

Advanced MCUs Bring AI to Cost-Sensitive Applications

Neural-network processors accelerate AI program execution while development tools help you get to market fast.

Microcontroller units (MCUs) with neural-network processors (NPs) bring edge artificial-intelligence (AI) capabilities to advanced applications extending from wearable devices to factory-automation systems. Eliminating the need for continuous cloud connectivity, such MCUs enable compact solutions with fast response times and minimal power consumption. In addition, AI development tools help you build intelligent applications in minutes rather than months.

Deploying AI Models at the Edge

[Examples of applications that can benefit from AI models deployed on MCUs range from wearable technology to industrial-equipment monitoring.](#) Wearable health trackers, for example, can leverage neural networks (NNs) to provide touch-free gesture recognition, which can illuminate the device's display at the flick of the wrist.

As shown in *Figure 1*, these devices employ accelerometers or gyroscopes to detect human motion. An analog front end (AFE) conditions the sensor output signals, and an analog-to-digital converter (ADC) digitizes the signals for conveyance to an MCU with integrated digital-signal-processing (DSP) and NN processing capabilities. The MCU

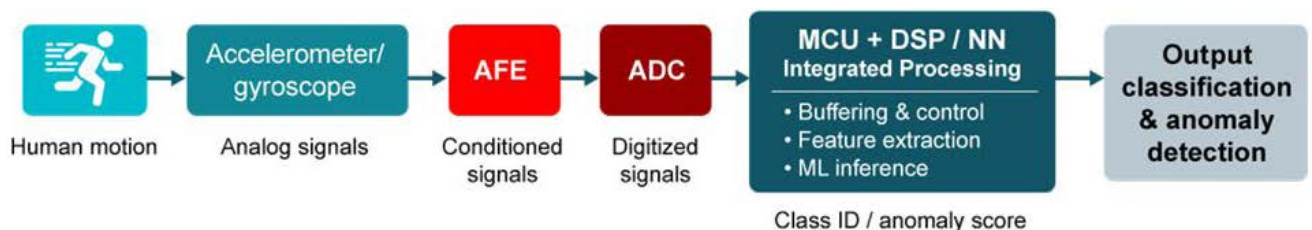
performs buffering and control, feature extraction, and machine-learning inference operations to provide output classification and anomaly detection.

Because health trackers must be compact and lightweight, they require compact MCUs with integrated analog and digital peripherals that occupy only a few square millimeters of PCB real estate.

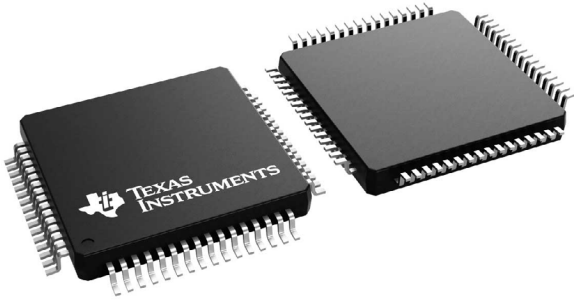
Other applications that can make use of NNs include smart home systems. Such systems often employ an audio subsystem charged with detecting a spoken wakeup word. This subsystem typically includes:

- An analog microphone to capture the user's voice.
- An AFE to provide amplification and filtering.
- An ADC to provide digitization.
- An Inter-IC Sound (I2S) or time-division multiplexing (TDM) interface to carry the digitized signal to the DSP- and NN-equipped MCU.

Because the function must operate continuously to listen for the wakeup word, its components must be low power yet sufficiently responsive that users need not repeat the wakeup word. If a valid wake-up word has been spoken, control can be transferred to a more powerful local processor or cloud-based AI model.



1. In a wearable device, accelerometers and gyroscopes detect human motion, with their outputs conditioned, digitized, and sent to an MCU for output classification.



2. The MSPM0G5187 with an Arm Cortex-M0+ CPU and TinyEngine NPU comes in a 64-pin package that measures 12 × 12 mm.

Texas Instruments reports that for such applications, it offers NPU-equipped MCUs that consume only tens of milliwatts, compared to the watts of power required for typical voice-processor integrated circuits. The company also reports that an AI keyword recognition model using a one-dimensional convolutional NN can reduce processing time by more than 90X compared to running the same model on a standard MCU.

Yet another edge-AI application is industrial-motor vibration detection, which can be used to schedule predictive maintenance for conveyers, pumps, or other motor-driven equipment. In such equipment, local AI models running on an MCU are able to extract time-domain anomalies such as impulse spikes and irregular periodicity that can indicate

impending failure.

In such systems, accelerometers monitor mechanical vibration, and as in wearable and smart home systems, an AFE, ADC, and DSP and NN-equipped MCU complete the signal chain necessary to provide output classification and anomaly detection.

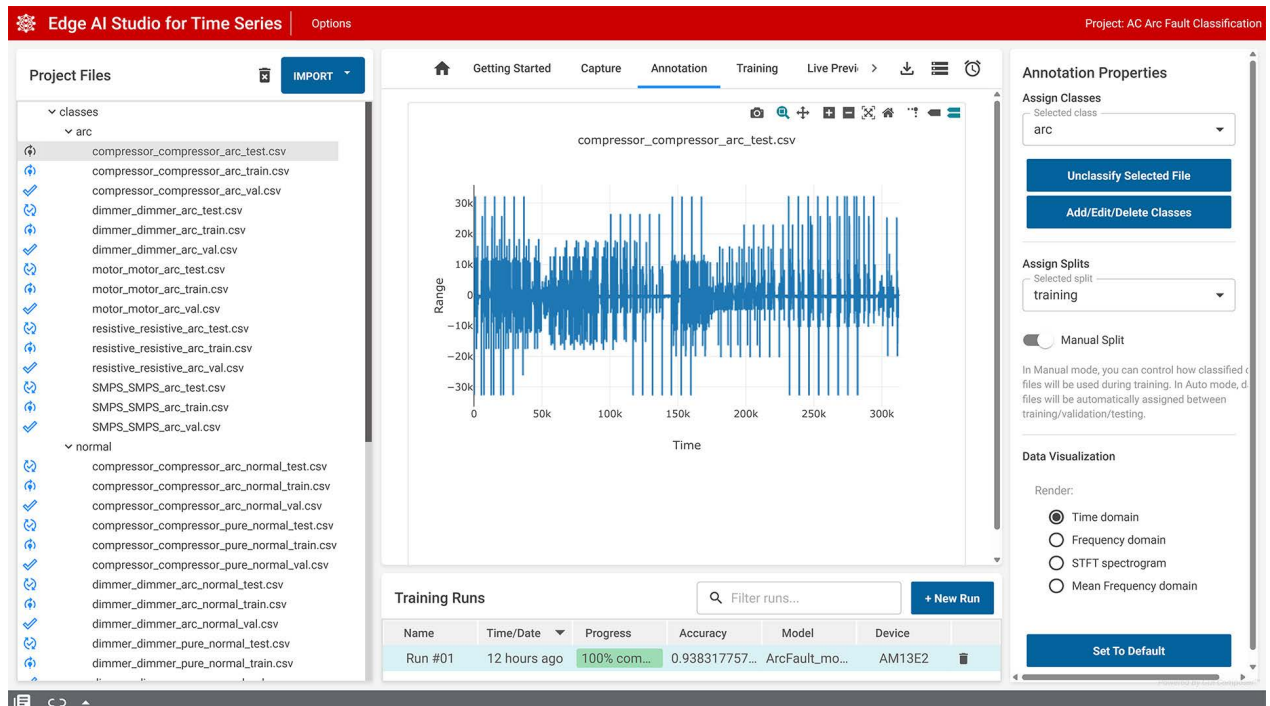
80-MHz MCU with NPU

To power such applications, [TI offers the MSPM0G5187 mixed-signal MCU with a TinyEngine NPU](#) (Fig. 2). This device includes a 32-bit Arm Cortex-M0+ CPU operating up to 80 MHz with up to 128 kB of flash memory and up to 32 kB of SRAM. Analog peripherals include a 12-bit, 1.6-MSPS ADC supporting up to 26 external channels, a high-speed comparator with integrated reference, a digital-to-analog converter (DAC), and two voltage references.

The device offers a range of optimized low-power modes, extending from an 88-nA shutdown mode with I/O wake-up capability to a run mode that draws 103 µA/MHz.

The MSPM0G5187's TinyEngine NPU, an optimized core for running deep convolutional neural networks (CNNs), lets you build applications that run fast. Operating at 80 MHz and autonomously from the Arm CPU, the NPU is a fully programmable hardware accelerator that works in conjunction with the CPU to provide high performance and low-power operation for CNN inference using pre-trained models.

The NPU can execute 2.56 gigaoperations per second



3. Edge AI Studio supports time-series analysis for monitoring and control applications.

(GOPs) and significantly reduces latency and energy usage per inference compared with software-only implementations. It enables you to use 8- or 4-bit input activations and 8-, 4-, or 2-bit weight parameters. The NPU supports generic convolutional layers, pointwise layers, depthwise layers, pooling layers (max/average), and residual layers.

Development Tools for Edge AI

To speed AI application development for devices such as the MSPM0G5187, [TI offers Edge AI Studio, part of the company's CCStudio development-tool ecosystem](#). In addition to supporting an MCU with a TinyEngine NPU, it works with a TI processor or system-on-chip (SoC) featuring a C7 NPU as well as an AI-capable device without dedicated accelerators. The Edge AI Studio provides a fully integrated solution for data management, model training, and deployment on a live development platform.

To help you quickly provide proof of concept, you can use Edge AI Studio with a model from TI's Model Zoo, a collection of pretrained models optimized to run efficiently on TI Edge AI processors. Alternatively, you may “bring your own data” (BYOD) to retrain an existing model, or you can even “bring your own model” (BYOM). Edge AI Studio also offers command-line tools to support open-source environments such as PyTorch and TensorFlow.

Edge AI Studio facilitates the real-time analysis of time-series data for monitoring and control applications, including motor-bearing and fan-blower-imbalance fault detection, allowing TI MCUs to promptly schedule preventative maintenance (*Fig. 3*). Edge AI Studio also supports machine-vision and radar applications.

Conclusion

To bring AI capability to a variety of home and factory applications, an MCU with integrated NPU provides a fast, compact, and low-power solution. Associated development tools let you get to market quickly, whether you choose a pretrained AI model or take a BYOD or BYOM approach.