

Flexible self-rectifying synapse array for energy-efficient edge multiplication in electrocardiogram diagnosis

Received: 13 June 2024

Accepted: 29 April 2025

Published online: 09 May 2025



Younghyun Lee¹, Hakseung Rhee¹, Geunyoung Kim¹,
Woon Hyung Cheong¹, Do Hoon Kim¹, Hanchan Song¹, Sooyeon Narie Kay¹,
Jongwon Lee²✉ & Kyung Min Kim¹✉

Edge computing devices, which generate, collect, process, and analyze data near the source, enhance the data processing efficiency and improve the responsiveness in real-time applications or unstable network environments. To be utilized in wearable and skin-attached electronics, these edge devices must be compact, energy efficient for use in low-power environments, and fabricable on soft substrates. Here, we propose a flexible memristive dot product engine (f-MDPE) designed for edge use and demonstrate its feasibility in a real-time electrocardiogram (ECG) monitoring system. The f-MDPE comprises a 32×32 crossbar array embodying a low-temperature processed self-rectifying charge trap memristor on a flexible polyimide substrate and exhibits high uniformity and robust electrical and mechanical stability even under 5-mm bending conditions. Then, we design a neural network training algorithm through hardware-aware approaches and conduct real-time edge ECG diagnosis. This approach achieved an ECG classification accuracy of 93.5%, while consuming only 0.3% of the energy compared to digital approaches, highlighting the strong potential of this approach for emerging edge neuro-morphic hardware.

Artificial intelligence (AI) has gained widespread attention due to its remarkable performance in diverse applications, such as data processing, image recognition, and classification¹. The advancement of AI involves more than just enhancing algorithms; it also entails creating increasingly complex and deeper neural networks. This necessitates the use of additional computing parameters, which strains conventional hardware. Consequently, there is a rising demand for energy-efficient AI accelerators capable of efficiently handling the growing volume of data.

Additionally, there has been a significant surge of interest in implementing AI on edge devices, often referred to as edge AI². This technology provides unique advantages in AI utilization; edge AI enables the use of AI in environments where networks are not

connected or in situations with security concerns regarding network usage. Furthermore, edge AI can conduct initial computations of cloud-based deep neural networks, simplifying the large volumes of data generated at the input stage before transmitting them to cloud AI. This helps alleviate data bottleneck issues between the input source and cloud AI. However, edge devices must be more energy-efficient, considering the inherent energy constraints in edge environments. Additionally, they should be wearable or attachable to ensure effective utilization in edge environments³.

In this regard, traditional CMOS transistor-based digital hardware faces limitations, and new hardware alternatives are needed^{4,5}. One of the most notable technologies is the memristive dot product engine (MDPE), which performs large-scale vector-matrix multiplication

¹Department of Materials Science and Engineering, Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea. ²Department of Semiconductor Convergence, Chungnam National University (CNU), 99 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea. ✉e-mail: jwlee80@cnu.ac.kr; km.kim@kaist.ac.kr

(VMM) operations using a memristive crossbar array^{6–13}. The MDPE is fundamentally very energy-efficient, as it performs VMM operations based on physics principles (i.e., Kirchhoff's law and Ohm's law). Additionally, flexible crossbar arrays have been actively investigated recently, and they have potential in diverse edge applications, including image recognition¹⁴, classifying input data measured from sensors^{15–17}, and diagnosing biosignals such as ECGs^{18,19}. These studies demonstrate the potential of MDPE for edge AI, and it is crucial to validate and demonstrate these capabilities using reliably fabricated MDPE.

In MDPE fabrication, selecting a memristor is crucial, and we suggest that a charge trap memristor is the best choice based on its various proven advantages^{20–26}. It has a low current analog operation capability and electroforming-free characteristics, allowing high reliability and uniformity. Furthermore, the self-rectifying characteristics of the device enable large-array operation without complex transistor or selector processes, and the device is highly compatible with the conventional CMOS process. In addition, the fabrication of charge trap memristor-based MDPE on flexible substrates should proceed at a low temperature because a high thermal budget destroys the substrate, which has not been demonstrated.

Furthermore, once the MDPE is ready, it is crucial to explore detailed methodologies for systematically utilizing this hardware in appropriate applications. Given that AI technology offers significant advantages in analyzing signal patterns, its application in electrocardiogram (ECG) diagnosis is particularly well suited^{27,28}. Typical ECG diagnosis primarily relies on the experience and judgment of medical professionals. However, ECG diagnosis requires long-term monitoring of intermittently occurring abnormal signals, imposing a significant burden in terms of both accuracy and efficiency²⁹. Therefore, recent efforts have been made to address this issue with the help of AI. However, AI-based analysis often relies on data collected over short periods or requires remote processing, posing certain limitations. For

a more comprehensive diagnosis, ECG signal collection and analysis must be carried out continuously and in real time, making this an optimal application for edge AI. Figure 1a shows the real-time ECG diagnosis system, where the edge AI device collects ECG data and detects arrhythmia heartbeats in real time.

In this study, we introduce the charge trap memristor-embodying flexible MDPE (f-MDPE) and propose its utilization in an edge AI system for ECG diagnosis. We fabricated a 32×32 crossbar array on a polyimide (PI) substrate containing a low-temperature ($<180^\circ\text{C}$) processed charge trap memristor as the f-MDPE. Figure 1b highlights the key advantageous features of the f-MDPE for edge AI accelerators. The device exhibits highly self-rectifying behavior in a simple capacitor structure, enabling easy, selector-less passive crossbar array integration and allowing accurate programming and reading operations. Additionally, the electroforming-free nature of these materials ensures high device-to-device and cycle-to-cycle uniformity. The f-MDPE can be integrated in a conventional fab environment at low temperatures, compatible with the PI substrate, demonstrating its scalability with CMOS processing. Moreover, the f-MDPE shows high electrical and mechanical stability even under 5 mm bending conditions, highlighting its suitability for wearable applications. Last, its low current operation enables inference with very minimal energy. To demonstrate these advantages of the f-MDPE, we conducted a real-time ECG diagnosis. Here, we devised an algorithmic approach during the training process that accommodates hardware specifications such as non-Ohmic I - V characteristics and quantized weight properties. We built and trained a single-layer perceptron (SLP) using the MIT-BIH ECG database on software and mapped it onto the f-MDPE hardware. Then, we conducted laboratory-scale real-time ECG diagnosis using the f-MDPE, demonstrating its potential for in situ ECG inference, as it consumes only 0.3% of the energy compared to digital approaches with a classification accuracy of 93.5%, highlighting its potential for edge AI hardware.

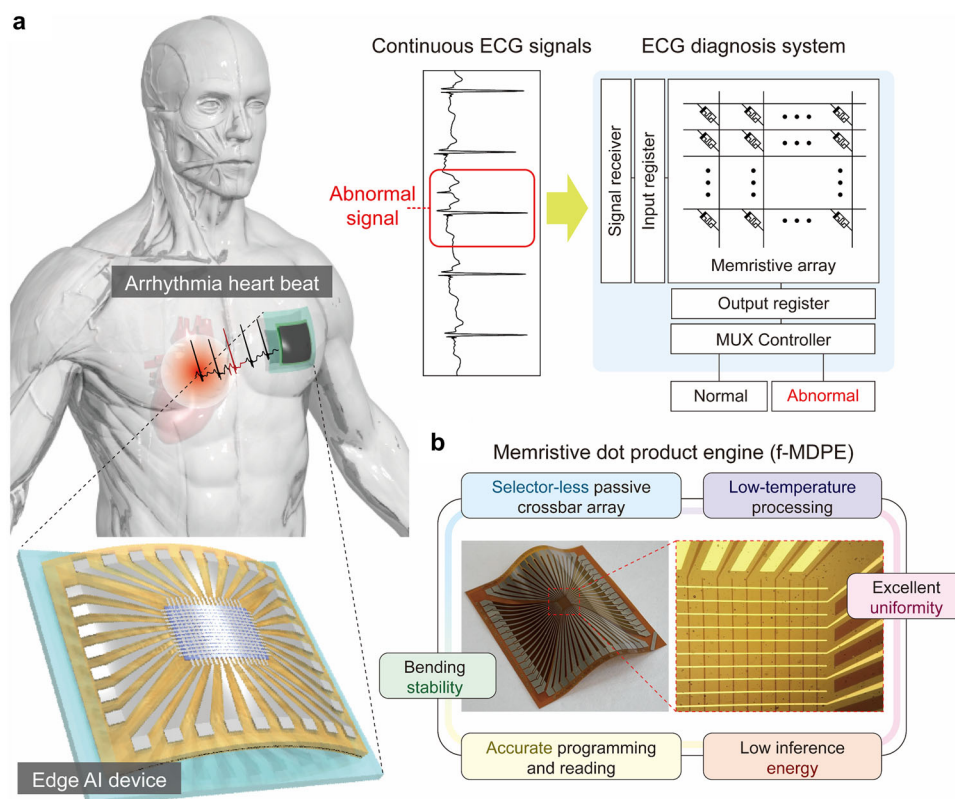


Fig. 1 | Overview of a flexible memristive dot product engine (f-MDPE) embodying a self-rectifying charge-trap memristor for ECG diagnosis. **a** Illustration of the real-time ECG monitoring system with the f-MDPE. **b** Key advantages of the f-MDPE for edge AI applications.

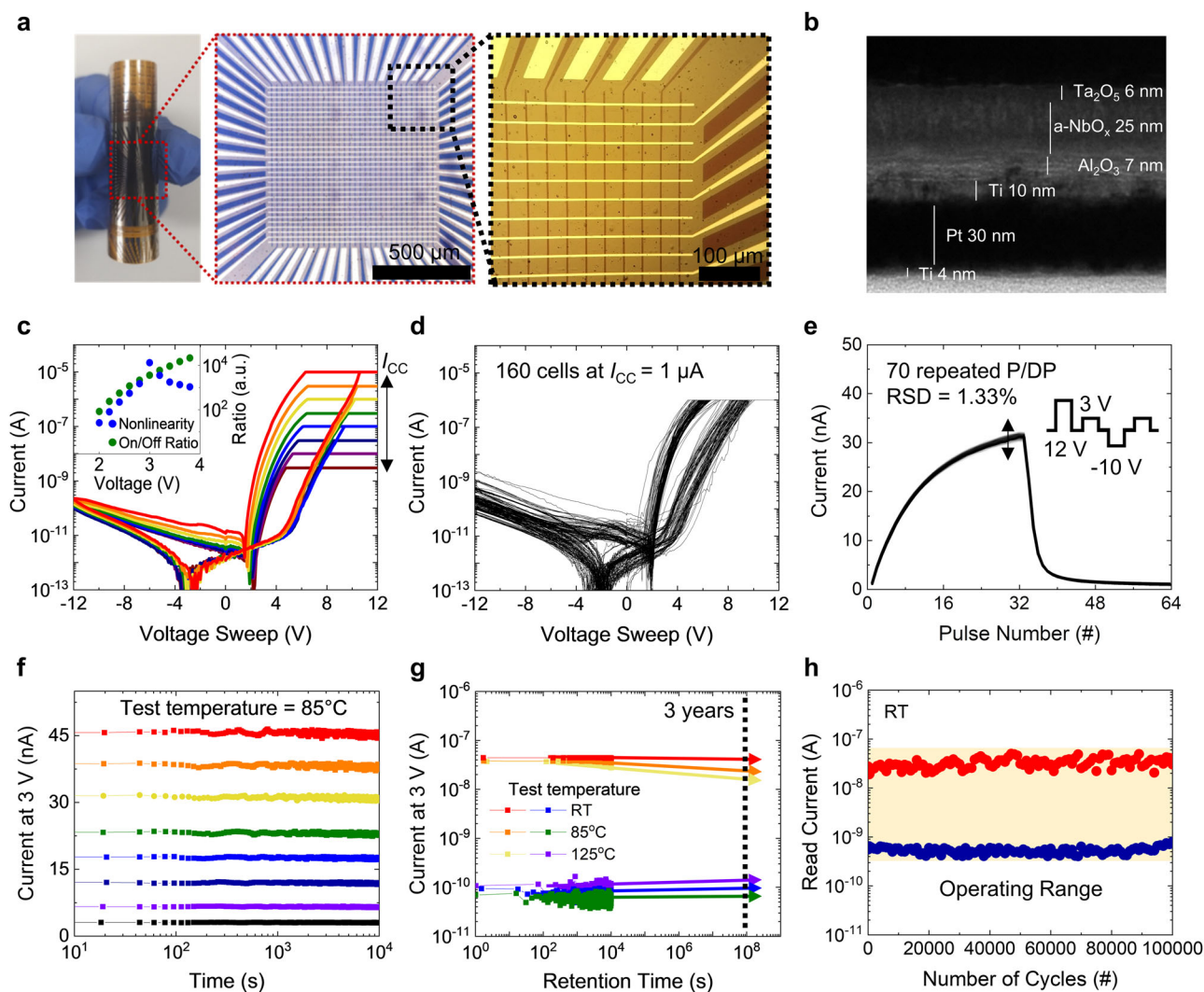


Fig. 2 | Electrical characteristics of the f-MDPE. a Optical images of the f-MDPE on a polyimide (PI) substrate. **b** Cross-sectional transmission electron microscopy (TEM) image of the device. A single representative TEM image is measured. **c** *I*-*V* curves of the device with various compliance currents (*I*_{CC}). The inset shows the nonlinearity and on/off ratio at 10 μA of *I*_{CC}. **d** *I*-*V* curves of 160 cells in the array

under 1 μA of *I*_{CC}. **e** Repeated potentiation and depression cycles of the device. **f** Retention characteristics of 8 conductance states at 85 °C. **g** Long-term retention characteristics at various temperatures. **h** Endurance performance for LRS (red) and HRS (blue) up to 10⁵ cycles at room temperature.

Results

Electrical characteristics, reliabilities, and bending durability of an f-MDPE array

Figure 2a shows optical images of the f-MDPE array. The layered structure of the charge trap memristor (Ti/Pt/Ti/Al₂O₃/NbO_x/Ta₂O₅/Pt, from bottom to top) was integrated on a flexible PI substrate. We have previously proposed this structure, but it was on a SiO₂/Si substrate²², so the device itself may not be novel. However, to implement this structure on a flexible substrate and overcome the slow operation speed of conventional devices, we introduced three key modifications in this study. First, since the PI substrate is vulnerable to high temperatures, we developed and applied low-temperature processes below 180 °C for all layers. Second, we replaced the oxidant for the Al₂O₃ tunneling layer from H₂O to O₃ to reduce the H-passivated traps^{30–33}, thereby improving the programming speed (see Supplementary Fig. 1 in the supplementary information (SI) for a detailed discussion on the changes in the Al₂O₃ deposition process and their impact on the programming speed). Third, we inserted a 30-nm-thick Pt layer in the middle of the Ti layer (i.e., Ti/Pt/Ti) to enhance the flexibility of the bottom electrode (BE), as Ti alone is

mechanically brittle. More details on the device fabrication process can be found in the Method section. The cross-sectional transmission electron microscopy (TEM) image in Fig. 2b confirms the thicknesses of the oxide layers: 7 nm for Al₂O₃, 25 nm for NbO_x, and 6 nm for Ta₂O₅. The NbO_x layer is non-stoichiometric and amorphous, which is a key factor that acts as a charge trap layer²². The Al₂O₃ and Ta₂O₅ layers act as tunneling and blocking oxides, respectively, stabilizing the trapped electrons.

The device exhibited analog set switching characteristics, with its conductance being tunable by a compliance current (*I*_{CC}), as shown in Fig. 2c, with a maximum on/off ratio of ~10⁴ at 3 V, as shown in the inset. The analog switching characteristics are attributed to the Schottky barrier height modulation accompanied by electron trapping and detrapping in the NbO_x layer²². Additionally, the device showed a rectifying ratio of ~10³ at a read voltage of ±3 V, as shown in the inset. Owing to its self-rectifying characteristics, the device may operate accurately even in a 1k × 1k array without sneak path disturbance^{34,35} (see Supplementary Fig. 2 in the SI for a discussion on the available array size and read margin of the f-MDPE array). To further evaluate the impact of voltage drop caused by wire resistance in a scaled-up device, we conducted cell voltage distribution simulations (see

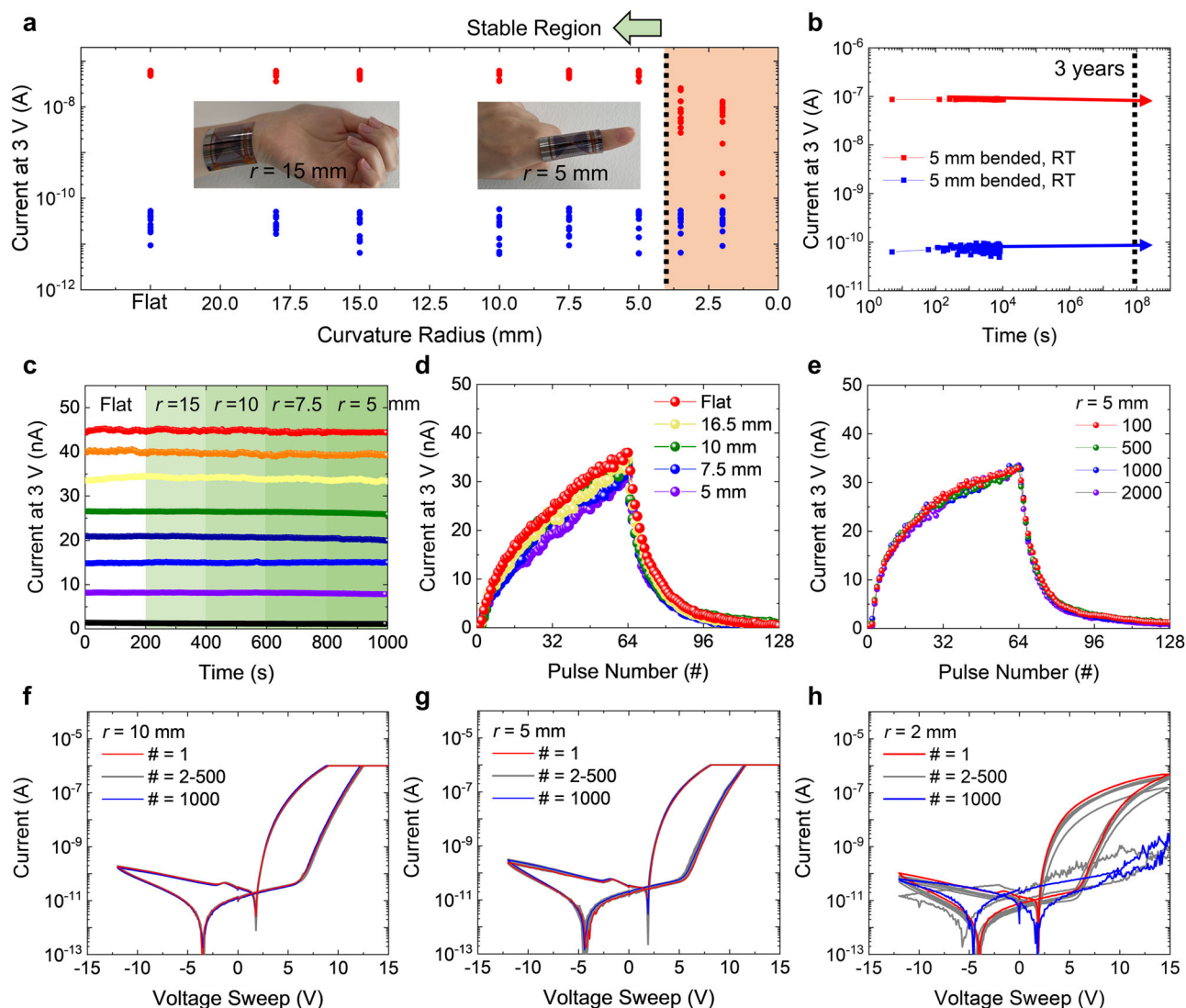


Fig. 3 | Bending stability of the f-MDPE. a Memory window for various curvature radii. **b** Long-term retention characteristics under a 5 mm curvature radius. **c** Retention characteristics of 8 conductance states and **d** potentiation and

depression under various curvature radii from flat to $r = 5$ mm. **e** Potentiation and depression after repeated bending cycles under a 5 mm radius. **f–h** I – V curves after repeated bending cycles under various curvature radii.

Supplementary Fig. 3). In the simulation, the conductance of each cell was randomly assigned a value between 1 nS and 15 nS, while the wire resistance between cells was assumed to be $10 \Omega^{36}$. A read voltage of 3 V was applied, and the voltage drop was analyzed across array sizes ranging from 32×32 to 1024×1024 . The results indicate that the wire resistance has little impact on f-MDPE operation, suggesting its highly reliable performance even in large-scale arrays.

Figure 2d shows I – V curves of 160 cells (32×5) in the array at an I_{CC} of 10^{-6} A, confirming a 100% yield and high array-level uniformity (see Supplementary Fig. 4 in the SI for the 160 raw I – V curves). Figure 2e shows 70 cycles of potentiation and depression (P/DP) curves from a representative cell. To obtain the P/DP characteristics, 100 μ s of +12 V, 100 μ s of –10 V, and 500 μ s of +3 V were used as the programming, erasing, and reading voltages, respectively, as shown in the inset. The results demonstrate the high feasibility and uniformity of these devices when they are suitable for analog devices.

The retention characteristics of eight different conductance states were confirmed at 85 °C, as shown in Fig. 2f. All conductance states were stable up to 10^4 s without significant changes. This gives an operating conductance window from 0.1 nS to 15 nS at 3 V. Figure 2g shows the retention characteristics of the high resistance

state (HRS) and low resistance state (LRS) for 10^4 s at different temperatures (RT, 85 °C, and 125 °C). The Arrhenius plot suggested that both the HRS and LRS could be maintained for more than 3 years (see Supplementary Fig. 5 in the SI for the Arrhenius plot with the failure time at 20% conductance change). Additionally, Fig. 2h shows the endurance of the device for up to 10^5 cycles, using 5 ms of +12 V pulse for programming and 5 ms of –10 V pulse for erasing, respectively. Here, the on/off ratio was smaller than in Fig. 2g, where a DC sweep was used, due to the shorter pulse duration. Although this is not the full on/off ratio, we still achieve an on/off ratio of approximately 70, which we believe is sufficient for verifying the endurance characteristic. These results confirm the practical feasibility of the f-MDPE approach.

We also examined the bending durability of the f-MDPE for wearable applications. Figure 3a shows the memory windows of ten randomly selected cells at various bending radii. The results showed that the cells operated well down to a curvature radius (r) of 5 mm but deteriorated to less than 5 mm. This is likely due to an increase in electrode resistance under bending conditions, which results in the inability to apply voltage to the cell. Nevertheless, considering that the curvatures of the human body are greater than 5 mm (insets,

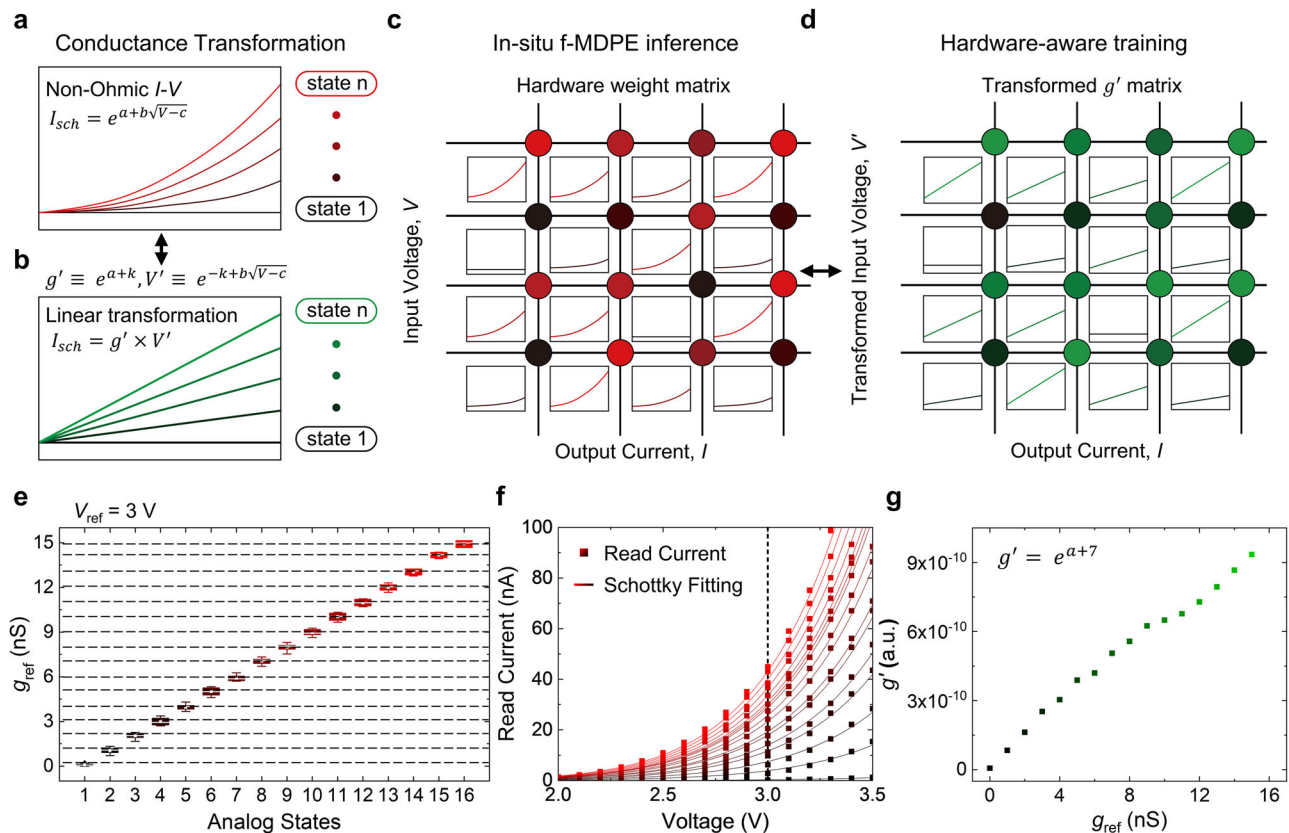


Fig. 4 | Hardware-aware training for addressing non-Ohmic conduction behavior. **a** Output currents of multiple states as a function of the input voltages of non-Ohmic devices following Schottky conduction (I_{sch}). **b** Output currents represented by the transformed voltage, V' . g' remains constant over the input voltages. **c** Crossbar array after linear transformation (V to V' and g to g'), enabling a formal VMM operation in the f-MDPE. **d** Crossbar array with non-Ohmic cells. In this case, the conductance changes according to the input voltage. **e** Four bits (i.e., 16 levels)

of multilevel conductance states at the reference voltage, 3 V, with 15 cells. Note that box plots are defined by the minimum, 25th percentile (Q1), median, 75th percentile (Q3), and maximum values, with the box indicating the interquartile range (IQR) and the whiskers representing the full data range. **f** I - V curves at the read voltage region at each analog state. **g** g' obtained by extracting values for each analog state using a compensation constant $k = 7$.

$r = 15$ mm at a wrist and $r = 5$ mm at a finger), the stability at $r = 5$ mm is practically meaningful. In addition, retention characteristics were examined at $r = 5$ mm, revealing that the bending stress did not negatively affect the stability over 3 years, as shown in Fig. 3b. Figure 3c shows the high stability of multiple conductance states while changing curvatures from flat to $r = 5$ mm for 200 s at each curvature (total 10^3 s). For this measurement, we first programmed the device into specific conductance states and measured the retention for 200 s at each bending radius while gradually decreasing the radius to 15 mm, 10 mm, 7.5 mm, and 5 mm. We repeated this measurement for 8 different conductance states.

Furthermore, the device exhibited reliable synaptic characteristics under bending conditions. Figure 3d shows the P/DP characteristics under various bending conditions (i.e., flat, 16.5, 10, 7.5, and 5 mm bending radii). Figure 3e confirms the high bending durability of the synaptic characteristics over bending cycles (100, 500, 1000, and 2000 cycles) at $r = 5$ mm.

Next, the bending cycle endurance for the electrical characteristics was investigated. Figure 2f-h shows the I - V characteristics during 10^3 bending cycles with various bending radii. The device operated well under $r = 10$ mm (Fig. 3f) and $r = 5$ mm (Fig. 3g) while maintaining a wide memory window. However, significant degradation was observed at $r = 2$ mm (Fig. 3h); the LRS degraded gradually after up to 500 cycles and eventually collapsed to the HRS after 1000 cycles. This was also associated with the fatality of the electrode, which is consistent with the previous results in Fig. 3a. In conclusion, the device exhibited sufficient electrical and mechanical stability.

Hardware-aware training for addressing non-Ohmic conduction behavior

Neural networks execute the multiplication of an input vector and a weight matrix, known as the VMM. The operations can be efficiently conducted in the MDPE through Kirchhoff's law and Ohm's law, which can be given by $I_j = \sum_{i=1}^n g_{ij} \cdot V_i$, where I is the output current, g is the conductance, V is the input voltage, i and j are row and column indices, respectively, and n is the number of inputs. In conventional MDPEs, the conductance (g_{ij}) is assumed to be constant regardless of the input voltage (i.e., Ohmic conduction), so multiple input voltages can be easily applied^{7,37}.

However, the proposed f-MDPE exhibited a nonlinear current response to input voltages stemming from non-Ohmic conduction. This is not a specific problem in our device but is an inherent challenge in low-current devices, where the conduction mechanism is non-Ohmic^{38,39}. Therefore, the effective conductance varies depending on the input voltage, as depicted in Fig. 4a. In this case, the trained weight using the conventional method cannot be directly mapped to g . Instead, training and inference must be performed considering that the weight is a function of the input voltage corresponding to the non-Ohmic behavior of devices⁴⁰. However, it necessitates additional partial derivative operations of the weight matrix during back-propagation, significantly increasing the computational complexity and thus presenting a highly intricate dilemma.

Here, we introduce a hardware-aware training approach to address this non-Ohmic behavior issue. This process involves converting the nonlinear weight function into a linear function

during training mathematically, allowing for effective network training. Then, the trained weights, reflecting the nonlinear characteristics, are converted to conductance, which can be directly mapped onto f-MDPE, enabling inference.

The first step of this process is to determine the non-Ohmic function. The f-MDPE device followed Schottky conduction (see Supplementary Fig. 6 in the SI for the detailed conduction mechanism fitting results). Therefore, the I - V characteristics of the device can be expressed by the following Schottky equation

$$I = I_{sch} = AT^2 \exp \left[\frac{-q(\Phi_b - \sqrt{-qE/4\pi\epsilon_r\epsilon_0})}{kT} \right] \quad (1)$$

where A is a constant, $q\Phi_b$ is the Schottky barrier height, T is the temperature, q is the electronic charge, E is the electric field, ϵ_0 is the permittivity in vacuum, ϵ_r is the optical dielectric constant, and k is the Boltzmann constant. Equation (1) can be simplified to the form of $I = \exp(a + b\sqrt{V - c})$, where a is the variable that depends on the conductance state related to Φ , and b and c are constants, assuming a constant T . Thus, the nonlinear current can be defined as a function of a and V :

$$I(a, V) = \exp(a) \cdot \exp(b\sqrt{V - c}) \quad (2)$$

Here, we put a compensation constant k that scales the value of each exponential term accordingly, addressing potential issues with scale consistency and ensuring that the neural network operates within a stable and efficient regime (see Supplementary Table 1 for scale between g' and V' with constant k).

$$I(a, V) = \exp(a + k) \cdot \exp(-k + b\sqrt{V - c}) \quad (3)$$

Then, by defining $g' = \exp(a + k)$ and $V' = \exp(-k + b\sqrt{V - c})$, Eq. (3) can be reformulated as a linear multiplication function:

$$I = g' \cdot V' \quad (4)$$

Consequently, the output currents can be expressed as a linear function of V' with a slope of g' , a function of the device state a , as depicted in Fig. 4b. Then, the current output at the J_{th} neuron during VMM can be given by the simple dot product between g' and V' matrices:

$$I_j = \sum_{i=1}^n g'_{ij} \cdot V'_i \quad (5)$$

Equation (5) indicates that the f-MDPE can perform the VMM by substituting g with g' and V with V' for neural network training. Consequently, non-Ohmic I - V curves can be converted to Ohmic I - V' curves (see Supplementary Fig. 7 for transformed I - V' curves), where the slope represents g' . By transforming linear VMM, unnecessary additional partial derivatives of Schottky functions during back-propagation in the training process are avoided, enabling energy-efficient neural network learning (see Supplementary Table 2 for the comparison of the network model, output form, and weight gradient function between three cases: ideal software case, hardware using Ohmic memristor case, and hardware using Schottky memristor case). Here, V were transformed to V' using the hardware parameters, retaining the features of the original signal (see Supplementary Fig. 8 for an example of the transformation from V to V'). Note that both V and V' are directly generated by receiving external signals through a controller unit. Therefore, the conversion from V to V' does not impose any additional circuit burden.

After hardware-aware training, one can obtain g'_{ij} matrix, as depicted in Fig. 4c. The next step is converting g'_{ij} to g_{ij} and mapping

g_{ij} onto f-MDPE for inference, as depicted in Fig. 4d. To achieve this, the relationship between g' and g must be defined, which can be obtained through the following process.

To consider the variation in analog conductance states, we defined 16 quantized conductance states as available weights, whose conductance at 3 V (g_{ref}) ranged from 0.1 nS to 15 nS with a 1 nS interval, as shown in Fig. 4e. The distinct g_{ref} were observed without overlap between states, indicating their suitability for the quantized weights of the neural network. Figure 4f shows the read currents for the 16 quantized states at various input voltages (square dots) alongside the Schottky fitting curves (solid lines) (see Supplementary Fig. 9 in the SI for the I - V curves of 16 states obtained by 20 devices demonstrating the stability of each state, and see Supplementary Fig. 10 and Supplementary Table 1 for the Schottky fitting results). For hardware-aware training, hardware parameters for each state are extracted from the fitting curves, and g' values are obtained from the hardware parameters. Figure 4g shows $g' - g_{ref}$ plot, which is used for converting the trained g' to g_{ref} to map the g_{ref} onto f-MDPE.

In summary, after obtaining the trained g' matrix, the corresponding g_{ref} map can be transferred to the f-MDPE, and by using V as input, the f-MDPE can perform the VMM. This hardware-aware approach can generally be applied to all memristors exhibiting non-Ohmic conduction behavior; furthermore, it can address any hardware nonideality issues⁴¹.

ECG dataset classification using the f-MDPE array

Here, we aim to demonstrate the feasibility of using f-MDPE as edge AI hardware by utilizing it for real-time ECG diagnosis. Previous studies have used AI for ECG signal analysis and proven its effectiveness, demonstrating that its implementation in real-time edge environments is novel in this work^{42–46}. Figure 5a illustrates the proposed ECG diagnosis process involving (i) ECG signal sensing, (ii) signal preprocessing, (iii) f-MDPE inference, and (iv) ECG diagnosis using the trained f-MDPE array. For training, we adopted the ECG dataset from the MIT-BIH arrhythmia database, which classifies the ECG signals into five categories⁴⁷: normal (N), supraventricular ectopic beat (S), ventricular ectopic beat (V), unknown beat (Q), and fusion beat (F). Among these categories, only N represents normal signals, which are observed most frequently, while the others denote intermittent abnormal signals. The MIT-BIH database contains raw data from continuous ECG signals. To utilize these data as input data for the f-MDPE, it is necessary to preprocess them into the ECG dataset and adjust them appropriately to fit the f-MDPE. Therefore, we extracted individual ECG signals and resized them into 32-time frames to align them with the number of input terminals of the 32×32 f-MDPE. Additionally, we adjusted the amplitude corresponding to the input voltages from 2.0 to 3.5 V (see Supplementary Table 3 for details on the data preprocessing procedure). Subsequently, we utilized this ECG dataset to train a single-layer perceptron (SLP) with a size of 32×5 on software, enabling it to distinguish between 5 categories of ECG patterns using an input size of 32. During training, we also implemented hardware-aware training methodologies reflecting non-Ohmic conduction behavior and quantized conductance characteristics, as described in the previous section (Fig. 4). The trained weight matrix contains both positive and negative values, which cannot be directly mapped as conductance, allowing only positive values. In such cases, a common approach is to represent one weight using a pair of columns, each representing positive or negative weight values^{8,48,49}. Thus, the 32×5 weight matrix was reconstructed to a 32×10 matrix. Figure 5b shows the weight matrix trained by the SLP with a size of 32×10 , where the weight values include both positive (red) and negative (blue) values, each having 16 quantized levels. The zeroth state (Level 0) denotes an unprogrammed conductance state (i.e., initial state, whose conductance was 1.6 pS, which is much smaller than the 0.1 nS of Level 1). Figure 5c shows the conversion of these conductance values into target current values at

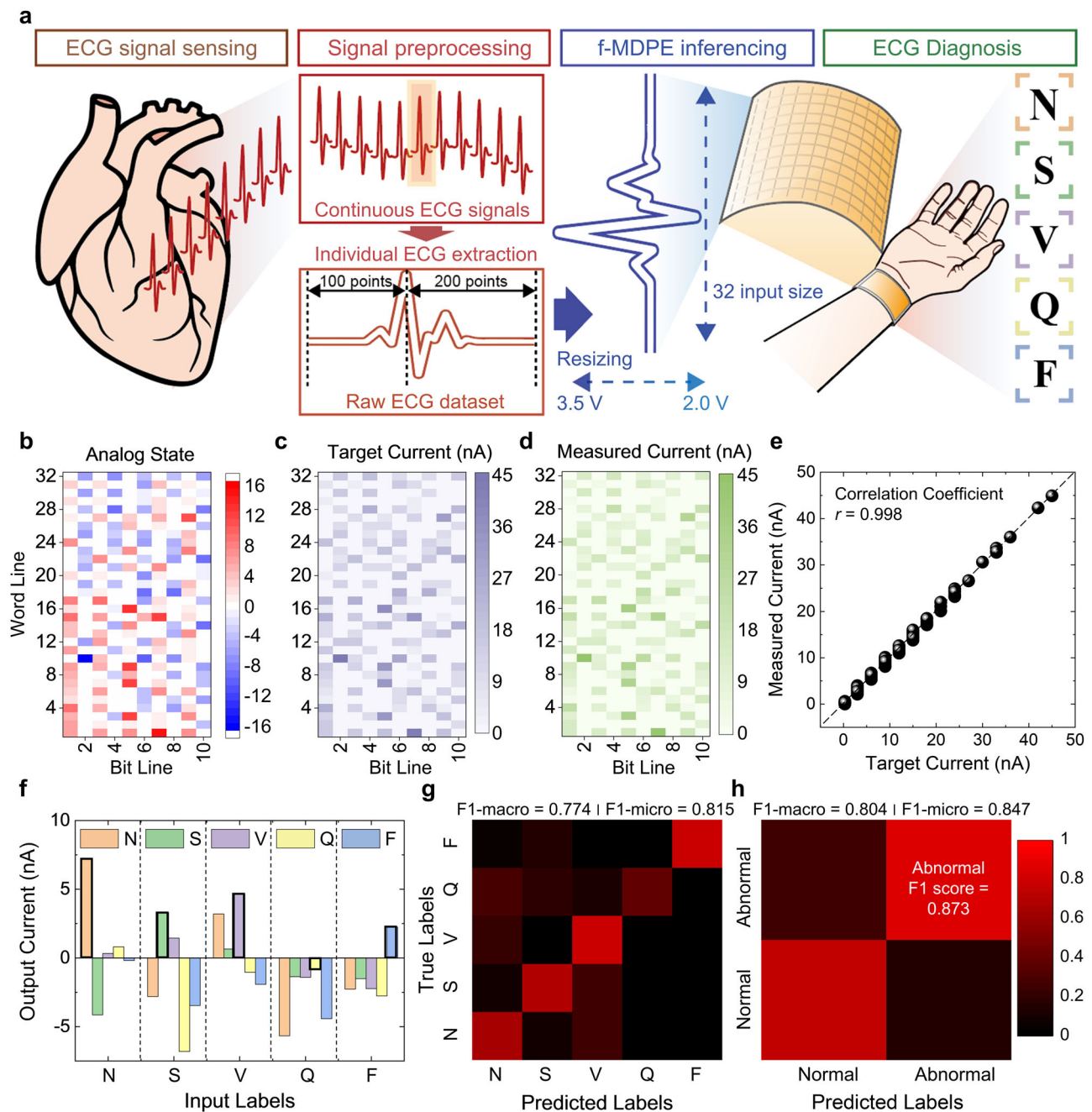


Fig. 5 | The f-MDPE-based hardware system for ECG heartbeat diagnosis.

a Schematics of in-situ inference with a pretrained f-MDPE array for ECG diagnosis. **b** Pretrained weight matrix of a 32×5 sized network, using positive and negative weights separately. **c** Target current at 3 V of the transferred weight matrix in (b). **d** Measured current at 3 V of the pretrained f-MDPE array. **e** Error distribution between the target and measured currents at each single point. **f** Summed output

currents for the given input heartbeats from the MIT-BIH heartbeat dataset represented as N, S, V, Q, and F. The neuron indicating each heartbeat has the highest output current, indicating a correct classification. **g**, **h** Confusion matrices of the network for 5-category classification (**g**) and normal/abnormal classification (**h**).

3 V for actual mapping onto the f-MDPE. Here, Levels 1–16 are converted from 0.3 nA to 45 nA at 3 V.

Next, we programmed the f-MDPE with the prepared weight matrix in Fig. 5b and examined its programmed accuracy. For this weight mapping, we used ISPP (incremental step pulse programming)-type programming and verification scheme to accurately programming the cells⁵⁰. Figure 5d shows the measured currents of the programmed f-MDPE read at 3 V. Figure 5e compares the target current (I_{target}) of Fig. 5c with the measured read current (I_{measured}) of Fig. 5d for the entire 320 weight cells. The correlation coefficient between I_{target} and I_{measured} was 0.998, confirming the accurate programming of the

trained weight matrix onto the f-MDPE array (see Supplementary Fig. 11 for more detailed statistics of the target and measured current difference).

Before we used the f-MDPE model for real-time diagnosis, we tested its ability to perform an inference test on the ECG dataset. Figure 5f shows five examples of outputs selected from different categories. Note that the total output current is the difference between the currents from the positive and negative weights. In these examples, the highest output corresponding to the correct category indicates that the inputs are accurately distinguished.

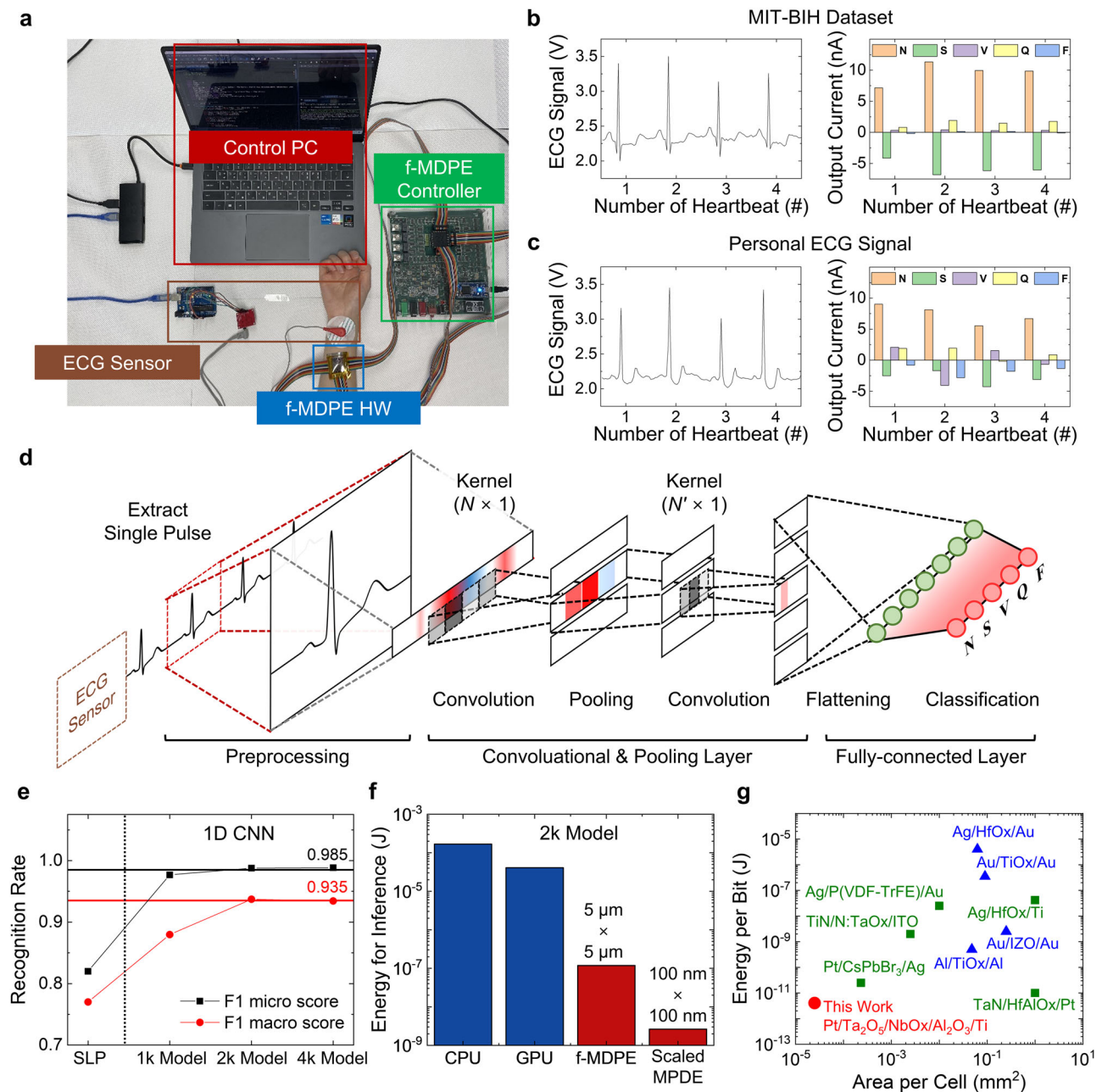


Fig. 6 | Experimental demonstration and scalability of f-MDPE for ECG diagnosis. **a** Photograph of the integrated system: ECG sensor, f-MDPE, and f-MDPE controller integrated on a control PC. ECG diagnosis of N pulses from the **b** MIT-BIH dataset and **c** personal signals acquired from the ECG sensor. **d** Schematics of the scaled-up 1D convolutional neural network (1D CNN). **e** Recognition rates of ECG diagnosis for the proposed 1D CNN architectures. The recognition rate is saturated

in the 2k model. **f** Energy consumption comparisons of the f-MDPE algorithm and conventional processors (CPU and GPU) for inferring a single heartbeat signal from the 2k model. **g** Energy consumption per bit and area per cell comparisons of the f-MDPE and flexible or stretchable memristive arrays. The green squares represent flexible arrays, and the blue rectangles represent stretchable arrays.

The estimated performance of the f-MDPE method for inferring signals from the ECG dataset was investigated. Figure 5g shows a confusion map of the inference results for the 4255 test dataset. The F1-macro score was 0.774, while the F1-micro score was 0.815 (see Supplementary Note for the details on the F1 score definition). This score represents the recognition rate for all the data. However, in ECG diagnosis, abnormal signals can be misidentified among abnormal signals, so the ECG pattern can be categorized as normal or abnormal, where abnormal includes S, V, Q, or F. Figure 5h shows the reconstructed confusion map, which gives an F1-macro score of 0.804 and an F1-micro score of 0.847. Furthermore, considering that classifying abnormalities is the most crucial aspect of diagnosis, when focusing on

abnormal signal detection, the accuracy of the abnormal dataset was 0.873, which is an attractive number considering that such abnormal ECG patterns may occur several times per day or hour⁵¹. In summary, although the f-MDPE is a prototype hardware with only a 32×32 size capable of driving the SLP, it can efficiently diagnose the ECG.

A real-time ECG diagnosis demonstration

Figure 6a shows a photograph of a laboratory-scale real-time ECG diagnosis system comprising a control PC, an f-MDPE controller, an f-MDPE, and an ECG sensor, where the ECG sensor and the f-MDPE are attached to a human wrist (see Supplementary Fig. 12 for the detailed electrical connection scheme between the f-MDPE and the controller).

The f-MDPE was programmed to perform only inference operations based on the weight matrix trained ex-situ through software. The control PC receives the ECG signals from the ECG sensor, preprocesses them to the input voltage signal for f-MDPE inference, and sends instructions to the f-MDPE controller. The f-MDPE controller sends the input voltages to the f-MDPE, receives the output currents, and returns them to the PC. Then, the PC determines the category of the ECG signals (see Supplementary Fig. 13 for the flowchart of the real-time ECG diagnosis system).

We attempted two real-time tests in this system: (1) using the MIT-BIH ECG dataset for reference (w/o the ECG sensor part) and (2) using the preprocessed ECG data collected in real-time from ECG sensors. The preprocessing of the real-time ECG data was performed via software (Python) and included signal denoising, resizing, and peak detection to ensure accurate dataset generation (see Supplementary Video for the real-time ECG diagnosis demonstration). Figure 6b shows an example of four consecutive N signals from the MIT-BIH dataset (left panel) and the f-MDPE's output currents for each signal (right panel), confirming that the system effectively categorized the input signals. Figure 6c shows the preprocessed real-time ECG signals collected from the ECG sensor (left panel) and the collected output currents (right panel). The results show that the system can accurately distinguish N pulses from a regular person. Note that the purpose of this testing is not to conduct actual diagnoses but to demonstrate the operation of the entire system. Therefore, we tested this approach with a regular person and observed N pulses.

A scale-up feasibility demonstration via virtual f-MDPE hardware

We adopted a small network size of 32×5 and a simple network structure of SLP to fit the f-MDPE. As a result, we achieved an F1 macro score of 0.774, indicating the potential of the ECG diagnosis system. However, there is also a need to achieve better performance at larger sizes. Therefore, we simulated a large-scale neural network using virtual f-MDPE hardware built into software and evaluated the behavior of the f-MDPE. Additionally, we adopted a method of partitioning a single piece of hardware to operate a multilayer perceptron. Here, we employed a one-dimensional convolutional neural network (1D CNN), whose network structure is shown in Fig. 6d, as it is known to be the best for pattern recognition tasks⁵². We built three 1D CNN network models, namely, 1k, 2k, and 4k, referring to the available number of weights in 32×32 (1k), 64×32 (2k), and 64×64 (4k) arrays of the virtual f-MDPE, respectively (details of 1D CNN architectures across varying network sizes are described in Supplementary Table 4). Although the 2k and 4k arrays were not physically fabricated, they were simulated assuming that the 1k-sized f-MDPE array was directly scaled up. This virtual scaling approach is widely applied in other studies as well^{13,53}, as it provides valuable insights into large-scale performance with high reliability in advance. Figure 6e shows the F1-macro score and F1-micro score for the proposed 1D CNN models. The table also includes the experimental SLP scores for reference. The results suggested that the recognition rate was saturated in the 2k model, with 0.99 for the F1-micro score and 0.94 for the F1-macro score. Supplementary Fig. 14 shows the confusion matrix obtained by the 2k model, revealing its high performance. The hardware-aware training method is also applicable to 2D CNNs, indicating its broad applicability across various applications (see Supplementary Fig. 15 for more details of hardware-aware training models).

Figure 6f shows the estimated energy consumption for ECG inference on various computing platforms when the 2k model is used. Note that this comparison focuses on the energy consumption of running neural networks, excluding other sources of energy consumption, such as ECG sensors. The energy consumption per inference by the conventional CPU and GPU approaches was 168 μ J and 41 μ J, respectively^{13,54}. However, the f-MDPE with a 5- μ m-line width required

only 120 nJ, far below that of conventional computing approaches. This energy consumption can be further reduced by scaling the array with a 100-nm-line width^{55,56}. We anticipate that the energy consumption can reach 2.64 nJ, outperforming GPU-based inference by more than four orders of magnitude (detailed energy consumption and latency are described in Supplementary Figs. 16–18 and Tables 5 and 6). Furthermore, compared with other flexible or stretchable memristive arrays, the f-MDPE offers advantages in terms of energy efficiency and size (Fig. 6g)^{15,16,18,57–62}.

Discussion

In summary, we reported energy-efficient neuromorphic applications utilizing the f-MDPE, characterized by its highly nonlinear and self-rectifying characteristics with a low-current operation, suitable for energy-efficient and reliable vector-matrix multiplication operations. Also, the device exhibited mechanical durability up to a 5-mm bending curvature, making it suitable for wearable applications. We successfully demonstrated the use of f-MDPE for in situ ECG classification by applying a hardware-aware training method that overcomes hardware characteristics such as non-Ohmic conduction and weight quantization. Furthermore, we simulated the proposed method using scaled-up 1D CNNs, confirming its feasibility for real-time ECG diagnosis with high diagnostic accuracy.

In edge computing, utilizing edge hardware plays a dominant role in reducing overall energy consumption. For example, in real-time ECG diagnosis, the overall process can be summarized as follows: (1) ECG signal acquisition, (2) preprocessing, (3) inferencing, and optionally, (4) wireless transmission (e.g., via Bluetooth). A single ECG pulse acquisition (assuming a 1-s acquisition) consumes around 6.1–17.4 μ J^{63–65}. The preprocessing stage requires negligible energy, around tens of nJ⁶⁶. Inferencing using a GPU requires 42 μ J, but this energy consumption can be reduced to 120 nJ by using f-MDPE, significantly alleviating this energy burden. Furthermore, edge computing processes data locally, minimizing external signal transmission. This drastically reduces the high energy consumption required for wireless transmission (real-time Bluetooth transmission typically consumes 22.7 mW⁶⁷). As such, the f-MDPE aligns well with future analog signal processing at edge devices, providing energy-efficient artificial intelligence solutions.

While the f-MDPE shows promise for real-time and long-term signal monitoring due to its energy efficiency and portability, several challenges remain to be addressed for further advancements. First, for its use as a fully neuromorphic processor, its functionality should be expanded beyond inference to training. To achieve this, a control unit must be integrated, which would require advancements in wearable electronics. Even in this scenario, the low programming energy of f-MDPE would remain a significant advantage. Second, improving the higher operating voltage and slower switching speed of f-MDPE through the development of materials and process optimizations could further enhance its usability. If these aspects can be enhanced in the future, it would significantly broaden the applicability of f-MDPE across a wider range of applications.

Methods

Device fabrication

For an f-MDPE array, the integrated device consisted of Pt/Ta₂O₅/NbO_x/Al₂O₃/Ti/Pt/Ti on a polyimide (PI) substrate. Polydimethylsiloxane (PDMS) was spin-coated on a 1-mm-thick glass substrate, and a 50- μ m-thick PI film was attached to it. Next, the f-MDPE device was integrated by the following steps: First, an adhesive of 4-nm-thick Ti followed by a 30-nm-thick Pt insertion and a 10-nm-thick Ti bottom electrode were deposited by e-beam evaporation (KVE-E2000) without breaking the vacuum on a bottom electrode pattern formed by a mask aligner (Midas MDA-600S). These layers were then patterned by a lift-off process. Next, a 7-nm-thick Al₂O₃ layer was deposited by thermal atomic layer deposition (ALD) at 180 °C using trimethylaluminum and O₃ as the Al

precursor and oxygen source, respectively. The NbO_x layer was deposited by reactive sputtering (Daeki Hitech cosputtering system) at 170 °C in an Ar and O₂ mixed gas ambient environment using an Nb target. Next, the Ta₂O₅ layer was deposited using plasma-enhanced ALD at 180 °C using Tris(diethylamido) (tert-butylimido)tantalum(V) and O₂ plasma for the Ta precursor and the oxidant, respectively. Then, a 60-nm-thick Pt top electrode was deposited by e-beam evaporation and patterned by a lift-off process. Finally, the f-MDPE was detached from the glass substrate. The line width for the active area in the array was 5 μm.

Electrical measurements

Electrical characterization was performed using a Keithley 4200A-SCS and an ArcOne f-MDPE controller. The pulse measurements were obtained by a Keithley 4200A-SCS and an ArcOne. During the measurement, the TE was biased, and the BE was grounded.

Measurement of personal ECG signals

This study received an exemption confirmation from the Institutional Review Board of Korea Advanced Institute of Science and Technology (KAIST IRB) (No. 2025-03-1-460). The personal ECG signals of the main authors, Y. Lee and G. Kim, were obtained by utilizing a commercial ECG sensor. Written consent was obtained from all participants prior to data collection. The purpose of this experiment was solely to collect data for hardware validation. No biometric data were stored, trained, or reused, and the study is entirely unrelated to any medical procedures. In selecting participants for the experiment, factors such as the number of volunteers, sex, and age were not considered.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All the data supporting the findings of this study are available within the article and its Supplementary Information. The raw data generated in this study are available from the corresponding author upon request.

Code availability

The simulation results were processed using Python, LTSpice, and Arduino software. All the relevant codes are available from the corresponding author upon request.

References

- LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
- Wu, H., Yao, P., Gao, B. & Qian, H. Multiplication on the edge. *Nat. Electron.* **1**, 8–9 (2018).
- Kwon, D. A. et al. Body-temperature softening electronic ink for additive manufacturing of transformative bioelectronics via direct writing. *Sci. Adv.* **10**, eadn1186 (2024).
- Conklin, A. A. & Kumar, S. Solving the big computing problems in the twenty-first century. *Nat. Electron.* **6**, 464–466 (2023).
- Lu et al. High-speed emerging memories for AI hardware accelerators. *Nat. Rev. Electr. Eng.* **1**, 24–34 (2024).
- Harabi, K. E. et al. A memristor-based Bayesian machine. *Nat. Electron.* **6**, 52–63 (2023).
- Hu, M. et al. Memristor-based analog computation and neural network classification with a dot product engine. *Adv. Mater.* **30**, 1705914 (2018).
- Li, C. et al. Efficient and self-adaptive in-situ learning in multilayer memristor neural networks. *Nat. Commun.* **9**, 2385 (2018).
- Li, C. et al. Long short-term memory networks in memristor crossbar arrays. *Nat. Mach. Intell.* **1**, 49–57 (2019).
- Ning, H. et al. An in-memory computing architecture based on a duplex two-dimensional material structure for in situ machine learning. *Nat. Nanotechnol.* **18**, 493–500 (2023).
- Park, J. et al. Implementation of convolutional neural networks in memristor crossbar arrays with binary activation and weight quantization. *ACS Appl. Mater. Interfaces* **16**, 1054–1065 (2024).
- Wang, S. et al. Echo state graph neural networks with analogue random resistive memory arrays. *Nat. Mach. Intell.* **5**, 104–113 (2023).
- Yi, S.-i., Kendall, J. D., Williams, R. S. & Kumar, S. Activity-difference training of deep neural networks using memristor crossbars. *Nat. Electron.* **6**, 45–51 (2023).
- Huang, J. et al. Flexible, transparent, and wafer-scale artificial synapse array based on TiO_x/Ti3C2Tx film for neuromorphic computing. *Adv. Mater.* **35**, 2303737 (2023).
- Cho, H. et al. Real-time finger motion recognition using skin-conformable electronics. *Nat. Electron.* **6**, 619–629 (2023).
- Kim, S. H. et al. A bioinspired stretchable sensory-neuromorphic system. *Adv. Mater.* **33**, 2104690 (2021).
- Liu, L. et al. Stretchable neuromorphic transistor that combines multisensing and information processing for epidermal gesture recognition. *ACS Nano* **16**, 2282–2291 (2022).
- Ham, S. et al. One-dimensional organic artificial multi-synapses enabling electronic textile neural network for wearable neuromorphic applications. *Sci. Adv.* **6**, eaba1178 (2020).
- Liu, Y. et al. Robust memristive fiber for woven textile memristor. *Adv. Funct. Mater.* **32**, 2201510 (2022).
- Gao, S. et al. Forming-free and self-rectifying resistive switching of the simple Pt/TaO_x/n-Si structure for access device-free high-density memory application. *Nanoscale* **7**, 6031–6038 (2015).
- Kim, K. M. et al. Low-power, self-rectifying, and forming-free memristor with an asymmetric programming voltage for a high-density crossbar application. *Nano Lett.* **16**, 6724–6732 (2016).
- Kim, G. et al. Retention secured nonlinear and self-rectifying analog charge trap memristor for energy-efficient neuromorphic hardware. *Adv. Sci.* **10**, 2205654 (2023).
- Yoon, J. H. et al. Pt/Ta₂O₅/HfO_{2-x}/Ti resistive switching memory competing with multilevel NAND flash. *Adv. Mater.* **27**, 3811–3816 (2015).
- Yoon, J. H. et al. Highly uniform, electroforming-free, and self-rectifying resistive memory in the Pt/Ta₂O₅/HfO_{2-x}/TiN structure. *Adv. Funct. Mater.* **24**, 5086–5095 (2014).
- Kim, G. et al. Threshold modulative artificial GABAergic nociceptor. *Adv. Mater.* **35**, 2304148 (2023).
- Kim, Y. et al. Nociceptive memristor. *Adv. Mater.* **30**, 1704320 (2018).
- Wu, J., Li, F., Chen, Z., Pu, Y. & Zhan, M. A neural network-based ECG classification processor with exploitation of heartbeat similarity. *IEEE Access* **7**, 172774–172782 (2019).
- Xu, C., Solomon, S. A. & Gao, W. Artificial intelligence-powered electronic skin. *Nat. Mach. Intell.* **5**, 1344–1355 (2023).
- Chua, S. K. et al. Comparison of arrhythmia detection by 24-hour Holter and 14-day continuous electrocardiography patch monitoring. *Acta Cardiol. Sin.* **36**, 251 (2020).
- Xiang, Y. et al. Oxidation precursor dependence of atomic layer deposited Al₂O₃ films in a-Si:H(i)/Al₂O₃ surface passivation stacks. *Nanoscale Res. Lett.* **10**, 1 (2015).
- Ok, K. C. et al. Effect of alumina buffers on the stability of top-gate amorphous InGaZnO thin-film transistors on flexible substrates. *IEEE Electron Device Lett.* **36**, 917 (2015).
- Goldstein, D. N., McCormick, J. A. & George, S. M. Al₂O₃ atomic layer deposition with trimethylaluminum and ozone studied by in situ transmission FTIR spectroscopy and quadrupole mass spectrometry. *J. Phys. Chem. C* **112**, 19530 (2008).

33. Yeom, H. I. et al. High-performance oxide thin-film diode and its conduction mechanism based on ALD-assisted interface engineering. *J. Mater. Chem. C* **11**, 1336 (2023).
34. Ni, R. et al. Controlled majority-inverter graph logic with highly nonlinear, self-rectifying memristor. *IEEE Trans. Electron Devices* **68**, 4897–4902 (2021).
35. Sung, S. H., Kim, T. J., Shin, H., Im, T. H. & Lee, K. J. Simultaneous emulation of synaptic and intrinsic plasticity using a memristive synapse. *Nat. Commun.* **13**, 2811 (2022).
36. Lee, Y. K. et al. Matrix mapping on crossbar memory arrays with resistive interconnects and its use in in-memory compression of biosignals. *Micromachines* **10**, 306 (2019).
37. Kim, H., Mahmoodi, M. R., Nili, H. & Strukov, D. B. 4K-memristor analog-grade passive crossbar circuit. *Nat. Commun.* **12**, 5198 (2021).
38. Ren, S. G. et al. Self-rectifying memristors for three-dimensional in-memory computing. *Adv. Mater.* **36**, 2307218 (2024).
39. Jeon, K. et al. Purely self-rectifying memristor-based passive crossbar array for artificial neural network accelerators. *Nat. Commun.* **15**, 129 (2024).
40. Joksas, D. et al. Nonideality-aware training for accurate and robust low-power memristive neural networks. *Adv. Sci.* **9**, 2105784 (2022).
41. Chiu, F. C. A review on conduction mechanisms in dielectric films. *Adv. Mater. Sci. Eng.* **2014**, 578168 (2014).
42. Acharya, U. R. et al. A deep convolutional neural network model to classify heartbeats. *Comput. Biol. Med.* **89**, 389–396 (2017).
43. Badertscher, P. et al. Clinical validation of a novel smartwatch for automated detection of atrial fibrillation. *Heart Rhythm O2* **3**, 208 (2022).
44. Chang, T. Y. et al. A high-precision deep learning algorithm to localize idiopathic ventricular arrhythmias. *J. Personal. Med.* **12**, 764 (2022).
45. Hajeb-M, S., Cascella, A., Valentine, M. & Chon, K. H. Deep neural network approach for continuous ECG-based automated external defibrillator shock advisory system during cardiopulmonary resuscitation. *J. Am. Heart Assoc.* **10**, e019065 (2021).
46. Puszkarski, B., Hryniów, K. & Sarwas, G. Comparison of neural basis expansion analysis for interpretable time series (N-BEATS) and recurrent neural networks for heart dysfunction classification. *Physiol. Meas.* **43**, 064006 (2022).
47. Singh, V., Tewary, S., Sardana, V., & Sardana, H. K. Arrhythmia detection—a machine learning based comparative analysis with MIT-BIH ECG data. In *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*. pp 1–5 (IEEE, 2019).
48. Ambrogio, S. et al. An analog-AI chip for energy-efficient speech recognition and transcription. *Nature* **620**, 768–775 (2023).
49. Yao, P. et al. Fully hardware-implemented memristor convolutional neural network. *Nature* **577**, 641–646 (2020).
50. Cheong, W. H. et al. Demonstration of neuromodulation-inspired stashing system for energy-efficient learning of spiking neural network using a self-rectifying memristor array. *Adv. Funct. Mater.* **32**, 2200337 (2022).
51. Moody, G. B. & Mark, R. G. The impact of the MIT-BIH arrhythmia database. *IEEE Eng. Med. Biol. Mag.* **20**, 45 (2001).
52. Ebrahimi, Z., Loni, M., Daneshmand, M. & Gharehbaghi, A. A review on deep learning methods for ECG arrhythmia classification. *Expert Syst. Appl.* **7**, 100033 (2020).
53. Jang, Y. H. et al. Spatiotemporal data processing with memristor crossbar-array-based graph reservoir. *Adv. Mater.* **36**, 2309314 (2024).
54. Li, D., Chen, X., Becchi, M., & Zong, Z. Evaluating the energy efficiency of deep convolutional neural networks on CPUs and GPUs. In *2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom) (BDCloud-SocialCom-SustainCom)*. pp 477–484. (IEEE, 2016).
55. Kim, M. G., Brown, D. K. & Brand, O. Nanofabrication for all-soft and high-density electronic devices based on liquid metal. *Nat. Commun.* **11**, 1002 (2020).
56. Kim, D. et al. Failure criterion of silver nanowire electrodes on a polymer substrate for highly flexible devices. *Sci. Rep.* **7**, 45903 (2017).
57. Kim, M. et al. Forming-less flexible memristor crossbar array for neuromorphic computing applications produced using low-temperature atomic layer deposition. *Appl. Mater. Today* **38**, 102204 (2024).
58. Zhu, S. et al. In-depth physical mechanism analysis and wearable applications of HfO_x-based flexible memristors. *ACS Appl. Mater. Interfaces* **15**, 5420–5431 (2023).
59. Wang, T. Y. et al. Flexible 3D memristor array for binary storage and multi-states neuromorphic computing applications. *InfoMat* **3**, 212–221 (2021).
60. Sun, L. et al. Self-selective van der Waals heterostructures for large scale memory array. *Nat. Commun.* **10**, 3161 (2019).
61. Wang, T. et al. Mechanically durable memristor arrays based on a discrete structure design. *Adv. Mater.* **34**, 2106212 (2022).
62. Sim, K. et al. Metal oxide semiconductor nanomembrane-based soft unnoticeable multifunctional electronics for wearable human-machine interfaces. *Sci. Adv.* **5**, eaav9653 (2019).
63. Nemati, E., Deen, M. J. & Mondal, T. A wireless wearable ECG sensor for long-term applications. *IEEE Commun. Mag.* **50**, 36 (2012).
64. Khayatizadeh, M. et al. A 0.7-V 17.4- μ S/W 3-lead wireless ECG SoC. *IEEE Trans. Biomed. Circuits Syst.* **7**, 583 (2013).
65. Altini, M. et al. An ECG patch combining a customized ultra-low-power ECG SoC with Bluetooth low energy for long term ambulatory monitoring. *Proc. 2nd Conf. Wirel. Health* **15**, 1 (2011).
66. Jain, A. et al. Real time system on chip based wearable cardiac activity monitoring sensor. *Measurement* **24**, 100523 (2022).
67. Huang, Hui et al. Energy-efficient ECG signal compression for user data input in cyber-physical systems by leveraging empirical mode decomposition. *ACM Trans. Cyber Phys. Syst.* **3**, 1 (2019).

Acknowledgements

This research was supported by NNFC (Grant No. 1711160154) and the National Research Foundation of Korea (NRF) (Grant Nos. RS-2023-00216619, RS-2023-00216992, 2022M3F3A2A01076569, 2022M3I7A4085484, and 2023R1A2C2005159). Ethical review by the Korea Advanced Institute of Science and Technology Institutional Review Board (KAIST IRB).

Author contributions

Y.L., H.R., and G.K. contributed equally to this work. Y.L., H.R., and G.K. conceived and designed the research. Y.L., H.R., G.K., D.H.K., W.H.C., and H.S. developed the methodology. Y.L. and G.K. curated the data. H. R. wrote the software. Y.L., H.R., G.K., W. H. C., and S.N.K. visualized the results. Y.L., H.R., and G.K. wrote the original draft. J.L. and K.M.K. supervised the work. All the authors reviewed and approved the final paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-59589-2>.

Correspondence and requests for materials should be addressed to Jongwon Lee or Kyung Min Kim.

Peer review information *Nature Communications* thanks Cheol Seong Hwang, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025