Check for updates

Advanced Electronic Materials

ADVANCED ELECTRONIC MATERIALS Open Access

www.advelectronicmat.de

RESEARCH ARTICLE OPEN ACCESS

Ion-Gating Reservoir Computing for Preprocessing-Free Speech Recognition from Throat Vibrations

Daiki Nishioka[1] | Akito Tateyama[2,3] | Hina Kitano[2,4] | Takashi Nakanishi[2,3] | Kazuya Terabe[2] | Takashi Tsuchiya[2,4]

[1]International Center for Young Scientists (ICYS), National Institute for Materials Science (NIMS), Tsukuba, Ibaraki, Japan | [2]Research Center for Materials Nanoarchitectonics (MANA), National Institute for Materials Science (NIMS), Tsukuba, Ibaraki, Japan | [3]Division of Soft Matter, Graduate School of Life Science, Hokkaido University, Sapporo, Japan | [4]Department of Applied Physics, Tokyo University of Science, Tokyo, Japan

Correspondence: Daiki Nishioka (NISHIOKA.Daiki@nims.go.jp) | Takashi Nakanishi (NAKANISHI.Takashi@nims.go.jp) | Takashi Tsuchiya (TSUCHIYA.Takashi@nims.go.jp)

Received: 12 January 2026 | **Revised:** 24 February 2026 | **Accepted:** 21 April 2026

Keywords: edge artificial intelligence | gel-electret mechanoelectric generator | graphene electric-double-layer transistor | ion-gating reservoir | physical reservoir computing | spoken-digit recognition | throat-vibration sensing

ABSTRACT

Physical reservoir computing offers a promising route to efficient edge-AI, yet most reservoir-based speech-recognition demonstrations still rely on microphones and digital preprocessing. Here, we report a real-time spoken-digit recognition system that directly processes throat vibrations by integrating a π–gel-electret mechano-electric generator (MEG) sensor with a multi-channel ion-gel/graphene ion-gating reservoir (IGR). The MEG, a free-deformable electret-based vibration sensor, converts articulation-induced throat-surface vibrations into voltage signals, which are then transformed into pulse sequences and applied directly to the IGR gate without any spectral or handcrafted feature extraction. Channel-dependent Dirac-point shifts, nonlinear electric-double-layer dynamics, and distinct relaxation behaviors in the IGR expand the one-dimensional MEG output into high-dimensional reservoir states, which are decoded using a simple linear classifier. This hybrid MEG–IGR system achieves an impressive real-time spoken-digit recognition accuracy of 96.8%, demonstrating that ion-gating reservoirs can efficiently extract discriminative spatiotemporal features from raw biomechanical signals. By eliminating frequency-domain preprocessing and enabling compact, low-power front-end computation, the approach provides a viable pathway for on-device speech interfaces. Beyond speech recognition, direct throat-vibration sensing opens opportunities for robust human–machine interfaces, assistive communication technologies, and secure authentication systems based on spoof-resistant biomechanical signatures.

1 | Introduction

Speech recognition has become a fundamental component of modern human–machine interfaces, enabling voice-controlled operation in mobile devices, wearables, smart appliances, and assistive technologies [1]. Conventional speech-recognition systems rely on microphones to capture airborne acoustic signals, followed by computationally intensive digital preprocessing—such as Fourier-based spectral decomposition, mel-frequency cepstral coefficient (MFCC) extraction, and deep neural network (DNN) inference [2–4]. While these methods have led to remarkable advances in accuracy, they remain vulnerable to environmental noise [5] and require considerable power and computational resources [6], making them unsuitable for low-power edge devices. Moreover, recent advances in generative AI have made it increasingly easy to synthesize natural-sounding speech, raising concerns about the security and authenticity of microphone-based voice interfaces [7, 8]. An alternative

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Advanced Electronic Materials* published by Wiley-VCH GmbH

approach to speech recognition is to measure the mechanical vibrations of the throat resulting from the motion of the vocal folds and surrounding tissues during articulation [9–14]. Recent wearable silent-speech interfaces and "artificial throat" devices have expanded this direction by employing diverse transduction mechanisms—including strain-/iontronic sensors and photonic implementations (e.g., optical-fiber–based wearable photonic artificial throats)—to decode speech content and enable silent communication from neck-surface motions [15]. More broadly, rapid progress in photonic biointerfaces—spanning flexible optical-fiber platforms, wearable photonic wristbands for pulse-wave–based cardiorespiratory monitoring and biometric identification, and chip-scale photonic sensors for cuffless blood-pressure measurements—highlights the expanding photonic toolbox for robust, miniaturizable physiological sensing at the edge [16–18]. Such biomechanical vibration signals are inherently robust against environmental acoustic noise because they do not rely on airborne sound propagation [9, 12, 14]. Furthermore, they reflect physiological characteristics that are extremely difficult to spoof, offering unique opportunities for secure and noise-resilient speech interfaces. However, even when vibration or contact sensors are employed, existing systems still depend on digital signal processing or deep-learning pipelines [9–14], which limit their suitability for compact, low-power, real-time applications.

Physical reservoir computing (PRC) has recently emerged as a promising paradigm for efficient edge-AI systems [19, 20]. By harnessing the intrinsic dynamics and nonlinear responses of physical devices, PRC enables high-dimensional feature expansion and temporal processing with minimal computational overhead, requiring only a linear readout for training. Among various physical reservoirs [21–40], ion-gating reservoirs (IGRs)—which exploit ionic gating–induced modulation of electronic material properties—are particularly attractive for edge-AI implementations due to their flexible material design, compact device structures, and strong computational capability [41–54]. For integrating PRC directly with sensors to construct fully physical edge-AI systems, however, the reservoir must additionally accommodate wide sensor bandwidths and maintain high computational efficiency to achieve strong performance with relatively small node counts. In this context, the ion-gel/monolayer-graphene electric-double-layer transistor (EDLT)–based IGR that we recently developed achieves exceptionally high performance, broad-band operation, and excellent computational efficiency [53]. These capabilities originate from the ambipolar transport of graphene, which provides rich nonlinear responses, combined with multiple relaxation processes arising from ionic motion in the gel electrolyte and the charging/discharging dynamics of the electric double layer (EDL). This combination yields a highly capable physical reservoir suitable for real-time processing of time-varying sensor signals. Despite these advances, demonstrations of reservoir-computing–based speech recognition have largely focused on microphone-acquired audio and typically rely on spectral or handcrafted feature extraction [21–28]. In contrast, the direct processing of biomechanical throat-vibration signals—without frequency-domain transformations or feature-engineering steps—has received relatively little attention. Furthermore, the behavior of ion-gating reservoirs when driven by such vibration signals and their suitability for real-time edge-AI implementations have yet to be systematically explored.

Here, we address these challenges by integrating a π–gel-electret mechano-electric generator (MEG) sensor [55] with a multichannel ion-gel/graphene IGR [53] to realize preprocessing-free, real-time spoken-digit recognition. The π–gel-electret is a viscoelastic gel material made by converting π-conjugated molecules (e.g., porphyrin [56], fullerene [57], pyrene [55], conjugated polymer [58], etc.) substituted with branched alkyl chains into a solvent-free liquid at room temperature, gelating them with the addition of a low-molecular-weight gelator, and then electretizing them through corona charging. The MEG is a free-deformable electret-based vibration sensor in which a corona-poled π–gel retains quasi-permanent charges, enabling sensitive detection of throat-surface vibrations through deformation-induced potential changes while remaining intrinsically robust to ambient acoustic noise. The resulting one-dimensional voltage stream is converted into pulse-voltage sequences and applied directly to the IGR gate. Through channel-dependent Dirac-point shifts, nonlinear EDL dynamics, and distinct relaxation behaviors, the IGR transforms this input into eight diverse current responses, which are further expanded into a high-dimensional reservoir state and decoded by a linear readout. Despite the MEG's standalone recognition accuracy being only 18.2%, using this hybrid MEG–IGR reservoir, we achieve real-time spoken-digit recognition accuracy of 96.8%, demonstrating that the IGR effectively extracts discriminative spatiotemporal features embedded in raw biomechanical vibration signals. By eliminating the need for spectral analysis, hand-crafted features, or deep-learning-based preprocessing, our approach enables a compact and low-power physical-computing architecture suitable for on-device speech recognition. Taken together, this study establishes an entirely physical, preprocessing-free reservoir-computing framework for real-time speech recognition. Beyond speech interfaces, the demonstrated framework highlights the potential of throat-vibration sensing and ion-gating reservoirs for wearable communication devices, assistive technologies for individuals with speech impairments, and secure authentication systems based on biomechanical signatures.

2 | Results and Discussion

2.1 | Concept and Task Design for MEG–IGR–Based Speech Recognition

Figure 1a illustrates the conceptual framework of the speech-recognition system developed in this study, which integrates an MEG–based vibration sensor with an IGR. Instead of a conventional microphone, a π–gel-electret MEG device [55] was used to detect vibrations around the throat (the Adam's apple) during articulation. The MEG converts these mechanical vibrations into voltage signals. These raw voltage traces were transformed into pulse-voltage waveforms at the sensor's original sampling rate, and the resulting pulses were applied directly to the gate terminal of the IGR [53]. The corresponding gate, source, and drain current responses were regarded as reservoir states and fed to a linear readout layer to classify spoken digits. Because the IGR nonlinearly expands the MEG signals into a high-dimensional spatiotemporal representation, it enables accurate speech classification using only a lightweight linear model. Importantly, this framework eliminates the need for computationally intensive preprocessing such as Fourier-based spectral analysis or

2199160x, 2026, 19, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/aelm.202600006 by Takashi Tsuchiya - National Institute For , Wiley Online Library on [14/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

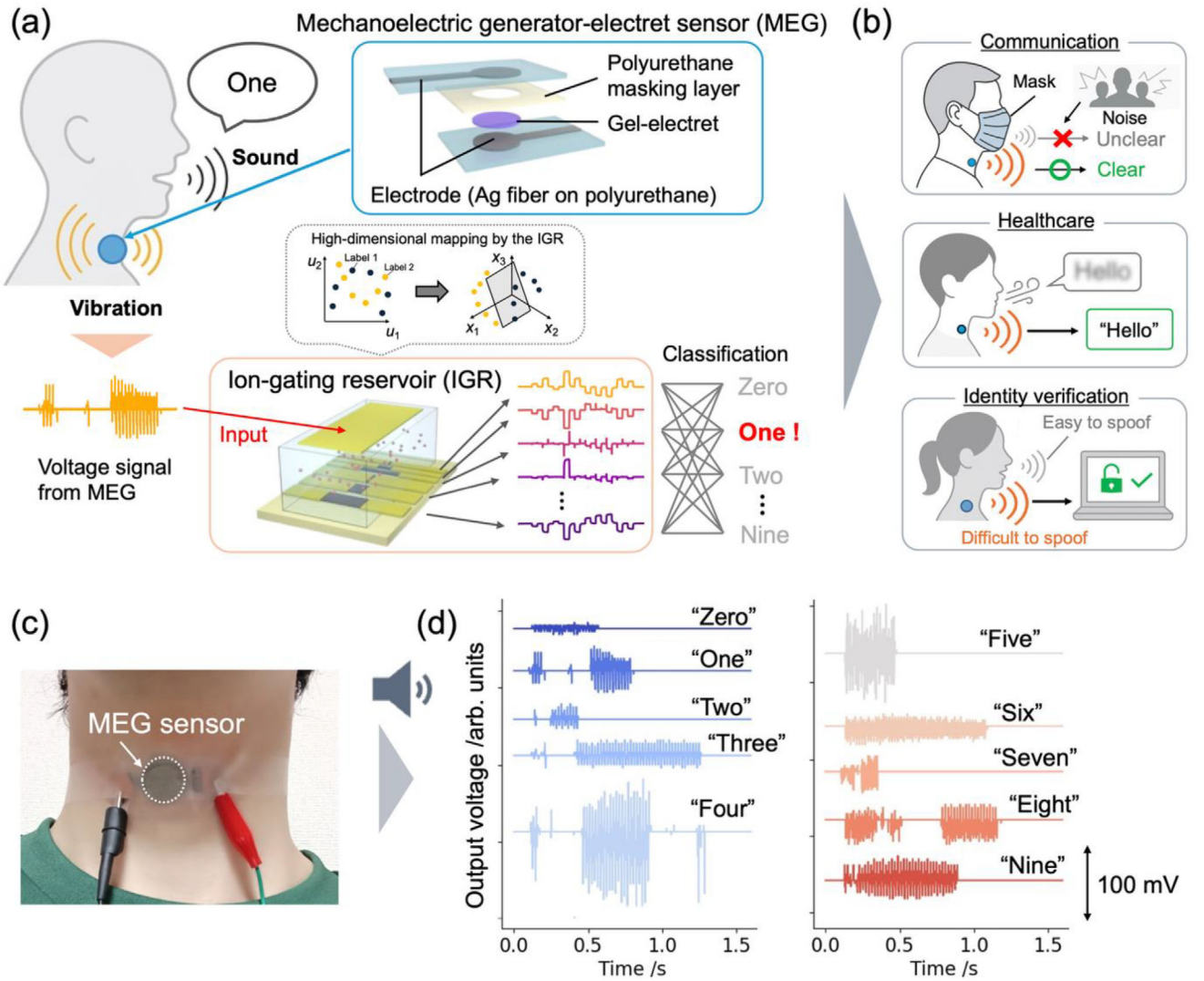


FIGURE 1 | Concept and data acquisition for MEG–IGR–based speech recognition. (a) Schematic illustration of the proposed speech-recognition framework integrating a π–gel-electret MEG vibration sensor with an IGR. MEG output voltages are converted into pulse signals and applied to the IGR, whose current responses form reservoir states for linear readout. (b) Potential applications of the MEG–IGR approach, including noise-resilient interfaces, assistive systems for individuals with speech impairments, and secure authentication based on throat-vibration patterns. (c) Placement of the MEG sensor on the throat, slightly below the laryngeal prominence, for acquiring vibration signals during digit articulation. (d) Representative MEG waveforms corresponding to spoken digits "zero" through "nine," demonstrating distinct temporal vibration patterns associated with each utterance.

handcrafted feature extraction, highlighting its potential for low-power, edge-AI applications. The anticipated application space is outlined in Figure 1b. Since the method relies on throat-vibration sensing rather than airborne sound, it naturally avoids acoustic coupling and is less susceptible to environmental noise [9, 12, 14]. Moreover, the ability to track biomechanical throat motion offers potential use for assistive interfaces supporting individuals with speech impairments [14]. As generative AI technologies increasingly enable the spoofing of audible speech [7, 8], the biomechanical nature of throat-vibration patterns—rooted in vocal-fold dynamics and tissue motion—also suggests applicability for secure personal identification systems.

To establish a dataset for evaluating the proposed framework, the MEG sensor was attached to the speaker's throat, slightly below the laryngeal prominence, as shown in Figure 1c. A total of 500 samples were collected by recording 50 repetitions each of the spoken digits "zero" through "nine." Representative MEG outputs for the ten digits are presented in Figure 1d. Although these waveforms differ fundamentally from acoustic speech signals—because they arise from mechanical vibrations of the throat surface—they nevertheless exhibit clear digit-dependent temporal patterns. The MEG used in this work is a deformable π–gel-electret–based mechanoelectric generator, constructed by infiltrating an alkyl-pyrene liquid gelated with a low–molecular-weight gelator into a polyurethane nonwoven matrix and subsequently corona-poling the structure to retain quasi-permanent charges [55]. Mechanical deformation of the electret layer modulates its internal and surface charge distribution, thereby generating voltage fluctuations proportional to the applied vibration. Owing to its electret-based transduction mechanism, the MEG sensor can detect subtle biomechanical signatures associated with articulation, providing rich temporal information that is highly discriminative across spoken-digit

2199160x, 2026, 19, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/aelm.202600006 by Takashi Tsuchiya - National Institute For , Wiley Online Library on [14/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

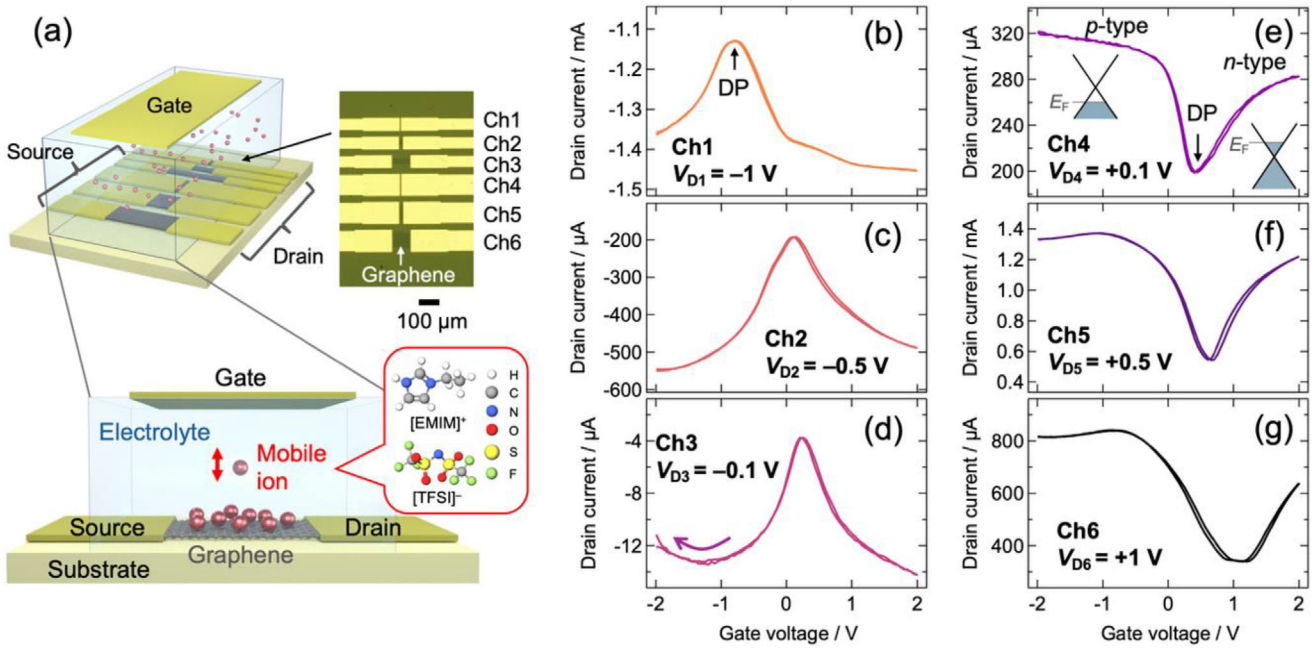


FIGURE 2 | Electrical characteristics of the ion-gel/graphene EDLT-based IGR. (a) Schematic illustration of the ion-gel/graphene IGR and optical micrograph of the multi-channel device comprising mono-layer graphene channels with different lengths (5 mm, 20 μm, 100 μm) and widths (30 and 80 μm), all sharing a common gate electrode. (b–g) Pulsed transfer characteristics of each graphene channel measured under distinct drain biases (V_{D1} = –1 V, V_{D2} = –0.5 V, V_{D3} = –0.1 V, V_{D4} = +0.1 V, V_{D5} = +0.5 V, V_{D6} = +1 V). Gate and drain voltages were applied as 10 ms pulses with a 1 s interval. The inset of (e) shows the Dirac cone of graphene, where modulation of the Fermi level E_F relative to the Dirac point by the applied V_G results in *p*-type or *n*-type behavior.

classes. Details of the fabrication procedure and material composition can be found in the Methods section and the prior report [55]. Despite the presence of such discriminative temporal features, conventional speech-recognition pipelines typically require substantial digital processing—including spectral decomposition, feature extraction, and deep neural inference—posing challenges for low-power edge implementation. In this study, we instead exploit the intrinsic nonlinear dynamics and computational richness of the IGR to process raw MEG outputs directly, without any explicit preprocessing. This approach allows us to evaluate the feasibility of an edge-compatible speech-recognition device in which a physical reservoir extracts and transforms the temporal structure embedded in throat-vibration signals.

2.2 | Electrical Characteristics of the Graphene-Based IGR and Its Nonlinear Mapping of MEG Outputs

Figure 2a illustrates the ion-gel/graphene IGR used as the physical reservoir layer in this study [53]. As shown in the optical micrograph inset, the device comprises multiple independent single-layer graphene channels with different geometries—lengths of 5 mm, 20 μm, and 100 μm, and widths of 30 and 80 μm—sharing a common gate electrode. Six drain terminals provide six independent drain-current (I_D) responses to a single gate input. When a gate voltage (V_G) is applied, mobile ions within the electrolyte migrate toward the graphene interface and form an EDL [49, 59–61], leading to highly efficient electrostatic carrier injection into graphene. The resulting modulation of channel conductance is detected as the drain current. Further details of the device physics underlying this ion-gel/graphene IGR can be found in our previous studies [53]. Figure 2b–g shows the transport characteristics of each channel. To isolate the EDL-driven carrier injection response from slow dynamical processes often observed in graphene transistors [62, 63], both the gate and drain voltages were applied as pulses, and the drain-current values at each pulse were plotted. With a pulse width of 10 ms and an interval of 1 s, all channels exhibit characteristic V-shaped transfer curves associated with the linear dispersion of monolayer graphene. Such strong nonlinear transfer characteristics directly contribute to the nonlinear mapping capability essential for reservoir computing. To further diversify the reservoir states, each drain terminal was biased at a different voltage (V_{D1} = –1 V, V_{D2} = –0.5 V, V_{D3} = –0.1 V, V_{D4} = +0.1 V, V_{D5} = +0.5 V, V_{D6} = +1 V). As evident in Figure 2b–g, the Dirac point (DP) systematically shifts with V_D, originating from the decrease in the effective gate voltage, $V_{G,eff} = V_G - V_D/2$. This channel-dependent DP shift provides multiple distinct nonlinear transfer functions that enrich the high-dimensional reservoir state space. In addition to DP shifts, the magnitude of V_D also affects the high-V_G plateau region. In graphene FETs, this plateau commonly arises from mobility degradation under high fields due to carrier scattering [64]. In ion-gated transistors, an additional effect emerges: pulsed gate voltages induce EDL charging/discharging currents that contribute to the measured gate current, and when the drain current is comparable in magnitude, a portion of this ionic current is superimposed on the drain current [41, 53]. This effect is particularly visible in high-resistance channels under small drain biases—for example, ch3 at V_D = –0.1 V. Under these conditions, relatively large negative gate pulses (approximately –1 to –2 V) generate gate currents that are no longer negligible

2199160x, 2026, 19, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/aelm.202600006 by Takashi Tsuchiya - National Institute For , Wiley Online Library on [14/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

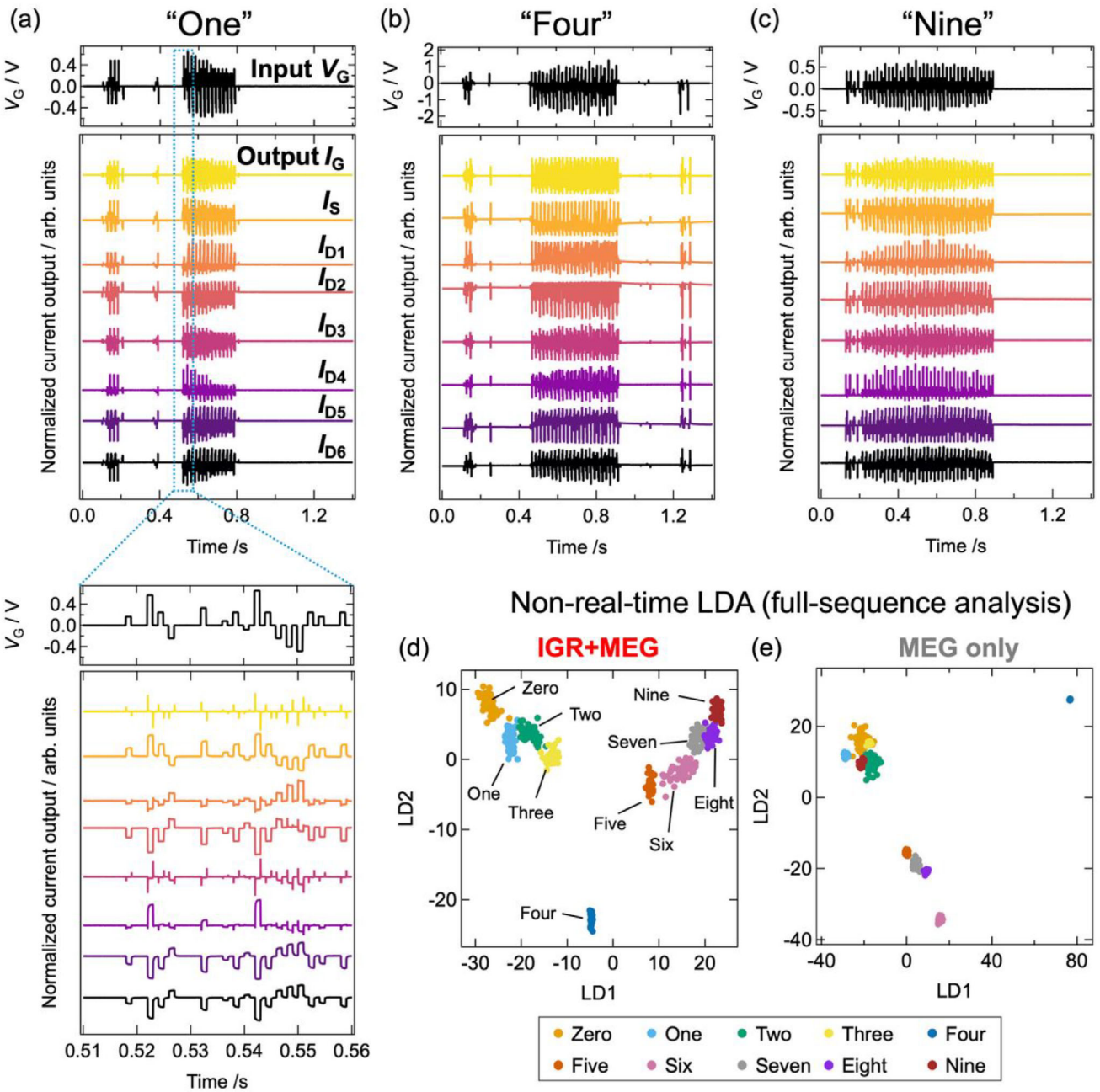


FIGURE 3 | IGR processing of MEG vibration signals and class separability. (a–c) Pulse-shaped V_G sequences converted from MEG outputs corresponding to the spoken digits one, four, and nine (upper panels), and the resulting IGR current responses (lower panels). The inset in (a) magnifies the gate-current spikes induced by EDL charging/discharging and the relaxation-type drain-current dynamics. (d) Conventional non-real-time LDA of the full time-evolution of the IGR current responses for 500 spoken-digit samples. The first two discriminant axes (LD1, LD2) yield apparent clustering of the ten digit classes. (e) LDA applied directly to the raw MEG sensor outputs, showing insufficient class separation. Comparison with (d) demonstrates that the IGR nonlinearly expands the MEG signal into a high-dimensional feature space, enabling linear discrimination.

compared with the drain current. As a result, as indicated by the purple arrow in Figure 2d, the drain current exhibits a strong nonlinear response that resembles an apparent increase in channel resistance as V_G becomes more negative. The coexistence of these device-state-dependent nonlinear dynamics indicates that the ion-gel/graphene IGR provides a rich nonlinear basis for reservoir computing [53].

Figure 3a–c presents the pulse-shaped gate-voltage sequences obtained from MEG outputs for the spoken digits one, four, and nine (upper panels), along with the corresponding IGR current responses (lower panels). The MEG sensor operated at 500 Hz, and the IGR was driven synchronously with 500 Hz V_G pulses (50% duty cycle). The pulse amplitude during the on-state was set to 20 times the MEG voltage (scaled to a maximum of 4 V), while the off-state was fixed at 0 V. During these measurements, six drain currents (I_{D1}–I_{D6}) and the gate and source currents (I_G and I_S) were recorded. Thus, the IGR expands a one-dimensional MEG input into eight-dimensional current responses that serve as the reservoir states. As seen in Figure 3a–c, the eight current

2199160x, 2026, 19, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/aelm.202600006 by Takashi Tsuchiya - National Institute For , Wiley Online Library on [14/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

responses exhibit distinct temporal patterns depending on the spoken digit. The inset of Figure 3a shows a magnified region of the input–output relationship: the gate current displays spike-like features associated with EDL charging and discharging [41], whereas the drain currents follow relaxation-type dynamics. Because each channel experiences a distinct Dirac point and nonlinear transfer characteristic due to its drain bias, the same gate input produces diverse current responses across channels. Under small drain biases (e.g., I_{D3} and I_{D4}), gate-current-induced spikes appear in the drain current due to the relatively large ionic current contribution. Whether such spikes emerge depends on device state and bias conditions, introducing additional state-dependent nonlinearities that further enrich the reservoir's computational diversity.

To use these current responses for real-time speech recognition, it is necessary to sequentially extract temporal features (reservoir states) from them and feed these features into a linear classifier. In this way, the linear classifier can output a temporal sequence of decisions. Although the details of this real-time speech-recognition scheme are discussed in the next section, we first examine the characteristics of the IGR current responses shown in Figure 3a–c—particularly their waveform shapes—and evaluate the latent separability among classes. For this purpose, we performed a non–real-time linear discriminant analysis (LDA) [65] using the IGR current responses corresponding to all 500 speech samples. For each speech sample, all temporal points across the eight current responses were treated as features. As shown in Figure 3d,e, the dimensionality was reduced to two discriminant axes (LD1 and LD2) that maximize the separation among the ten digit classes. In Figure 3d, the full temporal current-response features generated by the IGR form clearly separated clusters corresponding to each digit. In contrast, when LDA is applied directly to the raw MEG signals (Figure 3e), the separability is significantly degraded. These results indicate that the IGR current responses shown in Figure 3a–c nonlinearly transform the raw MEG outputs into a representation that is meaningful for class discrimination. Furthermore, classes that lie relatively close to one another in this discriminant space tend to be more difficult to classify, consistent with their waveform similarity. However, this LDA analysis utilizes the full time sequence of each response, making it unsuitable for real-time operation because classification can only occur after the entire utterance is complete. Conventional speech-recognition systems rely on sequential preprocessing (e.g., MFCCs or cochleagrams) [21–28], but these impose a substantial computational burden that limits deployment on edge devices. To address this, we feed the raw MEG signals directly into the IGR without any preprocessing and sequentially sample the resulting current responses, enabling real-time reservoir-state generation. This real-time spoken-digit recognition strategy is described in detail in the following section.

2.3 | Spoken-Digit Recognition Using MEG–IGR Physical Reservoir Computing

Figure 4a illustrates the framework of the spoken-digit recognition task implemented using the MEG sensor and the IGR. The voltage output from the MEG was converted into a 500 Hz pulse-voltage sequence—matching the MEG sampling rate—and applied directly to the gate terminal of the IGR. From the resulting eight-dimensional current responses (examples in Figure 3a–c), a set of N_V virtual nodes was sequentially sampled within a time window t_W. The resulting reservoir size N was therefore ($N = 8N_V$), and the interval between successive nodes was θ. For each time window, the reservoir-state vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_N, 1)$ was linearly transformed using a readout weight matrix W to generate a ten-dimensional temporal output $\boldsymbol{y} = (y_0, y_1, \ldots, y_9)$ as follows:

$$\boldsymbol{y} = W\boldsymbol{x} \tag{1}$$

The readout weights W were trained using ridge regression with one-hot encoded target vectors for digits 0–9, as described in the Methods section. For classification, the temporal outputs obtained across all time windows within a single sample were summed to form an integrated output vector $\tilde{\boldsymbol{y}}$. The predicted digit was then determined using a winner-take-all (WTA) rule [21–28], in which the index corresponding to the maximum component of the integrated vector was selected: Answer = $\mathrm{argmax}_i(\tilde{y}_i)$.

Figure 4b,c summarizes the dependence of the test accuracy on the t_W and node interval θ, obtained from five-fold cross-validation (400 training and 100 test samples). Under the condition $\theta = 500$ μs (Figure 4b), the MEG–IGR system achieved 84.2% accuracy at $t_W = 2$ ms, which increased with t_W and saturated at approximately 95% around $t_W = 16$ ms. This improvement stems from the fact that a larger t_W enables the reservoir to capture temporal correlations over longer time scales, effectively enhancing its memory. However, increasing t_W reduces real-time responsiveness and—under fixed θ—increases the number of virtual nodes, leading to higher computational cost. The condition indicated by the black arrow in Figure 4b corresponds to $t_W = 2$ ms, where four virtual nodes were extracted per pulse (i.e., $N = 32$). The corresponding confusion matrix and test accuracy are shown in Figure 4d. Remarkably, despite the extremely small reservoir size, the MEG–IGR system achieved 84.2% accuracy, outperforming the MEG-only baseline shown in upper panel of Figure 4e. The MEG-only result was obtained by directly feeding an equivalent number of points (32 points corresponding to $t_W = 64$ ms) into a linear classifier, effectively treating the MEG sensor as a physical reservoir. The low accuracy of the MEG-only baseline (≈18%; Figure 4e, upper panel) can be rationalized by the LDA result in Figure 3e, which indicates that the raw throat-vibration dynamics measured by the MEG do not provide a representation that is readily linearly separable across spoken-digit classes. Consequently, approaches that directly feed the sensor output into a linear model have limited discriminative power for this task. Importantly, however, the MEG waveforms still embed digit-dependent information in a form that can be exploited by nonlinear processing or suitable feature extraction. Consistent with this, a conventional digital baseline that extracts mel-frequency cepstral coefficients (MFCCs) from the MEG signals and applies a linear classifier achieves a test accuracy of 91.6% (Figure 4e, lower panel). Details of the MFCC-based pipeline are provided in Note S1 and Figure S1. These comparisons highlight that the IGR functions as a compact physical feature-expansion front end that renders the MEG signal more linearly decodable, while avoiding conventional digital preprocessing. The substantial improvement achieved with the IGR confirms its ability to extract and amplify discriminative features embedded in the MEG signal. This arises from the high information-processing

2199160x, 2026, 19, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/aelm.202600006 by Takashi Tsuchiya - National Institute For , Wiley Online Library on [14/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

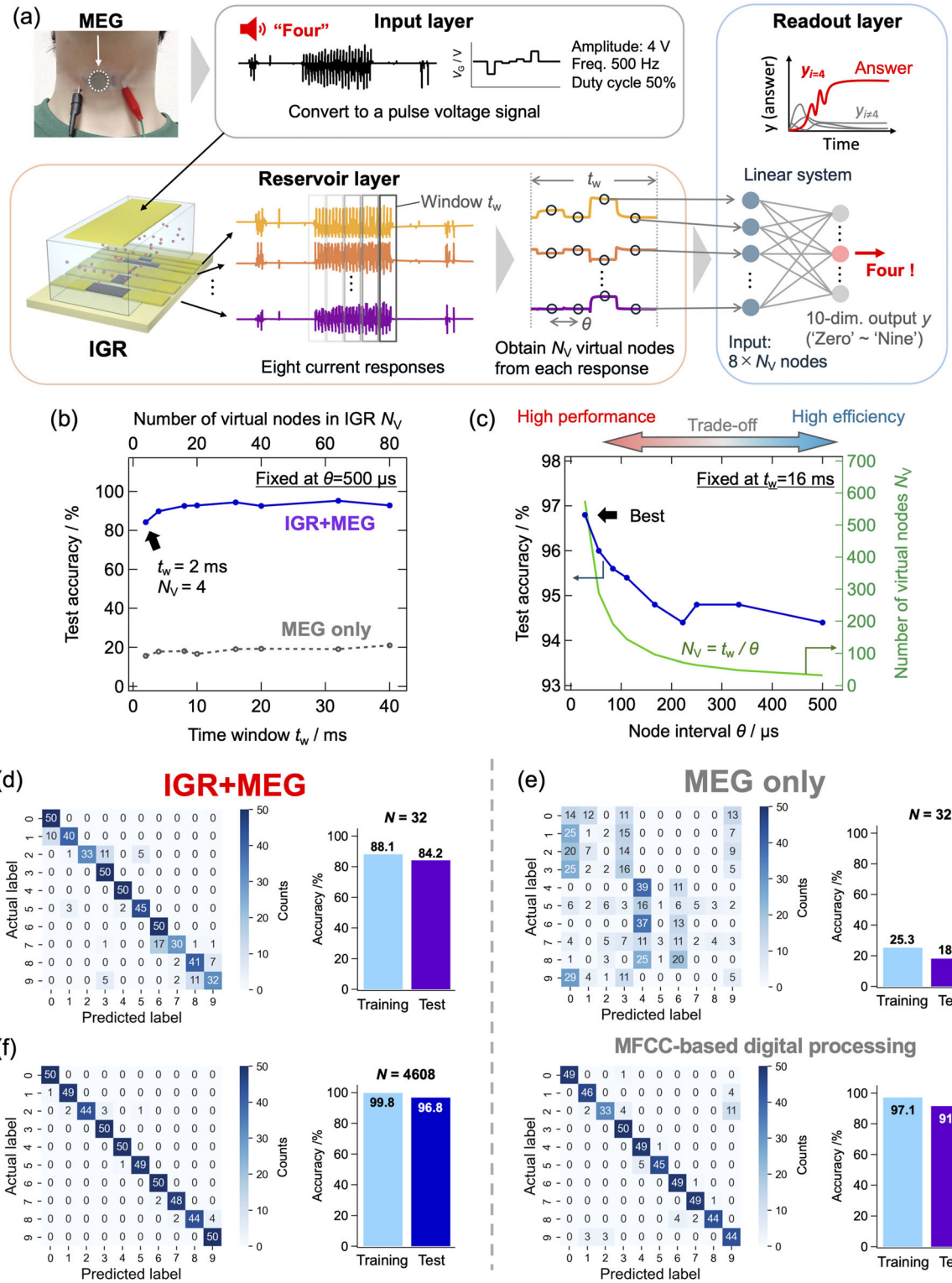


FIGURE 4 | Real-time spoken-digit recognition using the MEG–IGR reservoir. (a) Schematic of the real-time spoken-digit recognition framework. MEG outputs are converted into 500 Hz pulse-voltage sequences and applied to the IGR gate. From the resulting eight-dimensional current responses, N_V virtual nodes are sampled within a time window t_W with interval θ, forming a reservoir state vector used for linear readout. (b) Dependence of test accuracy on t_W at a fixed node interval $\theta = 500$ μs. The black arrow marks the minimal-node configuration ($t_W = 2$ ms, $N = 32$). (c) Dependence of test accuracy (left axis) and virtual node count N_V (right axis) on θ at a fixed $t_W = 16$ ms. Reducing θ increases temporal resolution and boosts performance, with the smallest $\theta = 27.8$ μs yielding the highest accuracy. (d) Confusion matrix for the minimal-node condition ($t_W = 2$ ms, $N = 32$), showing 84.2% test accuracy. (e) Baseline results using the MEG signal without the IGR. Upper: MEG-only PRC baseline obtained by feeding 32 MEG samples directly into a linear classifier. exhibiting significantly lower accuracy than the MEG–IGR system. Lower: Conventional digital baseline in which mel-frequency cepstral coefficients (MFCCs) are extracted from the MEG signals and classified using a linear model. (f) Confusion matrix for the best-performing condition ($t_W = 16$ ms, $\theta = 27.8$ μs, $N_V = 576$, $N = 4608$), achieving 99.8% training accuracy and 96.8% test accuracy.

2199160x, 2026, 19, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/aelm.202600006 by Takashi Tsuchiya - National Institute For , Wiley Online Library on [14/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

capacity (IPC) [66] —including nonlinear transformation and memory—and the intrinsic high dimensionality of the IGR [53]. By projecting the input into a space spanned by past states and their higher-order interactions, the IGR enables linear discrimination of spoken-digit patterns. Further details regarding the IPC characteristics of this IGR platform are provided in our previous studies. The confusion matrix in Figure 4d reveals that misclassifications occur predominantly among neighboring digit groups (e.g., 0–3 and 6–9). This is consistent with Figure 3d, where these digit classes occupy nearby positions in the discriminant subspace. Improved separability among such closely spaced classes can be achieved by increasing the temporal resolution of the reservoir states—i.e., decreasing θ and increasing N_V—which enables finer sampling of the temporal dynamics of the IGR currents.

Figure 4c plots the dependence of classification accuracy (left axis) and the number of virtual nodes N_V (right axis) on θ at a fixed time window of $t_W = 16$ ms (corresponding to eight MEG pulses). Across all θ, accuracies above 94% were consistently achieved. As expected from $N_V = t_W/\theta$, reducing θ increases the number of nodes and improves classification performance by allowing the reservoir to capture the spike-like and relaxation dynamics of the IGR output with higher temporal precision. The highest performance was obtained at the smallest $\theta = 27.8$ μs, indicated by the black arrow in Figure 4c, yielding 99.8% training accuracy and 96.8% test accuracy (Figure 4f). Notably, this test accuracy slightly exceeds the 96.2% accuracy recently reported for analog in-memory computing using a memristor-implemented deep neural network that extracts features from the TI-46 spoken-digit dataset without digital preprocessing [67]. In contrast to such multilayer nonlinear architectures, the MEG–IGR system achieves comparable or even superior performance using only a single physical reservoir layer followed by a linear readout. This highlights the strong nonlinear transformation capabilities of the ion-gating reservoir and its efficiency in extracting discriminative features directly from raw sensor signals. Although this condition yields a relatively large reservoir size ($N_V = 576$, $N = 4608$), Figure 4c demonstrates a clear trade-off between reservoir size and computational performance: by adjusting θ, the accuracy can be tuned between 94.4% and 96.8% while varying N_V from 32 to 576. Here, the number of physical nodes N_P was fixed at eight, and the overall reservoir size was tuned solely by modulating N_V through θ. This adjustment corresponds to controlling how extensively the intrinsic high dimensionality arising from the time evolution of the IGR current responses is exploited for information processing. Figure S2a reproduces the N_V dependence of the classification accuracy under the fixed physical node condition ($N_P = 8$; equivalent to Figure 4c), whereas Figure S2b shows the dependence on the number of physical nodes when N_V is fixed at 576. Comparison of these trends indicates that the number of physical nodes exerts a larger influence on the overall performance than N_V within the investigated range. This suggests that the dominant contribution to the reservoir's effective dimensionality originates from the diversity of the physical channels, arising not only from geometry induced variations in channel resistance (and thus relaxation times), but also from the diversity in transport characteristics (i.e., nonlinearity) introduced by independently applying distinct V_D biases. Consequently, modulating N_V (i.e., tuning θ) provides a practical knob to flexibly balance computational cost and performance while maintaining a high accuracy above 94%.

This adaptive tunability between reservoir size and classification accuracy is highly advantageous for edge-AI deployment. Depending on system requirements—desired accuracy, permissible power consumption, available computational resources, and whether training is performed on-device or in the cloud—the MEG–IGR reservoir can be flexibly configured to operate in either a low-resource or high-performance mode. Such configurability is essential for practical implementations of edge-intelligent sensing platforms. In addition, preliminary power characterization (Note S2 and Table S1) indicates that the IGR consumes an average electrical power of 5.09 mW under the present pulsed-bias conditions, which corresponds to an energy cost of 3.72 mJ per word inference (i.e., the time-integrated power over one utterance). Because the overall energy budget scales with both the applied drain bias and the number of sampled reservoir states, a low-power operating mode can be implemented by simultaneously reducing the drain voltages to $|V_D| \leq 0.1$ V and decreasing N_V via coarser time multiplexing (larger θ, e.g., $\theta = 500$ μs). Under this combined setting, the average power is expected to decrease to approximately 150 μW, corresponding to ~109 μJ per word.

Although the reservoir states and the readout outputs are generated sequentially in real time, the final classification in the present implementation is obtained by applying a WTA decision to the time-integrated output over the entire utterance. Consequently, an end-to-end decision latency arises until a word is fully articulated. In our dataset, the average waveform duration is 0.73 s; therefore, if the decision is made after integrating the readout over the full waveform, the latency is on the order of this duration (plus negligible computation time for the linear readout). To explore latency reduction, we evaluated an early-decision strategy in which the WTA decision is computed using only the initial portion of each utterance. Figure S3 plots the classification accuracy as a function of the decision time (i.e., the length of the initial segment used for WTA). While the accuracy is low immediately after speech onset, it increases rapidly and reaches approximately 83% at around 200 ms. This result indicates that the MEG–IGR system can perform fast classification before the utterance ends by leveraging early spatiotemporal features, providing a viable route to low-latency operation. To put these results in a broader context, Table S2 provides a side-by-side comparison with representative PRC approaches, summarizing key practical metrics including classification accuracy, preprocessing requirements, power/energy consumption, and decision latency. Overall, the MEG–IGR platform offers a competitive and highly configurable operating space—ranging from low-resource, low-power settings to higher-performance modes—while retaining the ability to extract discriminative features directly from raw sensor signals.

2.4 | Outlook

In practical operating environments, however, the classification task is expected to become more challenging due to noise and motion-induced artifacts (e.g., sensor displacement, skin motion, and mechanically induced baseline fluctuations). In such

2199160x, 2026, 19, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/aelm.202600006 by Takashi Tsuchiya - National Institute For , Wiley Online Library on [14/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

situations, one of the most effective approaches is to incorporate, as training data, as many samples as possible acquired under diverse noise levels and motion conditions. By exposing the readout to such variability during training, the system can learn more invariant decision boundaries and become more robust to real-world perturbations. Systematic validation of robustness under controlled noise and motion scenarios will be pursued as future work.

As further future directions, it is important to assess scalability beyond isolated spoken digits, including extension to different languages, larger vocabularies, and continuous-speech recognition. Notably, in our framework the MEG and IGR operate synchronously, and the readout produces outputs sequentially for each time window, enabling inherently streaming (real-time) inference. This property is expected to be advantageous for continuous recognition, where predictions can be updated window by window and combined with temporal segmentation/decoding strategies. At the same time, large-vocabulary natural-language processing (e.g., language modeling) typically exceeds the scope of conventional time-series tasks addressed by reservoir computing alone. In such cases, the MEG–IGR device is anticipated to function as a low-power front-end that extracts informative spatiotemporal features from raw biomechanical signals and interfaces with higher-level digital processors dedicated to more complex linguistic inference.

Another promising direction is on-device learning. By integrating the IGR with a high-performance memristive array as an in-memory readout network, it may become possible to implement local training and on-device weight updates. Such an approach could not only improve energy efficiency but also offer advantages in terms of privacy and security, and enable user-specific personalization of the speech-recognition module. Finally, multimodal integration—i.e., combining multiple task-specialized physical reservoir computing units—could provide a powerful machine–human interface for analyzing and processing biosignals from a wide range of advanced sensing devices beyond MEG. In particular, coordinated inference across functional RC units may help mitigate misrecognition caused by posture changes or motion-induced variations in throat vibrations.

3 | Conclusion

In this study, we demonstrated a real-time spoken-digit recognition system that integrates a throat-mounted π–gel-electret MEG sensor with an ion-gel/graphene-based IGR. The MEG captures mechanical vibrations generated during articulation, and the IGR directly processes these raw vibration signals without requiring any frequency-domain preprocessing such as MFCCs or cochleagrams. By exploiting the intrinsic nonlinear dynamics, high-dimensional state expansion, and memory inherent to the IGR, the system achieves effective linear separability of speech patterns within a compact physical reservoir. We first characterized the device properties of the multi-channel graphene IGR and showed that variations in channel geometry and drain bias yield diverse nonlinear transfer characteristics suitable for reservoir computing. When driven by MEG-derived pulse inputs, the IGR produced eight-dimensional current responses whose temporal evolution encoded digit-dependent features with rich nonlinear structure. These reservoir states enabled high classification performance even with minimal node counts. Using only 32 nodes, the system achieved 84.2% accuracy—substantially outperforming the MEG-only baseline—demonstrating the strong computational capacity of the IGR. By increasing the node sampling resolution, the accuracy further improved to 96.8% with 576 virtual nodes. The results highlight several key advantages of MEG–IGR physical reservoir computing: (i) elimination of preprocessing pipelines and spectral analysis, (ii) tunable trade-off between computational performance and reservoir size for edge deployment, (iii) robustness of throat-vibration sensing in challenging acoustic environments, and (iv) the potential for secure user authentication based on biomechanical signatures that are inherently difficult to spoof. These findings underscore the promise of IGR-based physical reservoirs as efficient and flexible computational front-ends for next-generation edge-AI sensing devices. Future work will explore integration of the MEG–IGR platform with on-device learning, multimodal sensing, and broader classes of vocal and non-vocal human–machine interfaces, further expanding its applicability to wearable, assistive, and secure edge-intelligent systems.

4 | Experimental Section/Methods

4.1 | Fabrication of the Gel-Electret MEG Sensor and Recording of Throat-Vibration Signals

The mechanoelectric generator-electret sensor utilized in this study was fabricated in a manner similar to a previously reported method [55]. We keep it brief here. An alkylated-pyrene liquid [55] was mixed with 1 wt.% of a gelator compound, *N*,*N*'-((1*S*,2*S*)-cyclohexane-1,2-diyl)didodecanamide [68], and soaked into polyurethane nonwoven fabric (MELTBLOWN NONWOVENS UC0040, Kuraray Kuraflex Co., Ltd., Japan). Corona-discharging parameters included an applied voltage of +80 kV, temperature of 100°C, heating for 30 min, and poling for 90 min. Positive discharging was performed with a corona discharging gun (GC90PN, Green Techno Co., Ltd., Japan). The sensors were fabricated on polyurethane (Breath Roll BR15, Nitto Co., Ltd., Japan, 13 μm thick) by applying silver-plated nylon fibers (AGposs, Mitsufuji Co., Ltd., Japan, fiber length = 3 mm), with a piece of copper tape attached at the end to connect to the wire (Figure 1a) [55]. Breath Roll was also used as the polyurethane masking layer. The flexible sensor was fixed just below the Adam's apple using Scotch mending tape (3 M Company, USA). The open-circuit voltage (V_{oc}) of the sensor during speech was recorded on an oscilloscope (MT-775, MotherTool Co., Ltd., Japan) equipped with a 1 MΩ input impedance probe. The bottom electrode of the sensor was connected to the probe, and the top electrode was grounded. Speech consisted of uttering each word from zero to nine 100 times, and recording the output on an oscilloscope each time.

4.2 | Fabrication of the Ion-Gel/Graphene IGR and Electrical Measurements

Monolayer graphene grown by chemical vapor deposition (CVD) and transferred onto a SiO_2/Si substrate (300 nm SiO_2) was

2199160x, 2026, 19, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/aelm.202600006 by Takashi Tsuchiya - National Institute For , Wiley Online Library on [14/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

purchased from Graphenea (Spain). The graphene was patterned by dry etching to form six independent channels: widths of 30 μm with lengths of 5, 20, and 100 μm (ch1–ch3), and widths of 80 μm with the same lengths (ch4–ch6), as illustrated in the inset of Figure 2a. Source and drain electrodes were defined by photolithography and electron-beam evaporation of Cr/Au (10/50 nm). The ion-gel electrolyte was synthesized by chemically gelling the ionic liquid 1-ethyl-3-methylimidazolium bis(trifluoromethylsulfonyl)imide (EMIm-TFSI). A polymer solution of poly(dimethylaminoethyl methacrylate) (PDMEMA) in toluene (200 μL) was mixed with EMIm-TFSI (1 mL) under stirring, followed by the addition of the cross-linker *N,N,N,N'*-tetra(trifluoromethanesulfonyl)-dodecane-1,12-diamine (C_{12}TFSA) (200 μL). After mixing for 30 min at 300 rpm, the solution was cast onto a Si wafer and heated at 100°C for 15 min to induce gelation. The resulting ion-gel sheet (~500 μm thick) was cut to size, laminated onto the graphene channels, and dried overnight in a vacuum chamber evacuated with a turbomolecular pump. A 1-μm-thick Au foil was placed on top of the gel and used as the common gate electrode. Electrical measurements were performed at room temperature in a vacuum cryostat, with all device contacts made via Al wire bonding. Gate voltage pulses were applied using the pulse measurement unit (PMU) of a 4200A-SCS semiconductor parameter analyzer (Keithley, USA), enabling time-resolved acquisition of the gate, drain, and source current responses from all six graphene channels.

4.3 | Readout Weight Training

As illustrated in Figure 4a, the temporal output of the reservoir for a given time window is obtained from the reservoir state vector $\boldsymbol{x}$ according to Equation (1). Here we describe the procedure used to compute the readout weight matrix W. Let $\boldsymbol{x}_1$ denote the reservoir state vector obtained from the first time window of the first training sample, and $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, … the vectors obtained sequentially from all subsequent time windows of all training samples. By stacking these vectors in order, we construct the reservoir state matrix $X = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_M)^{\mathrm{T}}$ where $M = \sum_{i=1}^{400} n_i$ and n_i is the number of time windows for the i-th training sample. For each state vector $\boldsymbol{x}_j$, we define a corresponding teacher vector $\boldsymbol{y}_{\mathrm{t},j}$ using a one-hot representation of the ground-truth digit label. By arranging these teacher vectors row-wise, we obtain the teacher matrix $Y_{\mathrm{t}} = (\boldsymbol{y}_{\mathrm{t},1}, \boldsymbol{y}_{\mathrm{t},2}, \ldots, \boldsymbol{y}_{\mathrm{t},M})^{\mathrm{T}}$ The readout weight matrix W was trained using ridge regression,

$$W = Y_{\mathrm{t}} X^{\mathrm{T}} \left(X X^{\mathrm{T}} + \lambda I \right)^{-1} \tag{2}$$

where $\lambda = 0.001$ is the regularization parameter and $I \in \mathbb{R}^{M \times M}$ is the identity matrix. Classification accuracy was evaluated using 5-fold cross-validation at the utterance (sample) level, with 400 samples used for training and 100 for testing in each split. Because all 500 utterances were acquired from a single speaker in a single recording session under identical sampling conditions, the train/test partitioning is not confounded by inter-speaker or inter-session variability. Nevertheless, to further strengthen the robustness of the evaluation and to eliminate any unintended bias, the utterances were randomly partitioned into five mutually exclusive folds while maintaining class balance (digits 0–9) in each split. Importantly, the partitioning was performed without overlap so that (i) no utterance appeared in both the training and test sets within a fold, and (ii) each utterance was used as test data exactly once across the five folds.

Acknowledgements

This research was in part supported by JST PRESTO Grant number JPMJPR23H4 and JSPS KAKENHI Grant Number JP24KJ0299 (Grant-in-Aid for JSPS Fellows), JP25K17941, and JP25H01264 (Grant-in-Aid for Transformative Research Areas "π-Molecular Complexity"). A part of this work was supported by "Advanced Research Infrastructure for Materials and Nanotechnology in Japan (ARIM)" of the Ministry of Education, Culture, Sports, Science and Technology (MEXT). This work was partially supported by Innovative Science and Technology Initiative for Security Grant Number JPJ004596, ATLA, Japan. We also acknowledge the World Premier International Research Center Initiative (WPI), MEXT, Japan. Proposal Number JPMXP1224NM5236 and JPMXP1225NM5329. D.N. acknowledges the International Center for Young Scientists (ICYS) at NIMS for providing ICYS fellowships. We thank Mr. Kota Fukuda (NIMS) for his assistance in acquiring data from the π–gel-electret MEG sensor.

Funding

This work was supported by JST PRESTO Grant number JPMJPR23H4, JSPS KAKENHI Grant Number JP24KJ0299 (Grant-in-Aid for JSPS Fellows), JP25K17941, JP25H01264 (Grant-in-Aid for Transformative Research Areas "π-Molecular Complexity"), and Innovative Science and Technology Initiative for Security Grant Number JPJ004596, ATLA, Japan.

Conflicts of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. J. M. Baker, L. Deng, J. Glass, et al., "Developments and Directions in Speech Recognition and Understanding, Part 1 [DSP Education]," *IEEE Signal Processing Magazine* 26 (2009): 75–80, https://doi.org/10.1109/MSP.2009.932166.

2. S. Davis and P. Mermelstein, "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences," *IEEE Transactions on Acoustics, Speech, and Signal Processing* 28 (1980): 357–366, https://doi.org/10.1109/TASSP.1980.1163420.

3. W. Han, C.-F. Chan, C.-S. Choy, and K.-P. Pun, "An Efficient MFCC Extraction Method in Speech Recognition," in *2006 IEEE International Symposium on Circuits and Systems (ISCAS)* (IEEE, 2006).

4. G. Hinton, L. Deng, D. Yu, et al., "Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups," *IEEE Signal Processing Magazine* 29 (2012): 82–97, https://doi.org/10.1109/MSP.2012.2205597.

5. Y. Gong, "Speech Recognition in Noisy Environments: A Survey," *Speech Communication* 16 (1995): 261–291, https://doi.org/10.1016/0167-6393(94)00059-J.

6. Q. Li, Y. Yang, T. Lan, et al., "MSP-MFCC: Energy-Efficient MFCC Feature Extraction Method with Mixed-Signal Processing Architecture for Wearable Speech Recognition Applications," *IEEE Access* 8 (2020): 48720–48730, https://doi.org/10.1109/ACCESS.2020.2979799.

21991160x, 2026, 19, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/aelm.202600006 by Takashi Tsuchiya - National Institute For , Wiley Online Library on [14/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

7. Y. Patel, S. Tanwar, R. Gupta, et al., "Deepfake Generation and Detection: Case Study and Challenges," *IEEE Access* 11 (2023): 143296–143323, https://doi.org/10.1109/ACCESS.2023.3342107.

8. Z. Wu, N. Evans, T. Kinnunen, J. Yamagishi, F. Alegre, and H. Li, "Spoofing and Countermeasures for Speaker Verification: A Survey," *Speech Communication* 66 (2015): 130–153, https://doi.org/10.1016/j.specom.2014.10.005.

9. S. Lee, J. Kim, I. Yun, et al., "An Ultrathin Conformable Vibration-Responsive Electronic Skin for Quantitative Vocal Recognition," *Nature Communications* 10 (2019): 2468, https://doi.org/10.1038/s41467-019-10465-w.

10. S. Gao, G. Li, S. Wu, et al., "High-Bandwidth Self-Powered Vibration Sensors Based on Triboelectric Particle-Surface Interactions," *Nano Energy* 119 (2024): 109060, https://doi.org/10.1016/j.nanoen.2023.109060.

11. Q. Zhang and K. Tong, "Vocal Cord Vibration Signal Recognition Model Based on Feature Engineering Preprocessing," *IEEE Sensors Journal* 23 (2023): 31380–31388, https://doi.org/10.1109/JSEN.2023.3321987.

12. T. Liu, M. Zhang, Z. Li, et al., "Machine Learning-Assisted Wearable Sensing Systems for Speech Recognition and Interaction," *Nature Communications* 16 (2025): 2363, https://doi.org/10.1038/s41467-025-57629-5.

13. J. Zhao, D. Chen, L. Zhao, et al., "Self-Powered Speech Recognition System for Deaf Users," *Cell Reports Physical Science* 3 (2022): 101168, https://doi.org/10.1016/j.xcrp.2022.101168.

14. S. Lee, H. Roh, J. Kim, et al., "An Electret-Powered Skin-Attachable Auditory Sensor That Functions in Harsh Acoustic Environments," *Advanced Materials* 34 (2022): 2205537, https://doi.org/10.1002/adma.202205537.

15. W. Li, R. Min, D. Zheng, et al., "Wearable Photonic Artificial Throat for Silent Communication and Speech Recognition," *ACS Applied Materials & Interfaces* 17 (2025): 11126–11143, https://doi.org/10.1021/acsami.4c21754.

16. K. Xiao, S. Zhou, H. Wang, X. Li, R. Min, and Z. Wang, "Photonic Biointerfaces: A New Paradigm for Seamless Integration with Living Systems Based on Flexible Optical Fibers," *Optics & Laser Technology* 197 (2026): 114776, https://doi.org/10.1016/j.optlastec.2026.114776.

17. R. Zhong, J. Pan, Q. Li, et al., "Cuffless Wearable Sphygmomanometers Using Dual Digital Optical Frequency Combs in Chalcogenide Chips," *Advanced Functional Materials* 35 (2025): 2417142, https://doi.org/10.1002/adfm.202417142.

18. W. Li, Y. Long, Y. Yan, et al., "Wearable Photonic Smart Wristband for Cardiorespiratory Function Assessment and Biometric Identification," *Opto-Electronic Advances* 8 (2025): 240254, https://doi.org/10.29026/oea.2025.240254.

19. H. Jaeger and H. Haas, "Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication," *Science* 304 (2004): 78–80, https://doi.org/10.1126/science.1091277.

20. G. Tanaka, T. Yamane, J. B. Héroux, et al., "Recent Advances in Physical Reservoir Computing: A Review," *Neural Networks* 115 (2019): 100–123, https://doi.org/10.1016/J.NEUNET.2019.03.005.

21. L. Appeltant, M. C. Soriano, G. Van der Sande, et al., "Information Processing Using a Single Dynamical Node as Complex System," *Nature Communications* 2 (2011): 468, https://doi.org/10.1038/ncomms1476.

22. Y. Paquot, F. Duport, A. Smerieri, et al., "Optoelectronic Reservoir Computing," *Scientific Reports* 2 (2012): 287, https://doi.org/10.1038/srep00287.

23. R. Martinenghi, S. Rybalko, M. Jacquot, Y. K. Chembo, and L. Larger, "Photonic Nonlinear Transient Computing with Multiple-Delay Wavelength Dynamics," *Physical Review Letters* 108 (2012): 244101, https://doi.org/10.1103/PhysRevLett.108.244101.

24. L. Larger, M. C. Soriano, D. Brunner, et al., "Photonic Information Processing beyond Turing: An Optoelectronic Implementation of Reservoir Computing," *Optics Express* 20 (2012): 3241–3249, https://doi.org/10.1364/OE.20.003241.

25. L. Larger, A. Baylón-Fuentes, R. Martinenghi, V. S. Udaltsov, Y. K. Chembo, and M. Jacquot, "High-Speed Photonic Reservoir Computing Using a Time-Delay-Based Architecture: Million Words per Second Classification," *Physical Review X* 7 (2017): 11015, https://doi.org/10.1103/PhysRevX.7.011015.

26. J. Torrejon, M. Riou, F. A. Araujo, et al., "Neuromorphic Computing with Nanoscale Spintronic Oscillators," *Nature* 547 (2017): 428–431, https://doi.org/10.1038/nature23011.

27. Y. Zhong, J. Tang, X. Li, B. Gao, H. Qian, and H. Wu, "Dynamic Memristor-Based Reservoir Computing for High-Efficiency Temporal Signal Processing," *Nature Communicators* 12 (2021): 408, https://doi.org/10.1038/s41467-020-20692-1.

28. S. Lilak, W. Woods, K. Scharnhorst, et al., "Spoken Digit Classification by In-Materio Reservoir Computing with Neuromorphic Atomic Switch Networks," *Frontiers in Nanotechnology* 3 (2021): 675792, https://doi.org/10.3389/fnano.2021.675792.

29. J. Moon, W. Ma, J. H. Shin, et al., "Temporal Data Classification and Forecasting Using a Memristor-Based Reservoir Computing System," *Nature Electronics* 2 (2019): 480–487, https://doi.org/10.1038/s41928-019-0313-3.

30. R. Midya, Z. Wang, S. Asapu, et al., "Reservoir Computing Using Diffusive Memristors," *Advanced Intelligent Systems* 1 (2019): 1900084, https://doi.org/10.1002/aisy.201900084.

31. C. Du, F. Cai, M. A. Zidan, W. Ma, S. H. Lee, and W. D. Lu, "Reservoir Computing Using Dynamic Memristors for Temporal Information Processing," *Nature Communicators* 8 (2017): 2204, https://doi.org/10.1038/s41467-017-02337-y.

32. Y. Usami, B. van de Ven, D. G. Mathew, et al., "In-Materio Reservoir Computing in a Sulfonated Polyaniline Network," *Advanced Materials* 33 (2021): 2102688, https://doi.org/10.1002/adma.202102688.

33. G. Milano, G. Pedretti, K. Montano, et al., "In Materia Reservoir Computing with a Fully Memristive Architecture Based on Self-Organizing Nanowire Networks," *Nature Materials* 21 (2022): 195–202, https://doi.org/10.1038/s41563-021-01099-9.

34. O. Lee, T. Wei, K. D. Stenning, et al., "Task-Adaptive Physical Reservoir Computing," *Nature Materials* 23 (2024): 79–87, https://doi.org/10.1038/s41563-023-01698-8.

35. J. Hochstetter, R. Zhu, A. Loeffler, A. Diaz-Alvarez, T. Nakayama, and Z. Kuncic, "Avalanches and Edge-of-Chaos Learning in Neuromorphic Nanowire Networks," *Nature Communications* 12 (2021): 4008, https://doi.org/10.1038/s41467-021-24260-z.

36. S. Kan, K. Nakajima, T. Asai, and M. Akai-Kasaya, "Physical Implementation of Reservoir Computing through Electrochemical Reaction," *Advanced Science* 9 (2022): 2104076, https://doi.org/10.1002/advs.202104076.

37. K. Toprasertpong, E. Nako, Z. Wang, R. Nakane, M. Takenaka, and S. Takagi, "Reservoir Computing on a Silicon Platform with a Ferroelectric Field-Effect Transistor," *Communications Engineering* 1 (2022): 21, https://doi.org/10.1038/s44172-022-00021-8.

38. W. Namiki, Y. Yamaguchi, D. Nishioka, T. Tsuchiya, and K. Terabe, "Opto-Magnonic Reservoir Computing Coupling Nonlinear Interfered Spin Wave and Visible Light Switching," *Materials Today Physics* 45 (2024): 101465, https://doi.org/10.1016/j.mtphys.2024.101465.

39. W. Namiki, D. Nishioka, Y. Yamaguchi, T. Tsuchiya, T. Higuchi, and K. Terabe, "Experimental Demonstration of High-Performance Physical Reservoir Computing with Nonlinear Interfered Spin Wave Multidetection," *Advanced Intelligent Systems* 5 (2023): 2300228, https://doi.org/10.1002/aisy.202300228.

40. W. Namiki, D. Nishioka, Y. Nomura, T. Tsuchiya, K. Yamamoto, and K. Terabe, "Iono–Magnonic Reservoir Computing with Chaotic Spin Wave Interference Manipulated by Ion-Gating," *Advanced Science* 12 (2024): 2411777, https://doi.org/10.1002/advs.202411777.

2199160x, 2026, 19, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/aelm.202600006 by Takashi Tsuchiya - National Institute For , Wiley Online Library on [14/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

41. D. Nishioka, T. Tsuchiya, W. Namiki, et al., "Edge-of-Chaos Learning Achieved by Ion-Electron–Coupled Dynamics in an Ion-Gating Reservoir," *Science Advances* 8 (2022): ade1156, https://doi.org/10.1126/sciadv.ade1156.

42. T. Wada, D. Nishioka, W. Namiki, T. Tsuchiya, T. Higuchi, and K. Terabe, "A Redox-Based Ion-Gating Reservoir, Utilizing Double Reservoir States in Drain and Gate Nonlinear Responses," *Advanced Intelligent Systems* 5 (2023): 2300123, https://doi.org/10.1002/aisy.202300123.

43. K. Shibata, D. Nishioka, W. Namiki, T. Tsuchiya, T. Higuchi, and K. Terabe, "Redox-Based Ion-gating Reservoir Consisting of (104) Oriented $LiCoO_2$ Film, Assisted by Physical Masking," *Scientific Reports* 13 (2023): 21060, https://doi.org/10.1038/s41598-023-48135-z.

44. D. Nishioka, Y. Shingaya, T. Tsuchiya, T. Higuchi, and K. Terabe, "Few- and Single-Molecule Reservoir Computing Experimentally Demonstrated with Surface-Enhanced Raman Scattering and Ion Gating," *Science Advances* 10 (2024): adk6438, https://doi.org/10.1126/sciadv.adk6438.

45. D. Nishioka, T. Tsuchiya, M. Imura, Y. Koide, T. Higuchi, and K. Terabe, "A High-performance Deep Reservoir Computer Experimentally Demonstrated with Ion-Gating Reservoirs," *Communications Engineering* 3 (2024): 81, https://doi.org/10.1038/s44172-024-00227-y.

46. M. Takayanagi, D. Nishioka, T. Tsuchiya, et al., "Ultrafast-Switching of an All-Solid-State Electric Double Layer Transistor with a Porous Yttria-Stabilized Zirconia Proton Conductor and the Application to Neuromorphic Computing," *Materials Today Advances* 18 (2023): 100393, https://doi.org/10.1016/j.mtadv.2023.100393.

47. Y. Yamaguchi, D. Nishioka, W. Namiki, et al., "Inverted Input Method for Computing Performance Enhancement of the Ion-Gating Reservoir," *Applied Physics Express* 17 (2024): 024501, https://doi.org/10.35848/1882-0786/ad2906.

48. W. Namiki, D. Nishioka, T. Tsuchiya, T. Higuchi, and K. Terabe, "Magnetization Vector Rotation Reservoir Computing Operated by Redox Mechanism," *Nano Letters* 24 (2024): 4383–4392, https://doi.org/10.1021/acs.nanolett.3c05029.

49. T. Tsuchiya, M. Takayanagi, D. Nishioka, W. Namiki, and K. Terabe, "Electric Double Layer Effect in the Vicinity of Solid Electrolyte/Diamond Interfaces and the Application to Neuromorphic Computing," *Journal of Solid State Electrochemistry* 28 (2024): 4367–4376, https://doi.org/10.1007/s10008-024-05937-z.

50. D. Nishioka, T. Tsuchiya, T. Higuchi, and K. Terabe, "Enhanced Synaptic Characteristics of H_xWO_3-Based Neuromorphic Devices, Achieved by Current Pulse Control, for Artificial Neural Networks," *Neuromorphic Computing and Engineering* 3 (2023): 034008, https://doi.org/10.1088/2634-4386/acf1c6.

51. H. Kitano, D. Nishioka, K. Terabe, and T. Tsuchiya, "Physical Reservoir Computing with Graphene-Based Solid Electric Double Layer Transistor and the Information Processing Capacity Analysis," *Applied Physics Express* 18 (2025): 024501, https://doi.org/10.35848/1882-0786/adb19b.

52. T. Tsuchiya, D. Nishioka, W. Namiki, and K. Terabe, "Physical Reservoir Computing Utilizing Ion-Gating Transistors Operating in Electric Double Layer and Redox Mechanisms," *Advanced Electronic Materials* 10 (2024): 2400625, https://doi.org/10.1002/aelm.202400625.

53. D. Nishioka, H. Kitano, W. Namiki, S. Souma, K. Terabe, and T. Tsuchiya, "Two Orders of Magnitude Reduction in Computational Load Achieved by Ultrawideband Responses of an Ion-Gating Reservoir," *ACS Nano* 19 (2025): 36896–36914, https://doi.org/10.1021/acsnano.5c06174.

54. D. Nishioka, K. Shibata, W. Namiki, K. Terabe, and T. Tsuchiya, "Physical Masking-Induced Enhancement of Information Processing Capacity in a Redox-Type Ion-Gating Reservoir," *Japanese Journal of Applied Physics* 64 (2025): 11SP25, https://doi.org/10.35848/1347-4065/ae1d84.

55. A. Tateyama, K. Nagura, M. Yamanaka, and T. Nakanishi, "Alkyl–π Functional Molecular Gels: Control of Elastic Modulus and Improvement of Electret Performance," *Angewandte Chemie International Edition* 63 (2024): 202402874, https://doi.org/10.1002/anie.202402874.

56. A. Ghosh, M. Yoshida, K. Suemori, et al., "Soft Chromophore Featured Liquid Porphyrins and Their Utilization toward Liquid Electret Applications," *Nature Communications* 10 (2019): 4210, https://doi.org/10.1038/s41467-019-12249-8.

57. R. K. Gupta, M. Yoshida, A. Saeki, Z. Guo, and T. Nakanishi, "Alkyl-C_{60} Liquid Electrets as Deformable Mechanoelectric Generators," *Materials Horizons* 10 (2023): 3458–3466, https://doi.org/10.1039/D3MH00485F.

58. A. Shinohara, M. Yoshida, C. Pan, and T. Nakanishi, "Stretchable π-Conjugated Polymer Electrets for Mechanoelectric Generators," *Polymer Journal* 55 (2023): 529–535, https://doi.org/10.1038/s41428-022-00725-w.

59. D. Nishioka, T. Tsuchiya, T. Higuchi, and K. Terabe, "Oxygen-Tolerant Operation of All-Solid-State Ionic-Gating Devices: Advantage of All-Solid-State Structure for Ionic-Gating," *Japanese Journal of Applied Physics* 59 (2020): SIIG09, https://doi.org/10.35848/1347-4065/ab7e12.

60. M. Takayanagi, T. Tsuchiya, D. Nishioka, et al., "Accelerated/Decelerated Dynamics of the Electric Double Layer at Hydrogen-Terminated Diamond/Li^+ Solid Electrolyte Interface," *Materials Today Physics* 31 (2023): 101006, https://doi.org/10.1016/j.mtphys.2023.101006.

61. T. Tsuchiya, M. Takayanagi, K. Mitsuishi, et al., "The Electric Double Layer Effect and Its Strong Suppression at Li^+ Solid Electrolyte/Hydrogenated Diamond Interfaces," *Communications Chemistry* 4 (2021): 117, https://doi.org/10.1038/s42004-021-00554-7.

62. H. Wang, Y. Wu, C. Cong, J. Shang, and T. Yu, "Hysteresis of Electronic Transport in Graphene Transistors," *ACS Nano* 4 (2010): 7221–7228, https://doi.org/10.1021/nn101950n.

63. A. Tries, N. Richter, Z. Chen, et al., "Hysteresis in Graphene Nanoribbon Field-Effect Devices," *Physical Chemistry Chemical Physics* 22 (2020): 5667–5672, https://doi.org/10.1039/D0CP00298D.

64. M. C. Lemme, T. J. Echtermeyer, M. Baus, and H. Kurz, "A Graphene Field-Effect Device," *IEEE Electron Device Letters* 28 (2007): 282–284, https://doi.org/10.1109/LED.2007.891668.

65. S. Zhao, B. Zhang, J. Yang, J. Zhou, and Y. Xu, "Linear Discriminant Analysis," *Nature Reviews Methods Primers* 4 (2024): 70, https://doi.org/10.1038/s43586-024-00346-y.

66. J. Dambre, D. Verstraeten, B. Schrauwen, and S. Massar, "Information Processing Capacity of Dynamical Systems," *Scientific Reports* 2 (2012): 514.

67. M. Zolfagharinejad, J. Büchel, L. Cassola, et al., "Analogue Speech Recognition Based on Physical Computing," *Nature* 645 (2025): 886–892, https://doi.org/10.1038/s41586-025-09501-1.

68. K. Hanabusa, M. Yamada, M. Kimura, and H. Shirai, "Prominent Gelation and Chiral Aggregation of Alkylamides Derived from Trans - 1,2-Diaminocyclohexane," *Angewandte Chemie International Edition in English* 35 (1996): 1949–1951, https://doi.org/10.1002/anie.199619491.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Supporting File: aelm70410-sup-0001-SuppMat.docx.

21999160x, 2026, 19, Downloaded from https://advanced.onlinelibrary.wiley.com/doi/10.1002/aelm.202600006 by Takashi Tsuchiya - National Institute For , Wiley Online Library on [14/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License