A Low-cost and Low-power Predictive Maintenance System Based on TinyML and Self-powered Sensor

Zijie Chen*a,d,e, Yiming Gaoa,b,c, Junrui Lianga,d

*aSchool of Information Science and Technology, ShanghaiTech University, Shanghai 201210, China
bShanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201210, China
cUniversity of Chinese Academy of Sciences, Beijing 100049, China
daShanghai Engineering Research Center of Energy Efficient and Custom AI IC, Shanghai 201210, China
eCorresponding author: chenzj1@shanghaitech.edu.cn

Keywords: Predictive maintenance, TinyML, Self-powered sensor, AIoT

In the coming Industry 4.0 era, increasing industry fields demand efficient monitoring and maintenance approaches. Predictive maintenance (PdM) proves a prominent strategy that can recognize the current status and predict the future trend of machines, preventing disastrous breakdowns and improving the reliability and efficiency of manufacturing systems. Such systems are mainly based on artificial intelligence (AI) algorithms that run on resource-rich and power-hungry personal computers (PC) or servers. In this paper, we propose an on-device PdM system based on the self-powered sensor (SPS) and tiny machine learning (tinyML), with an emphasis on low-cost and low-power characteristics, which remains unexplored. TinyML compensates for the distorted SPS data, equipping the end machine with intelligence for perceiving its situation in real-time with short latency, high accuracy, and low power consumption. Fig. 1 shows the architecture of the proposed system. When anomalies happen, the vibration information of the machine will change simultaneously. A lightweight piezoelectric cantilever as an SPS, as shown in Fig. 2(a), is utilized to substitute the inertial measurement unit (IMU) and form an endpoint device with a low-cost microcontroller (MCU). The real-world data for different machine statuses are collected firstly with an ultralow sampling rate for saving energy and used to pre-train a TinyML model. With the help of experts, some specific abnormal information can be obtained in advance; thus, the system...
can detect the occurrence and the corresponding type of anomalies. After deploying the compressed model on the MCU, the immediate real-world signal is sensed, preprocessed, and inferred. According to the application requirements, the MCU can send alarms only when a suspicious condition is detected. The system cuts off the close dependency on servers and unleashes a plethora of opportunities for edge intelligence.

A rich SPS dataset for similar vibration conditions has been collected in a simulated vibration environment, whose experimental setup is shown in Fig. 2(b). Six well-known machine learning algorithms, including Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR), K-Nearest Neighbor (KNN), and Deep Neural Network (DNN), are applied to the SPS data from different perspectives. RF and DNN are on par for identifying raw normalized SPS signals in five classes with an accuracy of up to 98%, when the required data length, data number, and sampling rate for each case are 8, 150, and 20 Hz, respectively, which the traditional fast Fourier transform (FFT) method proves tight to reach. The fine-tuning of the models is carried out to maintain accuracy when the size is optimized. The tradeoff between data size, model size, accuracy, development cost, and the energy consumption is discussed in detail. To thoroughly evaluate the power-saving capability of the proposed system, measurements of energy consumption are conducted. The energy and time cost of tinyML models prove lower than 40 µJ and 0.3 ms, respectively. Results demonstrate the SPS-based system can save 66.8% of energy for carrying out the same AI-based detection task compared to the IMU-based system. Fig. 2(c) illustrates the compact prototype and setup of the field test. A three-phase machine mounted with the prototype is controlled by the digital power supply. We set three different real-world vibration states, i.e., idle, normal, and abnormal, by changing the supplied voltage. Based on the data analysis, a small dataset of 100 seconds is sufficient for each class. The original sampled data for three cases are transmitted wirelessly to a PC for model training. A DNN model is generated with only 331 parameters for triple classifications and deployed on the MCU, which can detect the anomalies with 99% accuracy. The proposed system is general and can be applied to detecting irregularities and preventing severe events in other application scenarios with vibration signals, such as aero engines, industrial robots, launch vehicles, etc. It is hoped that this work can contribute to the successful development of TinyML.