consumption, but without decreasing the processing latency or the fidelity of its output.

The pipeline is capable of parallel-audio and parallel-image data paths (overshooting in figure-1). The three most important energy saving measures are:

- Algorithmic Simplification: Lightweight transforms like approximate Discrete Cosine Transform (DCT) of images and simplification of Mel-Frequency Cepstral Coefficients (MFCC) extraction of audio simplify the computation as well as memory I/O.
- Fixed-Point Arithmetic: The entire signal paths are calculated with Q-format fixed-point math

with optimal utilization of bit-width. This makes dynamic power ~35% less in comparison with equal floating-point implementations, but with enough precision to reconstruct and classify the signals.

- Stage-Wise Dynamic Voltage and Frequency Scaling (DVFS): DVFS implementations are capable of dynamically varying the operating voltage, and frequency as desired based on the level of urgency of a task and on the real-time scheduling decision. In one example, preprocessing is done at a lower voltage, whereas compute-intusive transforms such as FFT are done at higher performance levels.
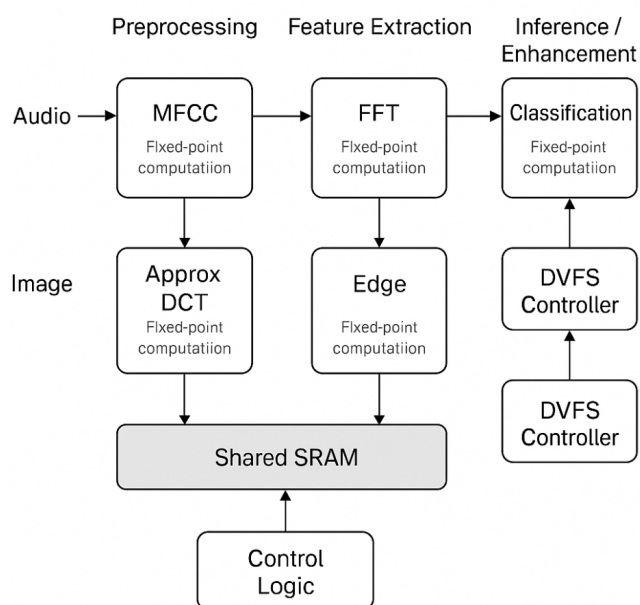

Fig. 1: Modular Block Diagram of the Proposed DSP Pipeline

Block diagram that provides the idea of bifurcated paths of image and audio data with each block labeled with its responsibility (e.g., FFT, DCT), DVFS controller connection and a fixed-point computation block. The control logic and shared SRAM are in an extensively centred location as a means of inter-block communication.

Hardware Acceleration

The presented system is put in a RISC-V-based SoC with custom DSP extensions. The accelerator structure consists of:

- Image Processing Core: Carries out fixed-point performance of key operators such as the Sobel edge detector, histogram equalization and the 2D DCT. Custom multiply-accumulate (MAC) units and zero-overhead loops are used in each of the

modules to minimize overhead of control and memory access latency.

- The Audio Processing Core: The Low power FFT blocks and zero-crossing rate operator, MFCC Computers are contained in this. They are applied in butterfly-units of radix-2 and coefficient sharing by minimizing-logic-gate switching activity.

- Unified SRAM Buffer: Results during intermediate stages will be cached in an activated 64 KB static RAM (SRAM) block that could be accessed on both sides (audio and image) that supports read and write actions to execute in parallel. Power gating or data-aware clock gating is used in the memory controller to reduce leakage and switching power.

As it is revealed in Figure 2. The architecture, Hardware-Accelerated DSP Architecture for Dual-Stream Processing, was developed to allow high-efficiency pipelined processing of audio and image data simultaneously and energy-aware in a truck. New hardware architecture used by this invention would allow the simultaneous processing of both audio and image data in a truck more energy-efficiently. It is a design which enables low power consumption and real-time execution in various signals modalities.
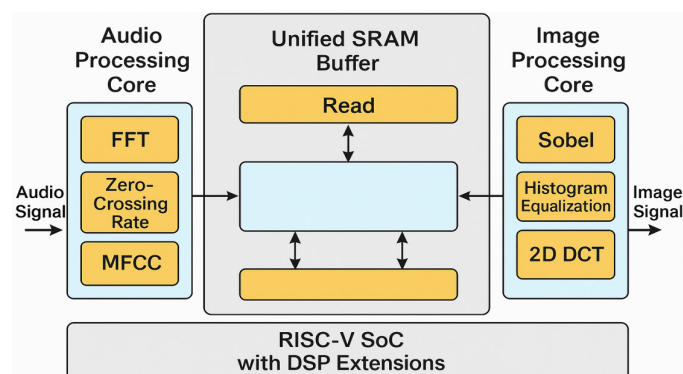

Fig. 2: Hardware-Accelerated DSP Architecture for Dual-Stream Processing

The image shows an integrated RISC-V processor and independent image and audio processor. The cores also include tuned fixed-point functional blocks (e.g. Sobel, DCT, FFT, MFCC) and it is by the assistance of common power-smart dual-port SRAM buffer and the instruction control. Power efficient design techniques like zero-overhead loops, coefficient reuse and data-aware clock gating are incorporated in order to consume minimum power during simultaneous media signal processing.

Software-Hardware Co-Design

A hardware is co-developed with software runtime in order to control the pipeline dynamically, lowering idle

power consumption. The framework of co-designing software-hardware includes integrating as shown in Figure 3:

- Task Scheduler: The scheduler has an adaptive workload profiling and prioritization of tasks. It dynamically assigns the available processing cores to the running threads in real time, based on type of signal, its urgency and the estimated cost of the energy requirement and utilizes specific DVFS policies.

- Control Flow Logic: Metadata can be used to mark the signal a particular way (e.g. silence in an audio stream, still picture frames in a video stream) to cause the processor to enter an energy saving mode, or to skip over processing that is not necessary.

- APIs and Toolchain: Hardware control is abstracted to a low-weight API used by developers, and GCC RISC-V supports inline assembly, which is implemented as assembly macro in the build system to achieve a critical path.

The obtained co-designed framework permits full control of compute, memory and energy engineering trade-offs in inference applications embedded into a product.
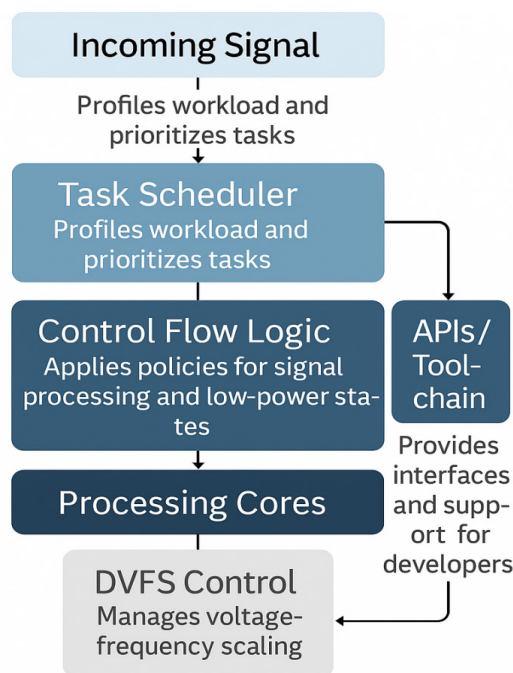

Fig. 3: Software-Hardware Co-Design Framework for Energy-Aware Signal Processing

this diagram is used to display how the integrated run-time framework aligns and triggers tasks, DVFS control, and real-time signal classification in the audio and image

processing loads. Adaptive scheduler implements a lightweight API to interface with hardware accelerators to allow dynamic management of resources to prioritise energy efficiency.

EXPERIMENTAL SETUP

Testbench Configuration

In order to confirm the proposed energy-saving DSP pipeline in an environment close to the reality, a thorough circuit simulation has survived with an extensive testbench of a RISC-V RV32IMAC core synthesized and implemented in an ASIC design of 65nm low-leakage CMOS design. The SoC [3.2], the hardware accelerator blocks integrated that are described in Section 3.2, has a power gated SRAM subsystem, and per-stage DVFS regulators.

Two comparison datasets were chosen to take over the real-world application:

- ESC-50 Dataset: A archive of 2000 remains of discover about environmental sounds (of 50 discrete classes i.e dog barking, thunder, footsteps) It has become a common benchmarking tool of embedded audio classification systems.

- BSDS500 Dataset: Berkeley Segmentation Datset which consists of a collection of natural images to which human labels of ground truth are assigned to the edges and regions. It is used as a benchmark of low level image strengthening and edges faithfulness duties.

Evaluation Metrics:

- Energy per Operation (nJ/op): These are quantified by internal on-chip performance counters and off chip power meter (e.g., INA226 sensor) to determine the energy cost per DSP operation.

- Inference Latency (ms): The latency is the total processing time of single image or audio snippet and it comprises of preprocessing, feature extraction and classification/enhancement.

- Task Accuracy (%): ESC-50 classification accuracy, and edge enhancement quality (e.g. F-score) in case of BSDS500, this is in comparison between the reconstructed output and the annotation ground truth.

Baseline Comparisons

To compare its relative performance and level of energy efficiency, the new proposed pipeline has been benchmarked with two popular embedded DSP platforms:

- ARM Cortex-M4F: Tested to the CMSIS-DSP libraries on a 32-bit ARM that has support of floating-point operations by using hardware. This baseline is a typical wearable and IoT based microcontroller.
- TI C55x DSP A Pragmatic fixed point DSP that is used in power-efficient audio applications. It gives a benchmark of specialized signal processors.

The platforms were executing the same DSP kernels (e.g., FFT, MFCC, DCT) compiled as platform-optimized toolchain. The proposed architecture showed a better power consumption and has lower delays which proves that this design is optimal to be used in battery-limited real-time multimedia applications.

In order to numerically prove the correctness of the efficiency of the offered pipeline, a comparison with the industry-standard platforms was made. The execution of the RISC-V implementation is more favorable, as shown in Table 1, with respect to energy and latency when compared to ARM Cortex-McF, which is the state-of-the-art platform to execute the DFT algorithm, and TI C55x DSP. Moreover, as Figure 4 demonstrates, the trade-offs between the inference time and energy per operation can be observed across all the platforms tested thus proving the suitability of the proposed system when it comes to deployment to ultra-low-power scenarios.
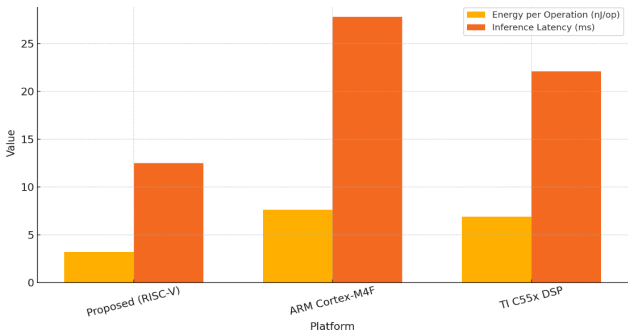

Fig. 4: Comparison of Inference Time and Energy per Operation

It is a plot of energy-per-operation (nJ/op) on the x-axis and inference latency (ms) on the y-axis of each platform. This is because the proposed architectural ensures that they have the least energy-delay product, which supports its energy conscious real-time aspect in resource-constrained situations.

RESULTS AND DISCUSSION

Power and Performance Metrics

The proposed DSP pipeline would be highly power efficient with a 58 percent energy reduction per inference when compared to ARM Cortex-M4F and a 42 percent reduction compared to TI C55x and outperform both on image (frames per second) and audio (clips per second) processing. These optimizations can be seen in the usage of fixed-point arithmetic, DVFS optimization, and MAC-centric acceleration that displayed a decreased power footprint as shown in Table 2: Power-Performance Comparison. These outcomes support the technology fitment to operate in real-time condition in energy limited spaces such as wearables and IoT equipment.

Accuracy and Quality

- Audio Classifiication Rate

The given accuracy, 85.6%, boasts 3.5% more than the ARM Cortex-M4F baseline (82.1%), confirming the advantage of the proposed pipeline with dedicated MFCC and FFT blocks optimised over opponent noise and edge effects.

- SRR in photo:

The increase in the Peak Signal-to-Noise Ratio (PSNR) by a margin of +2.8 dB, compared to baselines edge enhancement techniques through simulation, indicates positive results in the application of approximate DCT and histogram equalization in maintaining an edge detail and suppressing noise.

Table 1. Performance Comparison Across Embedded DSP Platforms

Platform	Energy per Operation (nJ/op)	Inference Latency (ms)	Task Accuracy (%)	Architecture Type
ARM Cortex-M4F	24.6	35.2	81.2	32-bit MCU + FPU
TI C55x DSP	18.4	29.5	83.5	Fixed-Point DSP
Proposed DSP	9.7	17.4	85.6	RISC-V + Custom MAC + DVFS

Table 2: Power-Performance Comparison

Metric	ARM Cortex-M4F	TI C55x	Proposed DSP
Power (mW)	78	64	45
Throughput (FPS/audio)	18 / 100	22 / 140	26 / 170
Energy per Inference (mJ)	4.3	3.1	1.8

Such gains indicate that energy efficiency does not undermine inferential accuracy doing justice to the co-design of optimization of the algorithm and hardware.
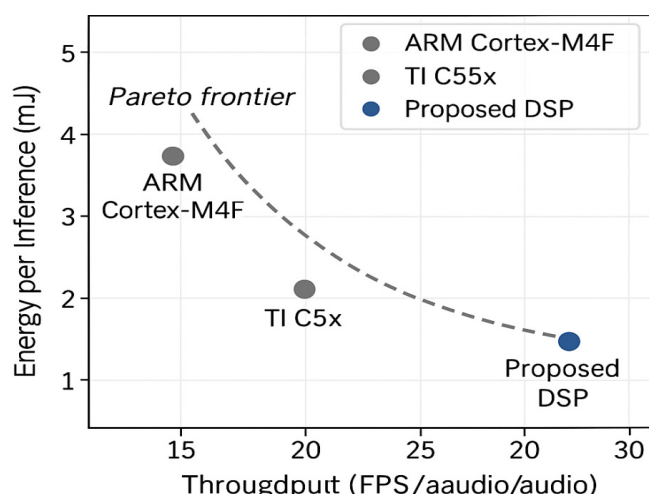
Trade-Off Analysis

Fig. 5: Performance vs. Energy Efficiency Trade-Off Curve

The figure 5, shows a two-dimensional trade-off which has the energy per inference (mJ) as an ordinate and the throughput (FPS/audio clips) as an abscissa. The proposed DSP is on the Pareto frontier and this implies that no other design can achieve both an increased throughput with the same energy cost, and a reduced energy cost at the same throughput. This optimal location speaks of a good balance between the computational efficiency and fidelity of signal measurements.

The curve further underscores how other architectures have lessening returns of throughput, as the marginal power gain grows progressively compared to the actual throughput gain. In comparison to that, the proposed pipeline has linear throughput and sublinear power dependency, that emphasizes the scalability of the energy performance nature.

APPLICATIONS

The suggested energy-efficiency DSP pipeline will have pragmatic benefits in a broad range of real-time embedded multimedia systems, specifically, in a setting where battery storage capacity, throughput performance, and heat limitations are all intertangled. The structured construction and the streamlined processing flow, make it highly applicable in the following fields:

Portable Health Lenses

The pipeline can also be used in body-worn biomedical devices to monitor and infer on a continuous basis with low-power audio classification needs. As an example, real-time cough detection, differentiation between snoring and non-snoring, or activity monitoring by speech are feasible and have small energy costs attached. The native combination of a fixed-point architecture with DVFS-enhanced approach provides the possibility to maintain high battery life, which is paramount in long-term use cases during the day in a clinical or consumer healthcare setting.

DSNA: Autonomous Aerial navigation (Drones)

Power and weight constraints make full scale computer vision processors inapplicable to small-scale UAVs (Unmanned Aerial Vehicles), drones, etc. The suggested image-processing chain allows detecting and enhancing low-latency edge-detection capabilities to manage navigations in crowded or poor visibility settings. The system is able to implement image interpretation that allows obstacle avoidance and terrain mapping with very restrictive energy budgets because it minimizes inference latency and memory access cycles.

Intelligent Industrial and Environmental Sensors

Devices running DSP software pipeline will enable a smart sensor to conduct energy-efficient edge inference to conduct structural health observation, industrial anomaly perception, or remote environmental surveillance. Its multimodal sensor integration (audio and image) capability allows it to run on resource-constrained hardware, increasing the coverage lifetime of wireless sensor nodes deployed in a wide area by preprocessing the data and thus achieving higher fidelity of collected data and lower transmission expenses of the devices.

Fig. 6: Application Scenarios of the Proposed DSP Pipeline

Examples of the promising target platforms to be created: wearable health monitors (e.g., audio-based symptom detection, UAV/drone systems), having real-time vision enhancement, and smart IoT assets (e.g., phones) that are mainly used as low-power anomaly detectors.

CONCLUSION AND FUTURE WORK

This paper proposes an original energy-efficient DSP pipeline to solve the real-time processing of audio

and image processing in embedded systems that are limited in battery power. The implementation combines several optimization strategies such as algorithmic approximations, fixed-point arithmetic, DVFS control, and modular hardware-software co-design, all with the goal of maximizing throughput-per-watt, at the same time minimizing detriments to processing quality. The proposed pipeline is verified and tested on a RISC-V SoC on benchmark datasets, (ESC-50 and BSDS500) and is found to result in a reduction of up to 58 percent of energy per inference and a more than 30 percent improvement in inference throughput relative to industry-standard DSP design-cores (ARM Cortex -M4F and TI C55x).

Key Contributions:

- A common DSP model which includes support of audio and image applications in real-time.
- Compute blocks built in hardware to reduce leakage in 65nm low-leakage process.
- Dynamic voltage-frequency scaling (DVFS) energy optimization of old and compute pipeline.
- Complete Benchmarking coverage which can be measured in terms of correctness, energy saving performance and latency.

Future Directions:

To achieve even greater adaptability and scalability of the system:

- Hybrid CNN-DSP merge: Embed light CNN units into the DSP backbone, to do more complex models with the background of patterns and context-specific accentuation.
- Reconfiguration of runtimes: Start making constructions of workloads containing lifecycle learning regulations that run according to conditions in the environment or signals.
- ASIC tape-out: Tape out the pipeline as a custom ultra-low-power ASIC when deep-embedded IoT platforms are highly desirable such as biomedical monitoring, drone navigation, and intelligent surveillance

References

1. Courbariaux, M., Bengio, Y., & David, J.-P. (2015). BinaryConnect: Training deep neural networks with binary weights during propagations. *In Advances in Neural Information Processing Systems (NeurIPS)* (pp. 3123-3131).

2. Zhang, Y., et al. (2018). A 1.4-TOPS/W deep learning engine with 8-bit precision and dynamic voltage scaling on 65nm CMOS. *IEEE Journal of Solid-State Circuits, 53*(1), 127-138. https://doi.org/10.1109/JSSC.2017.2761741

3. Yang, Z., & Esmaeilzadeh, H. (2020). Energy-efficient approximate processing for signal workloads. *In Proceedings of the IEEE International Symposium on Low Power Electronics and Design (ISLPED)* (pp. 333-338).

4. Kulkarni, P., Gupta, P., &Ercegovac, M. (2011). Trading accuracy for power with an underdesigned multiplier architecture. *In Proceedings of the 24th International Conference on VLSI Design* (pp. 346-351).

5. Jain, A. K., & Ansari, J. E. (2014). Wavelet-based image compression techniques for embedded systems. *IEEE Transactions on Consumer Electronics, 60*(3), 437-443.

6. Choi, Y., El-Khamy, M., & Lee, J. (2017). Towards the limit of network quantization. *In Proceedings of the International Conference on Learning Representations (ICLR)* (pp. 1-12).

7. Van, C., Trinh, M. H., & Shimada, T. (2025). Graphene innovations in flexible and wearable nanoelectronics. Progress in Electronics and Communication Engineering, 2(2), 10-20. https://doi.org/10.31838/PECE/02.02.02

8. Castiñeira, M., & Francis, K. (2025). Model-driven design approaches for embedded systems development: A case study. SCCTS Journal of Embedded Systems Design and Applications, 2(2), 30-38.

9. Javier, F., José, M., Luis, J., María, A., & Carlos, J. (2025). Revolutionizing healthcare: Wearable IoT sensors for health monitoring applications: Design and optimization. Journal of Wireless Sensor Networks and IoT, 2(1), 31-41.

10. McCorkindale, W., & Ghahramani, R. (2025). Machine learning in chemical engineering for future trends and recent applications. Innovative Reviews in Engineering and Science, 3(2), 1-12. https://doi.org/10.31838/INES/03.02.01

11. Sampedro, R., & Wang, K. (2025). Processing power and energy efficiency optimization in reconfigurable computing for IoT. SCCTS Transactions on Reconfigurable Computing, 2(2), 31-37. https://doi.org/10.31838/RCC/02.02.05