



A simplified wearable device powered by a generative EMG network for hand-gesture recognition and gait prediction

Received: 4 May 2025

Accepted: 27 September 2025

Published online: 01 December 2025

 Check for updates

Kyun Kyu Kim¹, Tomasz J. Zaluska¹, Six Skov², Yeongjun Lee¹, Hyunchang Park¹, Donglai Zhong¹, Muhammad Khatib¹, Yuya Nishio³, Yuanwen Jiang¹, Scott L. Delp² & Zhenan Bao¹  

There is an inherent trade-off between the design of comfortable wearable sensors and the need for high-resolution and large-area sensing. Here we address this challenge by introducing a generative electromyography (EMG) network (GenENet), a self-supervised generative representation learning framework combined with a wearable sensor that extrapolates limited sensor inputs to reconstruct muscle activity in unseen regions not covered by the sensor. This approach allows for gathering information equivalent to those from high-density EMG sensor networks, but using a more compact, wearable device with much reduced sensor counts, without sacrificing performance. For example, a 6-channel EMG device trained on a 32-channel dataset from low-impedance polymer electrodes shows similar performance to a 32-channel device in accuracy for predicting sign language and gait dynamics, highlighting the utility of the GenENet concept to reduce wearable-sensing-system complexity while maintaining prediction quality. Such electrophysiological data are essential for many applications, including motion detection and control, but traditionally require high-density, high-resolution sensors that can be cumbersome. This concept should be applicable to other types of electrophysiological mapping application across health monitoring, prosthetics, sports and human-machine interfaces, paving the way for more comfortable and efficient wearable devices.

Accurate information extraction in many sensing applications hinges on the deployment of high-density and high-resolution sensors^{1–3}. Wearable technologies are at the forefront of this development, delivering a rich, continuous stream of data that captures a broad range of biometric details. These range from kinematic information, such as movement and vibrations^{1,2,4–11}, to physiological signals^{12–14}, including temperature^{15,16} and various electrophysiological markers derived from muscle^{4,17}, brain^{12,18,19} and cardiac^{20–24} activities.

Electromyography (EMG) is widely used for non-invasive monitoring of muscle activity across different body regions^{3–5}. The complex interplay between muscle groups and the corresponding body movements calls for sophisticated high-density surface EMG instruments capable of capturing intricate kinematics, such as gesture recognition^{3,5} and gait analysis^{25,26}. Despite advancements in electrode design to reduce impedance and enhance flexibility, as well as in computational algorithms to improve signal processing and predictive accuracy, a

¹Department of Chemical Engineering, Stanford University, Stanford, CA, USA. ²Department of Mechanical Engineering, Stanford University, Stanford, CA, USA. ³Department of Electrical Engineering, Stanford University, Stanford, CA, USA. ✉ e-mail: zbao@stanford.edu

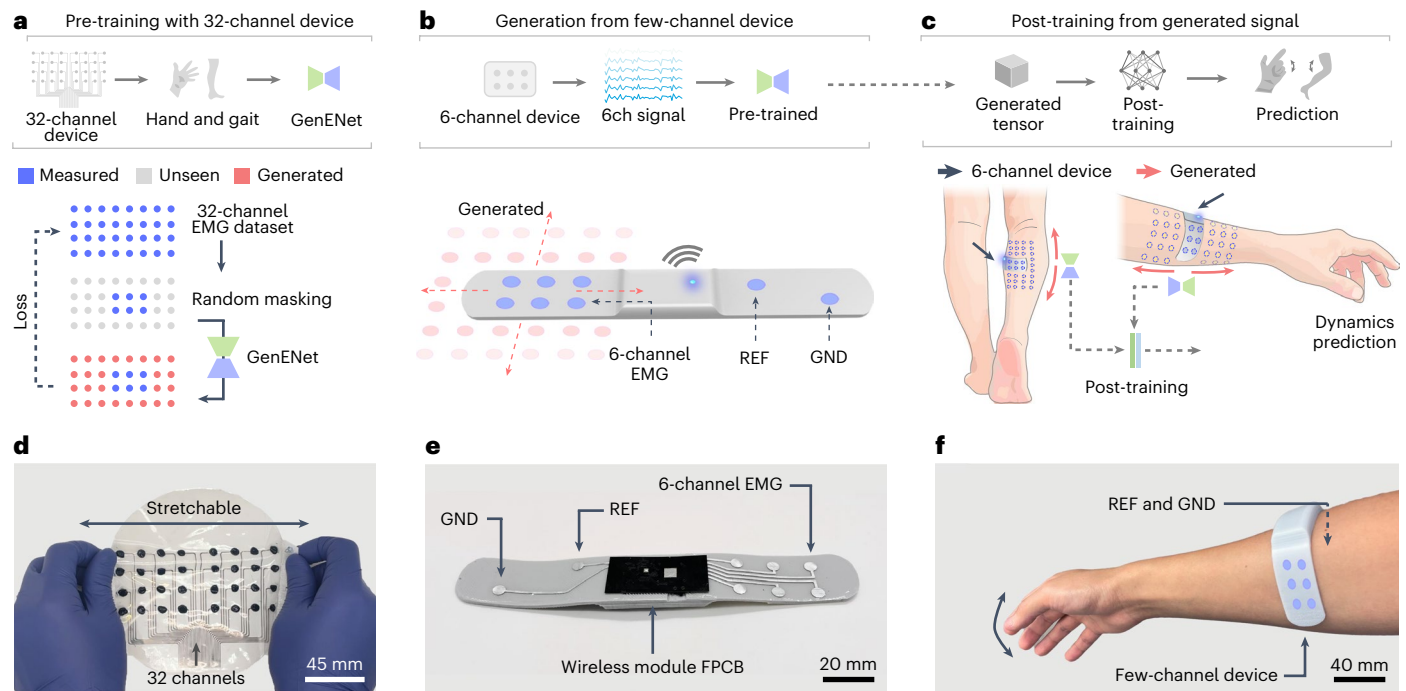


Fig. 1 | GenENet for sign language and gait prediction. **a**, Representation learning via GenENet using a 32-channel stretchable device, with random masking of input signals to reconstruct the original data. **b**, Use of a smaller 6-channel device, where the pre-trained network predicts muscle activity in unseen areas. Reference (REF) and ground (GND) electrodes are placed on the

right side. **c**, Post-training network, enabling the transfer of the pre-trained model to different applications and users. **d–f**, Images showing each stage corresponding to **a–c**, respectively. The wireless module consists of flexible printed circuit board (FPCB).

notable challenge remains: EMG devices with a large area and high electrode count are required for accuracy, but this increases the size of the devices and limits their practicality and adoption.

To address this challenge, here we introduce a generative EMG network (GenENet) that enables predicting EMG information from a smaller set and area of sensor inputs. GenENet is an autoencoder-based self-supervised generative representation learning algorithm trained to generate unseen sensor signals, discerning generalized patterns across sensor activations from a larger body area. Leveraging this prior-learned network of associations between sparse sensory inputs and providing a latent representation approximating the corresponding full-scale sensory mapping, our approach demonstrates an enhanced capability for a simple low-channel-count device to predict a spectrum of body kinematics previously reliant on high-density sensor arrays^{4,5,27}. To further enhance the efficacy of our system for wearable applications, we integrated it with stretchable sensors and low-impedance electrodes. These sensors offer superior conformity to the skin and facilitate the acquisition of low-noise signals that enable high performance of the generative network by reducing motion artefacts and contact impedance.

To develop GenENet, first, we designed a 32-channel soft EMG array for initial data collection to support prior learning, feeding this into our generative algorithm while omitting around 80% of the data. The model then learns to reconstruct the missing information by comparing the generated signals to the complete 32-channel signals (Fig. 1a). This pre-trained model is then integrated with a simplified 6-channel device, which, despite fewer sensing channels, generates a vector equivalent to that from a 32-channel array (Fig. 1b). The information is subsequently processed by the post-training network, which extracts crucial features of spatiotemporal muscle activities to classify static states and sequential movements, applicable to tasks such as American Sign Language (ASL) translation and gait dynamics prediction (Fig. 1c). Corresponding images of devices used for dataset collection are shown in Fig. 1d–f.

A stretchable sensor for a high-quality dataset

The objective of the pre-training generative algorithm is to reconstruct the original temporal and spatial patterns of muscle activity signals with high fidelity using minimal sensor inputs. Achieving high accuracy from the generative model is contingent on the quality of the training dataset. To facilitate the creation of a high-quality dataset, we have engineered a fully stretchable multi-array EMG device.

The device's layer is depicted in an exploded view in Fig. 2a. It comprises several layers: a polydimethylsiloxane (PDMS) substrate, a solvent resistant poly(acrylonitrile-co-butadiene) (NBR) protective layer, a gallium–indium eutectic (EGaIn) electrode, a poly(3,4-ethylenedioxythiophene):poly(styrenesulfonate) (PEDOT:PSS) gel and a poly(styrene-butadiene-styrene) (SBS) encapsulation. The PDMS layer provides a thin substrate with submillimetre thickness for flexible handling and easy adherence to various body contours. The NBR layer acts as a barrier to protect the PDMS from solvent-induced swelling during the fabrication of the multiple layers. The EGaIn liquid metal electrode, which is both stretchable and micro-patterned, was then applied, followed by a highly conductive PEDOT:PSS gel that serves to lower impedance when in contact with skin. The device is completed with a photo-patterned SBS layer that leaves only the electrode areas uncovered, ensuring skin contact and signal measurement. A biomedical adhesive (Skinister) is applied on the skin to ensure a conformal and secure attachment of the device before usage. Figure 2b illustrates the side view of the sensor array and the fabricated stretchable array. The device shows resilience to elongation, utilizing an intrinsically stretchable substrate along with electrodes made of EGaIn and PEDOT:PSS, as shown in Fig. 2c. It is then connected to a wireless device via a flexible flat cable, which is interfaced using anisotropic z-axis conductive tape (Fig. 2d and Extended Data Fig. 1).

Our conductive adhesive electrode is composed of an acrylamide (AAm) crosslinked gel with a PEDOT:PSS conducting polymer network. This provides both good electrical and mechanical properties, improving impedance characteristics and minimizing noise from movement.

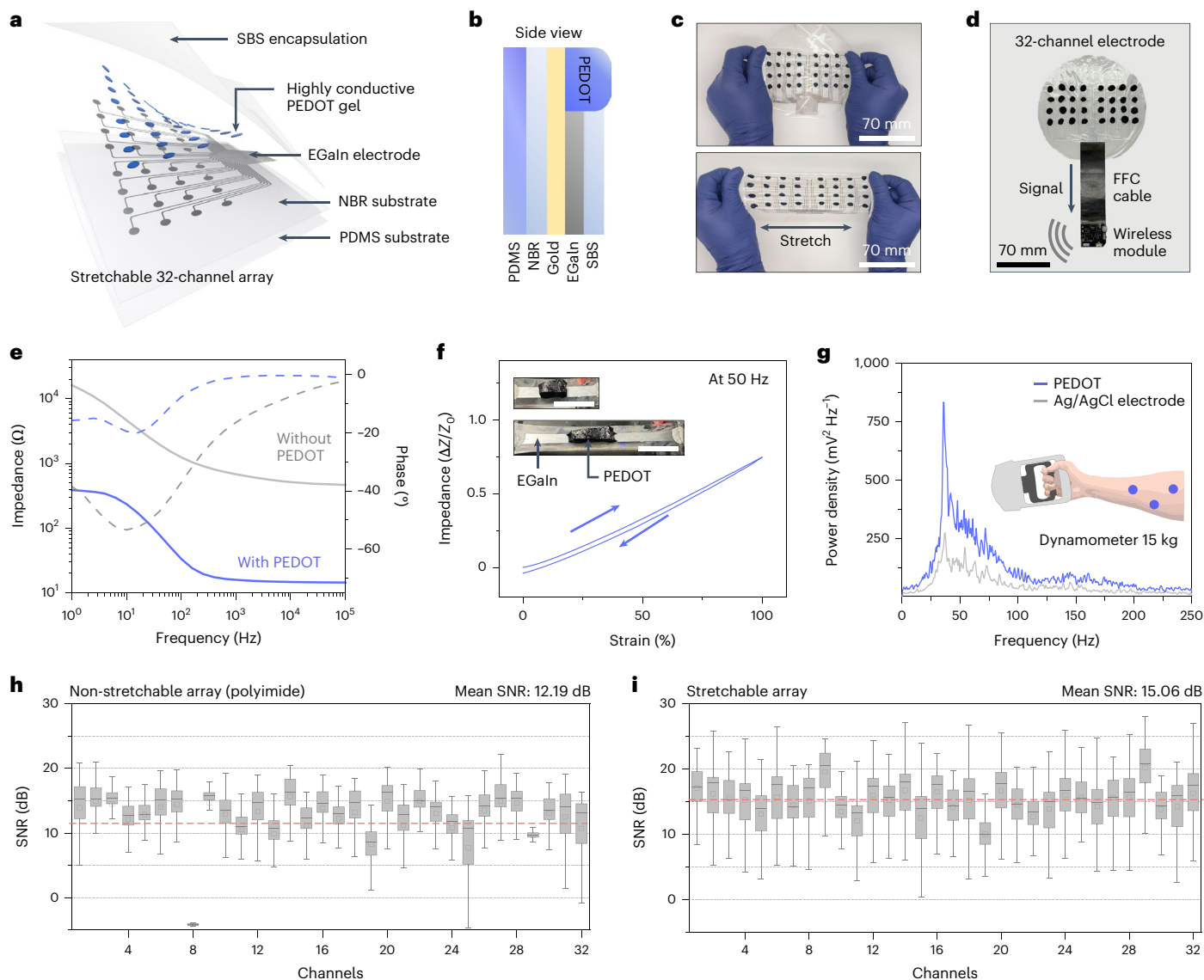


Fig. 2 | Stretchable sensory array for high-quality dataset generation. a, An exploded view of the 32-channel stretchable array, showing encapsulation, sensing electrodes, interconnections and substrates. **b,** Side view of the 32-channel device. **c,** Comparison of the device in its original and stretched states. **d,** The 32-channel device connected via a flexible flat connector (FFC) to a custom-made wireless module (Extended Data Fig. 1). **e,** Electrochemical impedance spectroscopy of hydrogels with (15 mg ml⁻¹ PEDOT:PSS with 150 mg ml⁻¹ AAm and 2.5 mg ml⁻¹ *N,N'*-methylenebisacrylamide) and without PEDOT:PSS. **f,** Impedance endurance plot under 100% strain. The inset image shows PEDOT interconnects under 0% and 100% strain. Z and Z_0 denote the

impedance under strain and at the unstrained state, respectively. Scale bars, 20 mm. **g,** Welch's power spectral density comparing PEDOT gel with Ag/AgCl electrodes of the same size. The inset illustrates the measurement setup using a dynamometer with electrodes attached to the forearm. **h,i,** SNR box plot of the device fabricated on a non-stretchable polyimide substrate (**h**) versus the device fabricated on a stretchable substrate (**i**), showing higher mean SNR across the 32 channels for the stretchable substrate. Data are presented as box plots showing the median (centre line), 25th–75th percentiles (box), and whiskers extending to the most extreme data points within 1.5× the interquartile range (IQR). The red dashed line indicates the mean value of the SNRs.

Our modifications to the AAm polymer network with the addition of PEDOT:PSS have resulted in a notable reduction in interfacial impedance when interfaced with phosphate-buffered saline (PBS), as seen in Fig. 2e. The impedance of the PEDOT-modified gel remains low at 100% strain at 50 Hz (Fig. 2f and Supplementary Fig. 1). Moreover, upon performing Welch's power spectral density estimate, the developed sensor array shows a higher power density compared with a standard Ag/AgCl electrode of the same size under 15 kg of grasp, as depicted in Fig. 2g.

To explore the capabilities of using a stretchable device, we fabricated the PEDOT electrodes above a non-stretchable thin polyimide substrate (20 μm thickness versus 200 μm used for the stretchable array, Supplementary Fig. 2). When subjected to a 20 kg grasp, applied at the same location on the wrist, the non-stretchable array showed a

lower mean signal-to-noise ratio (SNR) of 12.19 dB across the channel, as shown in Fig. 2h. In contrast, the stretchable array showed a consistently higher mean SNR (15.06 dB), as depicted in Fig. 2i.

Pre-training generalizable representations of EMG signals

The raw signals captured by the 32-channel sensory array (Fig. 3a, Supplementary Fig. 3 and Extended Data Fig. 2) were subjected to post-processing, which involved the calculation of root-mean-square (RMS) values across 32 time windows, using a sliding window with a size of 10, as depicted in Fig. 3b. The training dataset was selected from the arbitrary movement of fingers and gait movements attached to wrist (20 min; Methods) and calf (15 min walking; Methods). This large

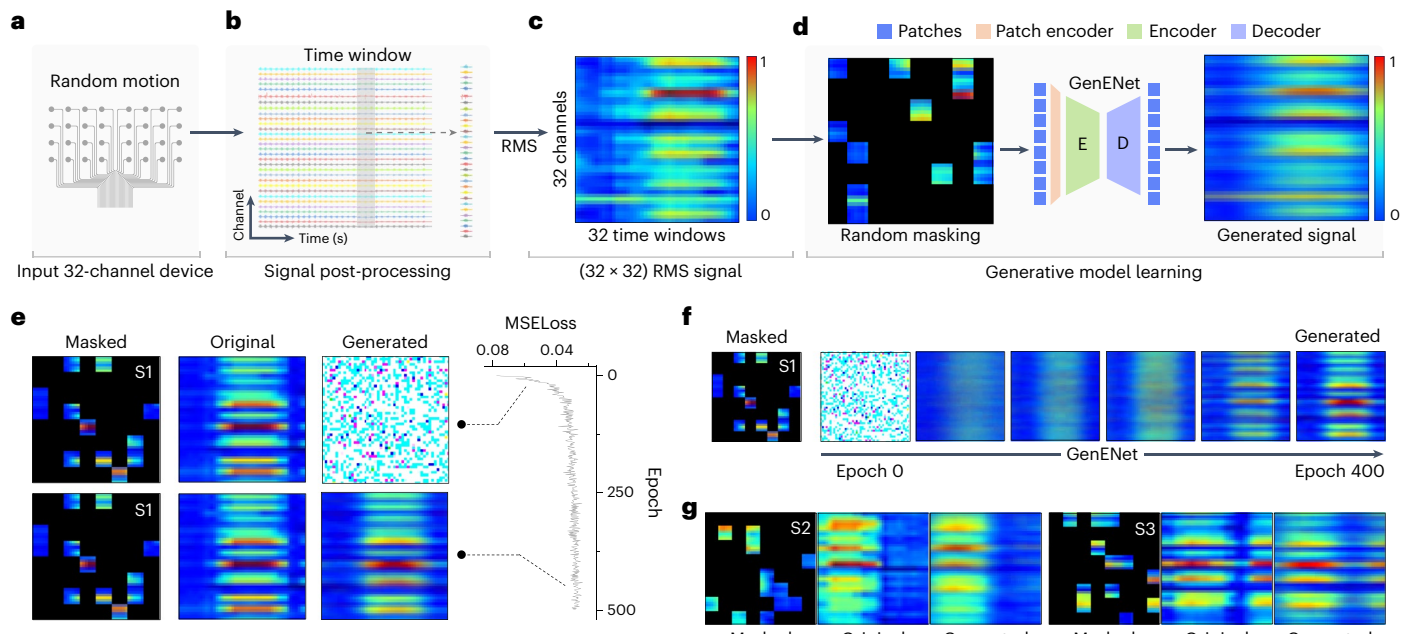


Fig. 3 | Pre-training of the GenENet. **a**, The 32-channel stretchable device captures muscle activation signals from the wrist or calf during arbitrary finger movements and walking. **b, c**, The signals undergo augmentation (**b**) and RMS processing (**c**) using a 32 time windows (32 time steps with 128 ms) with a sliding time (window interval; 10 time steps with 40 ms). The colour bars indicate normalized amplitude. **d**, Random masking (black regions) of the post-processed tensor, with GenENet trained to minimize the MSE loss between the generated and original signals. E and D denote the encoder and decoder modules of

GenENet, respectively. **e**, Representative signal from sample 1 (S1), showing the masked, original and generated signals across the training epochs. The first row shows the generated output from an early epoch and the second row shows the output after 450 epochs. The plot on the right shows the decrease in mean squared error (MSE loss) during training. **f**, Detailed plot of the generated signal over the course of training. **g**, Results for additional samples, sample 2 (S2) and sample 3 (S3) after 450 epochs. The samples were randomly selected; an additional 100 representative samples are shown in Extended Data Fig. 3.

dataset allows for pre-training of the generative algorithm to capture generalized muscle signal activities during the kinematic cycles.

To mitigate the high computational demands associated with conventional spectrogram conversion, we used RMS values as a simplified representation of the EMG signals (Methods and Supplementary Video 1). Consequently, an EMG signal tensor with dimensions of $32 \times 32 \times 1$ (32 time windows \times 32 channels \times 1 signal) is generated as shown in Fig. 3c. Subsequently, we deliberately masked some patches in the complete sensor data as shown in Fig. 3d. These masked data were then fed into GenENet. To strengthen the model's resilience against varying sensor attachment positions and orientations, we randomly distributed the masking locations, and the input of the tensor is augmented through resizing, random cropping and flipping for model optimization and generalizability (Methods). The architecture of our learning model adopts an autoencoder framework, principally divided into two sections: an encoder and a decoder. The tensor is divided into 64 squared patches, with 52 patches (81%) randomly masked. The encoder translates the intentionally random masked signals into a latent representation and the decoder reconstructs the missing patches in the pixel space. The encoder–decoder layers consist of multi-head attention, normalization and linear layers. The main structure is based on generalized denoising autoencoders, a scalable self-supervised learning approach used in computer vision^{28–30}.

Figure 3e illustrates that through the training cycles, GenENet successfully reconstructed the original signals using the masked inputs. A masked input, the original unmasked signal and the generated signal for sample 1 (S1) are shown across the training epochs. As the mean squared error loss (MSE loss) between the generated signal map and the original decreases during training, the generated signal map evolves from random noise to a much clearer representation (Supplementary Fig. 4). A detailed view of the generated signal map throughout the training epochs is depicted in Fig. 3f. The application of GenENet to

additional samples, S2 and S3, is shown in Fig. 3g, with a generated view of 100 additional samples presented in Extended Data Fig. 3. The fidelity of the reconstructed EMG signals is shown through signal distribution histograms over a training batch ($n = 128$) and uniform manifold approximation and projection visualization (Supplementary Fig. 5)

Downstream adaptation of GenENet for sign language prediction

To demonstrate practical deployment, we designed a compact six-channel wireless EMG watch, depicted in Fig. 4a, for gesture prediction (Extended Data Fig. 4). Our initial dataset consisted of EMG recordings corresponding to the 26 hand gestures representing the ASL alphabet, from A to Z. Using the pre-trained GenENet, we refined the model to enhance its predictive capabilities for ASL interpretation. Previously, achieving accurate predictions for a large number of gestures using EMG has required a dense array of electrodes, often exceeding 64 channels^{5,17,27} (Supplementary Table 1). Others used non-EMG sensors, specifically capacitive and piezoelectric sensors that cover the entire finger^{2,8,9}. While these glove-type sensors offer easier tracking of finger movements, they are limited in terms of wearable form factor and are affected by motion artefacts. In contrast, our approach requires only a band placed above the wrist.

Previously, EMG and speech-recognition tasks often used attention-based encoder–decoder frameworks^{31,32} or convolutional neural network (CNN)–transformer hybrids (for example, Conformer^{33,34}). In contrast, our method directly leverages the generalized representations learned through masked generative pre-training and feeds them into a lightweight long short-term memory (LSTM) classifier. This approach not only improves predictive performance but also exploits the masked autoencoding framework to operate effectively on intentionally masked hardware inputs, thereby reducing sensing complexity and enabling hardware-efficient deployment.

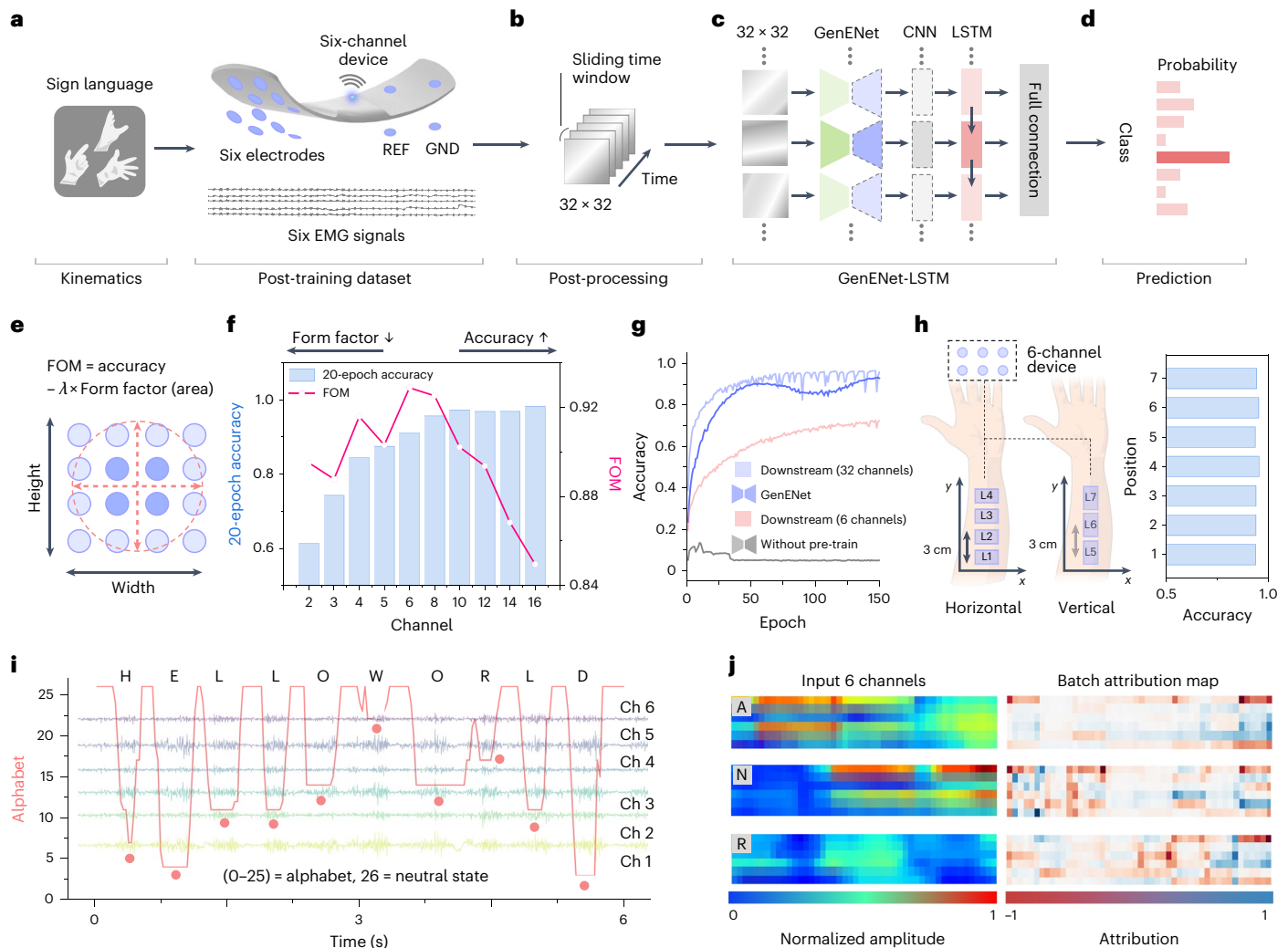


Fig. 4 | Prediction of ASL using the GenENet device. **a**, Sign language input signals captured through the six-channel device. **b**, Post-processing steps identical to pre-training, excluding data augmentation. **c**, Post-processed tensors are fed into GenENet, connected to a CNN, LSTM and dense layer. The dashed line of the decoder and CNN are only activated on regression modelling. **d**, Classification of sign language gestures. **e**, FOM measured by balancing model accuracy and total sensor area. **f**, FOM peaks in the six-channel region, where increasing channel count enhances accuracy but also enlarges the sensor area. **g**, Validation accuracy comparison between the pre-trained GenENet and the non-parameterized GenENet using six-electrode EMG array measurements for finger motion recognition. The dataset is divided into training and validation

datasets with a ratio of 8:2. **h**, Adaptability of the device to different locations and orientations on the wrist, showing negligible accuracy differences. L1–L7 indicate the location and orientation of the electrode array attachment. **i**, Sign language prediction using numeric values (0–25, that is 0 for A and 25 for Z) from 6-channel EMG inputs. The red plot represents the predicted numeric values, and the corresponding alphabet labelled on top, from the EMG input signals. **j**, Batch attribution map for representative letters A, N and R, with corresponding EMG signals and attribution maps. ‘-1’ indicates a negative contribution to prediction and ‘1’ indicates a positive contribution to prediction. Each row in the signal corresponds to one of the six sensor channels, with channel 1 at the bottom and channel 6 at the top.

As shown in Fig. 4b, the collected channel data undergo post-processing, which includes calculating RMS values across time windows and sliding windows, as described in the generative algorithm pre-training in Fig. 3. The remaining channel information is zero-padded and fed into the encoder to generate a latent vector, which is subsequently processed through an LSTM network with a sequence length of seven (Supplementary Fig. 6), as depicted in Fig. 4c. Finally, a dense layer with an output size of 26 is used to predict the entire alphabet classes (Fig. 4d).

The selection of 6 channels and its performance collaborating with the generative algorithm was investigated next. As shown in Fig. 4e, we identified a trade-off between the number of channels and performance, which is influenced by the total electrode area (assuming each electrode is of the same size) of the EMG device versus its accuracy. We experimented with different channel numbers ranging from 2 to 16 and compared the 20-epoch accuracy using the same dataset

(Supplementary Fig. 7). The figure of merit (FOM) was defined as $\text{Accuracy} - \lambda \times \text{Form factor}$, where the form factor is the area (mm^2) and $\lambda = 0.0005$. The parameter λ is chosen to appropriately scale the form factor (measured in mm^2) to align with the accuracy range. This ensures that the units and magnitudes of both components are balanced, making the FOM meaningful and interpretable within the desired range. The results showed that the FOM peaked with a moderate number of channels. A smaller number of EMG electrodes led to a substantial drop in performance due to limited signal information, while a larger number of electrodes improved performance but substantially increased the device's area, negatively impacting user comfort.

As shown in Fig. 4g, the pre-trained GenENet model, trained on a large dataset of random finger motions, outperformed the non-pre-trained model. The pre-trained GenENet, benefiting from unsupervised learning that captures the correlation of motion signal mappings, reached 93.6% validation accuracy within 150 transfer training epochs

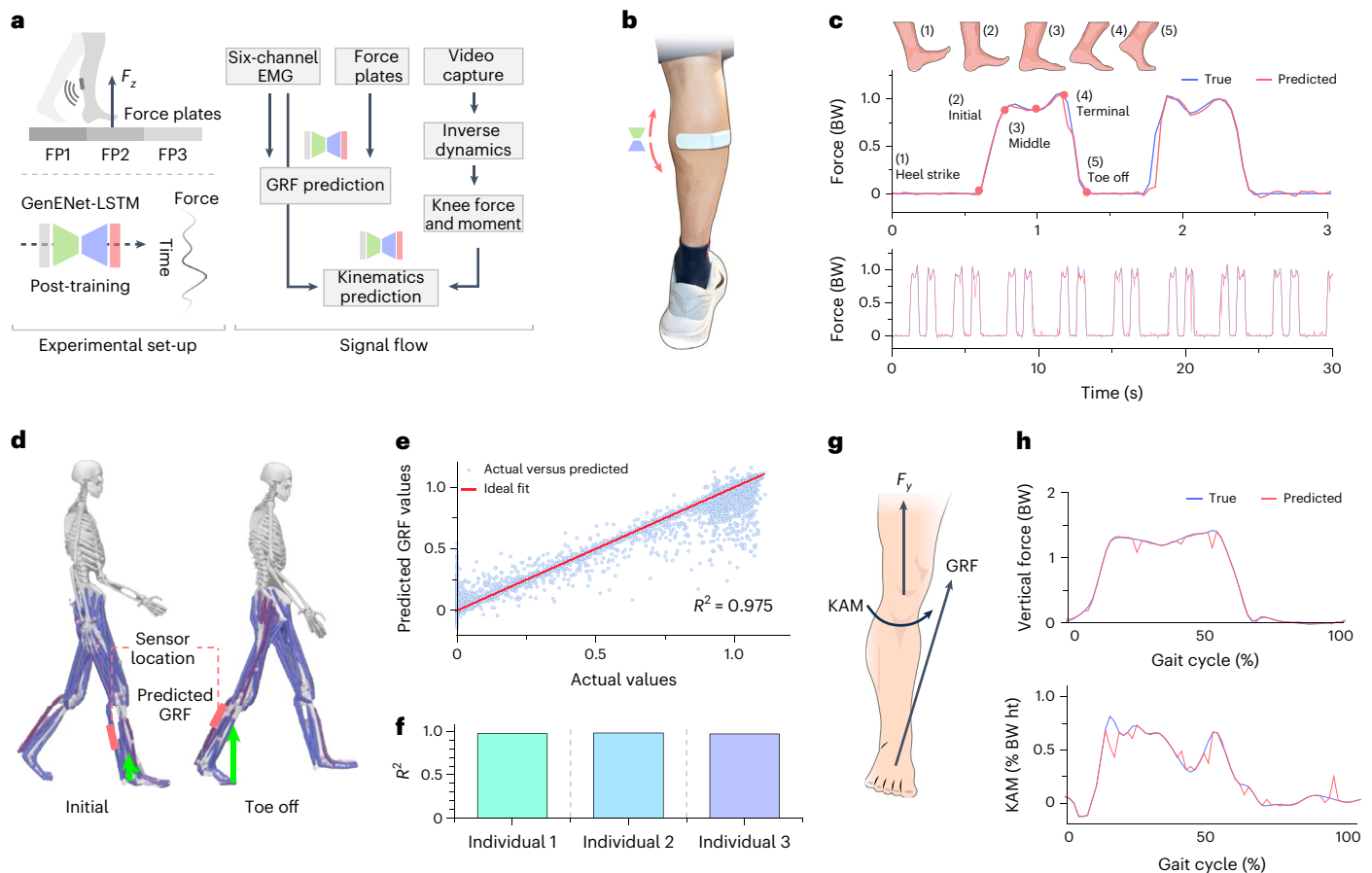


Fig. 5 | Prediction of gait dynamics using the GenENet device. **a**, Experimental set-up involving walking across three force plates (FPs) with simultaneous video capture. F_z denotes the vertical ground reaction force. The post-training network is used to predict the GRF, while vertical knee force and moment are calculated through inverse dynamics based on the video data, which are incorporated into the kinetic post-training dataset. **b**, Schematic of the six-channel EMG device attached to the calf. **c**, GRF prediction during the gait cycle, showing five distinct phases where predicted values closely match the true values obtained from the

video data. **d**, Snapshots of the real-time prediction of GRF on musculoskeletal model. The green arrow indicates the GRF and the device is attached to the right leg as shown in the red region. **e**, R^2 coefficient of 0.975 for GRF prediction. **f**, Adaptation to different individuals, showing a consistent R^2 coefficient across them. **g**, Illustration of GRF and KAM vector directions. **h**, Predicted y-axis knee joint force and KAM over specific time intervals. Muscle contributions were not included in the inverse dynamics calculation of joint forces.

(Supplementary Figs. 8 and 9). In contrast, the non-pre-trained model, which shares the same architecture but is trained from scratch without access to the full 32-channel dataset, remains below 10% accuracy. The model trained directly on the full 32-channel input achieves the highest validation accuracy because it benefits from richer sensor information. In contrast, the model trained directly on the six-channel input without any pre-training shows notably lower accuracy. This clearly demonstrates that the six-channel input alone lacks sufficient information to fully support complex class prediction (Supplementary Fig. 10).

The above-described capability allowed the model to transfer pre-trained knowledge to newly attached positions, as demonstrated by placing sensors in 7 different locations within the region where the original 32-channel sensor was positioned to capture the dataset, as depicted in Fig. 4h. The real-time application of the six-channel set-up for sign language translation is illustrated in Fig. 4i. The graph shows live EMG signals from the six-channel set-up along with the predicted alphabet characters in numeric labels. During the performance of the sign language phrase ‘Hello World’, the post-trained GenENet translated the six-channel information into corresponding alphabet values (Supplementary Video 3). The red lines indicate numeric values from 0 to 25, representing the alphabet from A to Z.

An attribution map is shown during prediction in Fig. 4j, where the model interprets which muscle features contribute to specific sign language hand postures. Each row in the signal corresponds to one

of the six sensor channels. Blue regions closer to 1 indicate positive contributions to the prediction and red regions closer to -1 negatively impact the prediction. Attribution maps for all 26 classes are provided in Supplementary Fig. 11. A higher attribution value in a specific channel may suggest a stronger correlation between the corresponding muscle group in the electrode’s position at each learning time frame.

Downstream adaptation of GenENet for gait dynamics prediction

To extend our system’s applicability to different body locations, we attached the device to the calf to monitor kinetic information during the gait cycles. Continuous monitoring of gait kinetics provides valuable insights into potential musculoskeletal disorders, aiding in the identification of risk factors for falls, the need for rehabilitation and optimization of athletic performance^{35–37}.

Previous work used video capture and calculation through inverse dynamics to determine knee moments and corresponding forces³³. However, this requires a specialized lab set-up, making it difficult to apply in daily life. When EMG arrays were used, 32 electrodes could be used for gait dataset capturing while a 6-electrode array was subsequently used for gait force, knee force and moment prediction. For gait kinetic prediction, a 6-electrode array was minimally required to achieve average R^2 value of 0.972 and relative root mean square error (RMSE) of 6.09 % (Supplementary Table 2).

For the initial pre-training dataset in gait kinetic prediction, we used the 32-channel EMG signal data collected during a normal gait cycle (15 min, ~1,700 gait cycles; experimental set-up shown in Supplementary Fig. 12). In the subsequent post-training phase, using the smaller device, we recorded six-channel EMG signals in conjunction with ground reaction forces (GRFs) captured through three force plates during a normal gait cycle, as illustrated in Fig. 5a. The measured GRF data were fed into the model, where we used MSELoss for continuous force prediction. Simultaneously, gait movements were video-captured using OpenCap³⁸, and knee moments and forces were calculated through inverse dynamics (Methods). The attached 6-channel device, shown in Fig. 5b, was then positioned within the 32-channel measurement area as the boundary.

We identified three key gait states—heel strike, mid-stance and toe off—using the GRF data, as shown in Fig. 5c. The GRF signals including these states were then input into the pre-trained GenENet encoder with the 32-channel stretchable EMG array for further post-training. As shown in Fig. 5d, our model successfully predicted continuous gait forces (expressed in units of body weight (BW)) throughout the gait cycle. To visualize the predicted GRF across the gait cycle, we mapped the OpenCap data to an OpenSim musculoskeletal model and associated it with the predicted GRF values over two representative gait cycles, as shown in Supplementary Video 2. The video provides two perspectives: the top panel shows a diagonal view and the bottom panel presents a side view of the motion. A green arrow represents the relative strength of the predicted GRF, with the device attached to the right leg highlighted in the red region. As depicted in Fig. 5e, the model achieved an R^2 coefficient of 0.975 with relative RMSE of 6.21%, and the model was able to transfer learning to different individuals, as shown for three individuals in Fig. 5f and Supplementary Figs. 13 and 14. For the gait regression task, we incorporated decoder and combined it with an additional CNN to help reconstruct and refine feature representations, capturing richer local patterns and temporal dependencies. The incorporation of the CNN-LSTM block enabled more precise reproduction of GRF signals, as shown in Supplementary Fig. 15.

To evaluate robustness under sensor placement variation, we conducted experiments involving 7 different locations on the calf, 4 placed horizontally and 3 placed vertically, each spaced 30 mm apart. As shown in Supplementary Fig. 16, the ground-truth model showed relatively high average loss across all locations. However, after applying a lightweight fine-tuning step (30 epochs) for each of the 7 locations, the overall MSE for the gait force prediction task was substantially reduced, even after just a few epochs of adaptation. These results indicate that while some performance degradation occurs due to sensor displacement, it can be effectively mitigated with minimal post-training. Inter-session and inter-individual variability are also assessed in Supplementary Fig. 17.

Understanding gait forces is essential for analysing human locomotion and assessing biomechanical health, aiding in the identification of risk factors such as fall risks, rehabilitation needs and injury prevention^{35,36,39–41}. The correlated EMG signals provide further predictive power for assessing these risk factors. The use of compact wearable EMG arrays enables such analysis in daily life without the requirement of specialized video set-up labs as currently used.

Through inverse dynamic simulations powered by OpenCap³⁸ (Methods), we extracted knee joint forces and the knee adduction moment (KAM) and correlated them with the corresponding EMG signals. As illustrated in Fig. 5g, KAM represents the moment acting on the joint in the frontal plane, causing medial rotation of the tibia on the femur⁴². Higher KAM is often associated with the development or progression of medial knee osteoarthritis⁴³. As shown in Fig. 5h, the GenENet combined with CNN-LSTM successfully predicted the progression of GRF and KAM during the gait cycle, particularly within the 4.5–5.5 s range depicted in Fig. 5h. The resulting peak KAM values fell within 0.5–1% BW ht (body weight × height), which is lower than

previously reported values (1–3% BW ht)^{44,45}, without indicating deviation from expected physiological trends. This approach highlights notable advancements in KAM prediction by relying solely on EMG signals, unlike previous methods that require a combination of EMG and inertial measurement unit devices. In addition, while other systems often rely on distributed EMG arrays or with extensive sensor placements^{46–49}, our compact six-channel EMG array achieves comparable predictive performance, providing a clear advantage in ease of use and practicality (Supplementary Table 2). Moreover, the use of this compact, wearable EMG array shows a potential reduction in power consumption by approximately 71% compared with 32-channel systems (Methods). This power saving underscores the feasibility of deploying such systems in portable, long-term monitoring set-ups without sacrificing predictive accuracy and without the need for cumbersome large electrode arrays.

Conclusion

We have developed an approach that enables the use of compact and low-power-consumption few-electrode arrays to predict signals equivalent to those from a much larger-area and high-electrode count system. We combine a generative representation-learning algorithm with a wearable device to extrapolate limited sensor information to reconstruct muscle EMG activities in unseen regions to expand information collected, demonstrating its practical application in alphabet recognition and gait dynamics predictions. Precise gestures and gait dynamics predictions previously required increasing footprint and body coverage, such as large EMG grids with 64 to 256 electrodes, or relied on fewer electrodes combined with external sensors or distributed attachments. This approach reduced sensor count, footprint and power consumption for data transmission while maintaining performance.

In the pre-training phase, the generative representation learning network was trained on a high-quality, multichannel dataset obtained from low-impedance polymer electrodes and higher-SNR stretchable interconnections. This network was then integrated with an LSTM layer to enable sequential processing of kinetic inputs.

The successful demonstration of a compact six-channel EMG device for complex task predictions, such as individual alphabet recognition and gait dynamics predictions, highlights the potential for practical applications. Going forward, the resolution of the training array can be further enhanced to produce even more detailed information, while adapting the GenENet approach can ease the complexity of fabrication and lower the production cost of the wearable versions while maintaining performance. We anticipate that this platform could be extended to accommodate other types of input that typically require high-density sensory arrays while there are correlations between the signals, including strain, temperature, other electrophysiological sensing, such as electrocardiogram and electroencephalogram, photodiodes, ultrasonic sensors, and even chemical sensors.

While our system shows robust performance, it could be further improved by implementing on-device processing of signals, enabling reduced communication bandwidth and lower-radio-power requirements. In addition, the current model is limited to data obtained from a 32-channel area. Expanding the dataset to include a larger sensor coverage area would enhance the capabilities of the system, allowing for greater flexibility in sensor placement across more arbitrary regions of the body.

Moreover, integrating a sensor fusion strategy that combines an inertial measurement unit with EMG data alongside electrophysiological signals may enrich the dataset and improve the network's training²⁶. In addition, proper calibration is crucial when adapting the system for multiple users, as individual variations in muscle physiology and skin impedance may require personalized adjustments to ensure optimal performance. Special attention should also be given to abnormal conditions, such as users with partial muscle dysfunction or neuromuscular disorders, where signal patterns can deviate from typical profiles.

In such cases, calibration techniques such as personalized adaptive baseline subtraction and filtering may be necessary to address signal inconsistencies and enhance the system's reliability.

This advancement paves the way for a wide range of applications with improved form factors and reduced data collection and transmission energy consumption while preserving the quality of multi-array signals. Potential applications include health monitoring (for example, blood pressure, respiration rate, pulse), prosthetics (for example, prosthetic limbs, gait analysis), sports (for example, posture monitoring) and human–machine interfaces (for example, gesture recognition, facial movement recognition, virtual reality).

Methods

Materials

The following materials were all obtained from Sigma Aldrich: dextran (product 09184), NBR (product 180912), cyclohexane (product 179191), cyclohexanone (product 398241), PEDOT:PSS (product 739332), SBS (product 182877), phenylbis(2,4,6-trimethylbenzoyl) phosphine oxide (BAPO; product 511447), pentaerythritol tetrakis(3-mercaptopropionate) (PETMP; product 381462), EGaIn (product 495425), PBS (product 806552), poly(pyromellitic dianhydride-co-4,4'-oxydianiline), amic acid solution (product 575828) and AAM (product 800830), and *N,N'*-methylenebisacrylamide (MBAA; product 101546). Aqueous hydrogen peroxide (30%) was obtained from Fisher Scientific. Ascorbic acid was obtained from TCI. PDMS (Sylgard 184) was purchased from Dow.

Fabrication of the 32-channel dataset-generation device

Solution and substrate preparation. A 60 mg ml⁻¹ solution of NBR was prepared in cyclohexanone. This solution was left to dissolve overnight on a hotplate set to 150 °C. The fabrication process began with oxygen plasma etching (March Instruments PX-250 Plasma Asher) of a silicon wafer at 150 W and an oxygen flow rate of 2 s.c.c.m. for 100 s. Following this, a 10 wt% dextran aqueous solution was spin-coated at 800 rpm for 60 s onto the wafer and baked at 100 °C for 5 min to form a sacrificial layer. To create the flexible substrate, PDMS (1:10 weight ratio) was spin-coated onto the wafer at 800 rpm for 45 s, followed by annealing at 100 °C for 30 min. The NBR solution (60 mg ml⁻¹ in cyclohexanone solution), mixed with BAPO and PETMP (4 wt% of base polymer) for photo-crosslinking, was then spin-coated onto the PDMS substrate (oxygen plasma etched 150 W for 50 s) at 800 rpm for 60 s. Then this was ultraviolet-cured for 20 min in a nitrogen environment to complete the substrate formation with post-baking of 120 °C for 15 min.

Electrode patterning. Electrode patterning was initiated by applying photoresist (AZ 1512) at 2,000 rpm for 60 s followed by baking at 90 °C for 1 min. The desired pattern was created using masked patterning exposed in ultraviolet light for 10 s in air (Open Cure 365-nm LED UV; Supplementary Fig. 18) and developed in MF219 developer. Following patterning, chromium (3 nm) and gold (40 nm) layers were deposited via thermal evaporation. EGaIn was then applied on top of the patterned electrodes. The lift-off process was carried out by immersing the patterned device in acetone for approximately 3 h to remove the photoresist and leave the desired metal patterns.

Encapsulation preparation. For encapsulation, an 80 mg ml⁻¹ solution of SBS with BAPO and PETMP (4 wt% of base polymer) in toluene was prepared and spin-coated onto the device at 800 rpm for 60 s. A photomask was used to define the encapsulation pattern, curing with 365 nm ultraviolet lamp in 3 s, and the unexposed areas were dissolved by washing with cyclohexane.

PEDOT electrode patterning. To prepare the PEDOT electrodes, 5 ml PEDOT:PSS was combined with 1.5 g AAM and 5 mg MBAA. To initiate

gel formation, 250 µl of 20% (wt/vol in water) ascorbic acid and 250 µl of 30% (v/v in water) hydrogen peroxide were added to the mixture. The solution was then dispensed onto an Ecoflex mould (4.5 mm in diameter and depth), where it formed a gel and created a defined pattern on top of the previously prepared electrodes. Then, after attaching the electrodes to the EGaIn electrodes, the Ecoflex mould was detached, where the final device was created as shown in Supplementary Fig. 19. A six-channel array was fabricated using the same method. A low-impedance, highly conductive electrode paste (Elefix V Electrode Paste, Nihon Kohden) was applied to the six-channel electrode array to enhance the signal quality. The device was then securely attached to the skin using a biomedical adhesive (Skinister). To ensure long-term stability and fully prevent oxidation, the electrode surface facing the PEDOT:PSS layer can be selectively exposed as gold, as shown in Supplementary Figs. 20 and 21. This approach requires additional fabrication steps, as illustrated in Supplementary Fig. 19. Stages a to c involve the patterning of Au for both electrode and wiring layouts. Supplementary Fig. 20g represents a faster prototyping method in which EGaIn is applied over the entire gold surface, enabling quicker data collection. In contrast, Supplementary Fig. 20d shows an additional photoresist patterning step that masks the electrode area so that EGaIn is applied only to the wiring regions.

A miniaturized wireless EMG device for dataset generation

The device integrates a flexible printed circuit board featuring an analogue-to-digital converter sensing element, a Bluetooth low energy module, a lithium polymer battery, a multiplexer, an amplifier and a 32-channel EMG module. The 32-channel EMG system is connected to the wireless module via an anisotropic conductive film, enabling reliable signal transfer. Analogue EMG signals are amplified and multiplexed through a fully integrated electrophysiology chip (Intan Technologies) before being processed by the analogue-to-digital converter. These signals are digitized and transmitted at a 250 Hz data rate via Bluetooth to the receiver. The compact wireless module, through conformal attachment to the skin, ensures accurate motion detection without compromising user comfort. The system is programmed using a system-on-chip (CC2650, Texas Instruments) in Code Composer Studio, which converts the packets of analogue sensor data into UART format for transmission.

Mould fabrication through 3D printing

To visualize and evaluate the conformability of the device, we have 3D-printed (Ender 3 V2) a corrugated surface. A detached device is placed on top of the surface as shown in Supplementary Fig. 22. Also, as shown in Supplementary Fig. 23, a 6-channel watch-type mould is created through 3D printing and poured and cured the encapsulation polymer (Elkem RTV 4420; part A and part B, mixed with 5% of silicone opaque dye).

Visualization using *t*-SNE

To visualize and evaluate the results of the trained dataset, we applied *t*-distributed stochastic neighbour embedding (*t*-SNE). The *t*-SNE visualization reveals clusters in the projected two-dimensional space, offering insights into sample relationships and the potential separability of classes in high-dimensional space. *t*-SNE was performed on both ASL and GRF predictions, as shown in Supplementary Figs. 9 and 14.

Dataset collection for pre-training

Dataset collection was driven by the custom-made wireless device. The dataset was collected at Stanford Human Performance Lab. Wrist movements were collected with randomly performed different finger gestures for 20 min from a single individual. Gait motions are collected while attached to the calf and with normal walking cycles with 15 min from a single individual. This protocol is approved through Stanford IRB-54795.

Inverse kinematic simulation

Post-processing of human movement kinematics collected using OpenCap were done through OpenCap processing (<https://github.com/stanfordnmb/opencap-processing>). Multiple sessions of OpenCap measurements were processed to estimate kinetic information of joint force and moment.

Details on the transformer encoding and decoding block

The overall architecture is shown in Supplementary Fig. 24. The input signal is a 32×32 array of 32-channel EMG signals with 32 time windows (-0.12 s), for which the tensor size of input will be $x \in \mathbb{R}^{H \times W \times C}$, where H , W and C represent height (32), width (32) and channels (1), respectively. The RMS amplitude of EMG signal is shown in a one-channel array. This tensor will be augmented through resizing and random cropping into a 48×48 array, followed by random horizontal flipping for augmentation.

The image is then divided into patches of size $P \times P$, resulting in $N = H \times W / P^2$ patches, each with $P \times P \times C$ values, which defines the patch area, where the patch size P is 6.

Each patch is then embedded using a linear projection:

$$x_{\text{patch}} = f_{\text{proj}}(x_{\text{patch}}) \in \mathbb{R}^{N \times D}$$

where f_{proj} denotes a linear projection layer that maps each flattened patch to a D -dimensional vector (128).

Before encoding the signals into a motion feature space with the attention mechanism, we add positional embedding to the embedded vectors so that the model can understand the relative position of input sequences while encoding them in parallel. Positional embedding is one of key features of transformer architecture that allows the model to avoid iterative computation for each time frame.

For N patches, the positional embeddings are represented as:

$$\text{pos}_i \in \mathbb{R}^D, i \in [1, N]$$

The positional embeddings are then added to the patch embeddings to be conveyed to the transformer:

$$x_{\text{patch_pos}} = x_{\text{patch}} + \text{pos}$$

A subset of patches is masked before being fed into the transformer, where the randomly selected indices of patches will be masked that will be fed into the transformer decoder, where the rest of unmasked patches are passed into the transformer encoder. The transformers' self-attention plays a crucial role in capturing long-range dependencies and global context across the EMG signals.

The encoder processes several layers of multi-head attention, layer normalization and feed-forward networks.

$$\hat{x}_{\text{patch}} = \text{LayerNorm}(x_{\text{patch}})$$

$$\text{query}_{ij} = W_q \hat{x}_{i_{\text{patch}}}, \text{key}_{tj} = W_k \hat{x}_{t_{\text{patch}}}, \text{value}_{tj} = W_v \hat{x}_{t_{\text{patch}}}$$

$$\text{attn}(i, j) = \text{softmax}\left(\frac{\text{query}_i}{\sqrt{32}} \times \text{key}_{tj}^T\right)$$

where i and j index the query and key tokens, respectively; W_q , W_k and W_v indicate the query, key and value projection matrices; t denotes the time step. Then the residual connection ensures the original input to be preserved:

$$x_{\text{enc}} = x_{\text{patch_pos}} + \text{attn}(i, j)$$

After applying residual connection, the output passes through a feed-forward layer consisting of two linear layers, a Gaussian error

linear unit (GeLU) activation and dropout. The output of the encoder then further processed through the input of the decoding layer.

The decoder processes the same architecture with the encoder, but including a linear layer in the end to reconstruct the original EMG signal. The output from the encoder is first projected into the decoder space:

$$x_{\text{dec_p}} = W_{\text{in}}(x_{\text{enc}})$$

where

$$W_{\text{in}} \in \mathbb{R}^{D_{\text{enc}} \times D_{\text{dec}}}$$

The projected input is then added with mask tokens, where mask tokens replace and represent the originally masked patches. These tokens are learnable parameters that initialized randomly from normal distribution:

$$x_{\text{token}} \sim \mathcal{N}(0, 1) \in \mathbb{R}^{P \times P \times C}$$

Positional embeddings are then applied to both of $x_{\text{dec_p}}$ and x_{token} , where the final decoder input will be:

$$x_{\text{dec_in}} = x_{\text{dec_p_pos}} + x_{\text{token_pos}}$$

As the encoder, the decoder processes layers of multi-head attention, layer normalization and feed-forward networks. The output of the decoder is flattened and passed through a linear layer to reconstruct the original EMG signal:

$$x_{\text{flat}} = \text{Flatten}(x_{\text{dec}}), \hat{x}_{\text{recon}} = W_{\text{out}} \times x_{\text{flat}}$$

Final reshaping of \hat{x}_{recon} ensures that the output matches the original dimension of multi-array EMG signal.

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N_{\text{mask}}} \sum_{i \in \text{mask}} (p_{\text{orig}}^i - p_{\text{recon}}^i)^2$$

EMG data processing with tensor generation

The raw EMG signals are filtered using a fourth-order Butterworth bandpass filter with cut-off frequencies between 1 Hz and 100 Hz. Filtering reduces noise and retains relevant frequency components. Sliding window is used to segment the EMG signal into overlapping frames. Each frame consists of 32 time steps (128 ms), corresponding to 1 sliding window. The sliding window advances with a step size of 10 (40 ms), ensuring sequential windows overlap which helps maintain temporal continuity in the data.

Within each window, the RMS of the signal is computed. The RMS features are then converted into a three-dimensional tensor via colourmap conversion. Each tensor is labelled with the corresponding gesture classes or force signal.

Gradient-based attribution visualization

To investigate the model's interpretability across 26 classes, we utilize Gradient SHAP from Captum to generate attribution maps for representative examples from each class. This process highlights which portions of the input sequence contribute most strongly to the predictions.

Input tensors and their corresponding labels are iteratively retrieved from the dataset. For each alphabet class $c \in \{1, 2, \dots, 26\}$, a single representative sample x_c is selected from the evaluation set. Given a batch of input sequences $X = \{x_1, x_2, \dots, x_B\}$ and corresponding labels $Y = \{y_1, y_2, \dots, y_B\}$, the representative sequence is stored.

Gradient SHAP estimates feature importance by comparing the input example against a baseline tensor. The attribution values highlight the features that contribute most to the model's decision.

The attributions are arranged in a 26×2 grid, where each row corresponds to a specific class. In Fig. 4j, the left panel shows the original input example, and the right panel shows the corresponding attribution map.

Details on the backend CNN-LSTM block for gesture and kinematic prediction

The CNN-LSTM network processes a sequence of EMG signals to predict target values (for example, force and finger classes). This network combines the generated information from the GenENet, where the model extracts features from individual frames and captures temporal dependencies with an LSTM.

Given a sequence $x \in \mathbb{R}^{B \times T \times C \times H \times W}$, where B is the batch size, T is the sequence length, C is the number of channels, and $H \times W$ is the signal resolution, each signal frame is processed individually. Each frame is passed through the transformer-based GenENet, resulting in tensor from the decoder z_i .

The output is passed through a CNN module to extract spatial features, which includes 3×3 convolution, Rectified linear unit (ReLU) activation and 2×2 max pooling. The output is flattened and further reduced by a fully connected compression layer. $z_{\text{comp}} \in \mathbb{R}^{B \times S}$, and S is the dimension of the compressed feature.

The sequence of compressed features is passed through an LSTM network, where the last hidden state is fed through a fully connected layer to produce the final output. Depending on the task, MSE loss or cross entropy loss is applied for force or class prediction, respectively. The final model achieves an average inference latency of 46.95 ms, with 7.7 million multiply-accumulate operations and 0.428 million parameters.

Comparative analysis of theoretical time complexity

To analyse and compare the theoretical time complexities of two signal processing techniques, RMS and spectrogram, the time complexities were calculated relative to the signal size. This offers insights into the computational demands of each method. The RMS computation requires a single pass through the signal, making its time complexity linear in n , the signal size. The time complexity for a spectrogram depends on the signal size n , the window size w , and the frequency resolution achieved through Fourier transform within each window.

A synthetic EMG signal is generated to simulate a realistic input for each processing technique. Two methods, RMS and spectrogram, are evaluated across 10,000 iterations to assess the average processing time. The average processing times across all iterations are then computed as $\bar{t}_{\text{RMS}} = 1/\text{iter} \sum_{n=1}^{\text{iter}} t_{\text{RMS}}^{(n)}$ and $\bar{t}_{\text{spectrogram}} = 1/\text{iter} \sum_{n=1}^{\text{iter}} t_{\text{spectrogram}}^{(n)}$.

The average processing times for each method are visualized in a bar plot for direct comparison as shown in Supplementary Fig. 25, showing RMS as an efficient approach for quick signal characterization.

Power-saving estimation

The potential power savings of using a 6-channel device compared with a 32-channel device can be calculated based on their transmission characteristics:

$$P_{\text{total}} \propto P_{\text{tx}} \times T_{\text{tx}}$$

where P_{tx} is the power during transmission and T_{tx} is the transmission duration. The transmission payload for 32-channel and 6-channel devices can be determined, with each channel contributing 2 bytes to the payload:

32-channel payload: $2 \times 32 + 2 = 66$ bytes

6-channel payload: $2 \times 6 + 2 = 14$ bytes

Here, 2 bytes are used for timestamp data. The transmission duration can be expressed as:

$$T_{\text{tx}} = \text{Payload}/\text{Data rate}$$

As $T_{\text{tx}} \propto \text{Payload}$, the power savings can be estimated as:

$$1 - T_{\text{tx},6}/T_{\text{tx},32} \approx 0.71$$

Reporting summary

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

Data availability

The collected finger datasets for various daily tasks performed in this study are available via GitHub at <https://github.com/nature-sensors/GenENet>. Further data that supports the plots within this paper and other findings of this study are available from the corresponding author upon request.

Code availability

The code used in this study is available via GitHub at <https://github.com/nature-sensors/GenENet>.

References

1. Tang, L., Shang, J. & Jiang, X. Multilayered electronic transfer tattoo that can enable the crease amplification effect. *Sci. Adv.* **7**, eabe3778 (2021).
2. Sundaram, S. et al. Learning the signatures of the human grasp using a scalable tactile glove. *Nature* **569**, 698–702 (2019).
3. Caesarendra, W., Tjahjowidodo, T., Nico, Y., Wahyudati, S. & Nurhasanah, L. EMG finger movement classification based on ANFIS. *J. Phys. Conf. Ser.* **1007**, 012005 (2018).
4. Tacca, N. et al. Wearable high-density EMG sleeve for complex hand gesture classification and continuous joint angle estimation. *Sci. Rep.* **14**, 18564 (2024).
5. Moin, A. et al. A wearable biosensing system with in-sensor adaptive machine learning for hand gesture recognition. *Nat. Electron.* **4**, 54–63 (2021).
6. Sun, Z., Zhu, M., Shan, X. & Lee, C. Augmented tactile-perception and haptic-feedback rings as human-machine interfaces aiming for immersive interactions. *Nat. Commun.* **13**, 5224 (2022).
7. Gao, X. et al. A wearable echomyography system based on a single transducer. *Nat. Electron.* **7**, 1035–1046 (2024).
8. Tashakori, A. et al. Capturing complex hand movements and object interactions using machine learning-powered stretchable smart textile gloves. *Nat. Mach. Intell.* **6**, 106–118 (2024).
9. Zhou, Z. et al. Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays. *Nat. Electron.* **3**, 571–578 (2020).
10. Kim, K. K. et al. A deep-learned skin sensor decoding the epicentral human motions. *Nat. Commun.* **11**, 2149 (2020).
11. Kim, K. K. et al. A substrate-less nanomesh receptor with meta-learning for rapid hand task recognition. *Nat. Electron.* **6**, 64–75 (2023).
12. Hu, H. et al. Stretchable ultrasonic arrays for the three-dimensional mapping of the modulus of deep tissue. *Nat. Biomed. Eng.* **7**, 1321–1334 (2023).
13. Kweon, O. Y., Lee, S. J. & Oh, J. H. Wearable high-performance pressure sensors based on three-dimensional electrospun conductive nanofibers. *npj Asia Mater.* **10**, 540–551 (2018).
14. Vinoth, R., Nakagawa, T., Mathiyarasu, J. & Mohan, A. V. Fully printed wearable microfluidic devices for high-throughput sweat sampling and multiplexed electrochemical analysis. *ACS Sens.* **6**, 1174–1186 (2021).
15. Oh, Y. S. et al. Battery-free, wireless soft sensors for continuous multi-site measurements of pressure and temperature from patients at risk for pressure injuries. *Nat. Commun.* **12**, 5008 (2021).

16. Han, S. et al. Battery-free, wireless sensors for full-body pressure and temperature mapping. *Sci. Transl. Med.* **10**, eaan4950 (2018).
17. Sîmpetru, R. C. et al. Learning a hand model from dynamic movements using high-density EMG and convolutional neural networks. *IEEE Trans. Biomed. Eng.* **71**, 12 (2024).
18. Shin, J. H. et al. Wearable EEG electronics for a Brain–AI Closed-Loop System to enhance autonomous machine decision-making. *npj Flex. Electron.* **6**, 32 (2022).
19. Mahmood, M. et al. Fully portable and wireless universal brain–machine interfaces enabled by flexible scalp electronics and deep learning algorithm. *Nat. Mach. Intell.* **1**, 412–422 (2019).
20. Sunwoo, S.-H. et al. Ventricular tachyarrhythmia treatment and prevention by subthreshold stimulation with stretchable epicardial multichannel electrode array. *Sci. Adv.* **9**, eadf6856 (2023).
21. Kim, S. et al. Three-dimensional electrodes of liquid metals for long-term, wireless cardiac analysis and modulation. *ACS Nano* **18**, 24364–24378 (2024).
22. Kim, D.-H. et al. Materials for multifunctional balloon catheters with capabilities in cardiac electrophysiological mapping and ablation therapy. *Nat. Mater.* **10**, 316–323 (2011).
23. Choi, S. et al. Highly conductive, stretchable and biocompatible Ag–Au core–sheath nanowire composite for wearable and implantable bioelectronics. *Nat. Nanotechnol.* **13**, 1048–1056 (2018).
24. Park, J. et al. Electromechanical cardioplasty using a wrapped elasto-conductive epicardial mesh. *Sci. Transl. Med.* **8**, 344ra386 (2016).
25. Pincheira, P. A. et al. Regional changes in muscle activity do not underlie the repeated bout effect in the human gastrocnemius muscle. *Scand. J. Med. Sci. Sports* **31**, 799–812 (2021).
26. Dimitrov, H., Bull, A. M. & Farina, D. High-density EMG, IMU, kinetic, and kinematic open-source data for comprehensive locomotion activities. *Sci. Data* **10**, 789 (2023).
27. Montazerin, M. et al. Transformer-based hand gesture recognition from instantaneous to fused neural decomposition of high-density EMG signals. *Sci. Rep.* **13**, 11000 (2023).
28. He, K. et al. Masked autoencoders are scalable vision learners. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 16000–16009 (2022).
29. Li, T. et al. MAGE: MASKed Generative Encoder to unify representation learning and image synthesis. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 2142–2152 (2023).
30. Van Hoorick, B. Image outpainting and harmonization using generative adversarial networks. Preprint at <https://arxiv.org/abs/1912.10960> (2019).
31. Gong, Y., Chung, Y.-A. & Glass, J. AST: audio spectrogram transformer. In *Interspeech 2021* https://www.isca-archive.org/interspeech_2021/gong21b_interspeech.pdf (2021).
32. Radford, A. et al. Robust speech recognition via large-scale weak supervision. In *Proc. 40th International Conference on Machine Learning* 28492–28518 (PMLR, 2023).
33. Wang, C. et al. fairseq S2T: Fast speech-to-text modeling with fairseq. In *Proc. 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and 10th International Joint Conference on Natural Language Processing: System Demonstrations* <https://doi.org/10.18653/v1/2020.aacl-demo.6> (2020).
34. Guo, P. et al. Recent developments on espnet toolkit boosted by conformer. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 5874–5878 (2020).
35. Carvajal-Castaño, H., Lemos-Duque, J. & Orozco-Arroyave, J. Effective detection of abnormal gait patterns in Parkinson's disease patients using kinematics, nonlinear, and stability gait features. *Hum. Mov. Sci.* **81**, 102891 (2022).
36. Grabiner, M. D. & Kaufman, K. Developing and establishing biomechanical variables as risk biomarkers for preventable gait-related falls and assessment of intervention effectiveness. *Front. Sports Act. Living* **3**, 722363 (2021).
37. Haugen, T. et al. On the importance of “front-side mechanics” in athletics sprinting. *Int. J. Sports Physiol. Perform.* **13**, 420–427 (2018).
38. Uhlrich, S. D. et al. OpenCap: human movement dynamics from smartphone videos. *PLoS Comput. Biol.* **19**, e1011462 (2023).
39. Ferber, R., Osis, S. T., Hicks, J. L. & Delp, S. L. Gait biomechanics in the era of data science. *J. Biomech.* **49**, 3759–3761 (2016).
40. Meurisse, G. M., Dierick, F., Schepens, B. & Bastien, G. J. Determination of the vertical ground reaction forces acting upon individual limbs during healthy and clinical gait. *Gait Posture* **43**, 245–250 (2016).
41. Pogorelec, B., Bosnić, Z. & Gams, M. Automatic recognition of gait-related health problems in the elderly using machine learning. *Multimedia Tools Appl.* **58**, 333–354 (2012).
42. Maly, M. R. et al. Knee adduction moment relates to medial femoral and tibial cartilage morphology in clinical knee osteoarthritis. *J. Biomech.* **48**, 3495–3501 (2015).
43. Froughi, N., Smith, R. & Vanwanseele, B. The association of external knee adduction moment with biomechanical variables in osteoarthritis: a systematic review. *Knee* **16**, 303–309 (2009).
44. Shull, P. B. et al. Six-week gait retraining program reduces knee adduction moment, reduces pain, and improves function for individuals with medial compartment knee osteoarthritis. *J. Orthop. Res.* **31**, 1020–1025 (2013).
45. Miller, R. H., Krupenevich, R. L., Pruziner, A. L., Wolf, E. J. & Schnall, B. L. Medial knee joint contact force in the intact limb during walking in recently ambulatory service members with unilateral limb loss: a cross-sectional study. *PeerJ* **5**, e2960 (2017).
46. Sakamoto, S.-I., Hutabarat, Y., Owaki, D. & Hayashibe, M. Ground reaction force and moment estimation through EMG sensing using long short-term memory network during posture coordination. *Cyborg Bionic Syst.* **4**, 0016 (2023).
47. Mengarelli, A. et al. Mapping lower limb EMG activity to ground reaction force in free walking condition. In *International Symposium on Medical Measurements and Applications (MeMeA)* 1–6 (2024).
48. Stetter, B. J., Krafft, F. C., Ringhof, S., Stein, T. & Sell, S. A machine learning and wearable sensor based approach to estimate external knee flexion and adduction moments during various locomotion tasks. *Front. Bioeng. Biotechnol.* **8**, 9 (2020).
49. Tan, T., Wang, D., Shull, P. B. & Halilaj, E. IMU and smartphone camera fusion for knee adduction and knee flexion moment estimation during walking. *IEEE Trans. Ind. Inf.* **19**, 1445–1455 (2022).

Acknowledgements

This work was partially supported by the Army Research Office Bionic Electronics Program (grant no. W911NF-23-1-0282). Part of this work was performed at the Stanford Nano Shared Facilities (SNSF) and Stanford Nanofabrication Facility (SNF). Part of this work was performed at Stanford Human Performance Lab. K.K.K. acknowledges postdoctoral fellowship support from Wu Tsai Human Performance Alliance. Z.B. is an investigator of the CZ Biohub-San Francisco and an Arc Institute innovation investigator. Z.B. acknowledges support from the Stanford Wearable Electronics Initiative (eWEAR) seed funding, the Tianqiao and Chrissy Chen Ideation and Prototyping Lab. We thank Y. M. Liu and S. F. Fung for their generous support of the Bao group's research at Stanford University.

Author contributions

K.K.K., S.L.D. and Z.B. conceived of the project and designed the experiments. K.K.K. designed and fabricated the devices,

and performed the mechanical and electrical measurements. K.K.K. developed and designed the training algorithms. T.J.Z. fabricated the devices and performed electrical measurements. S.S. and S.L.D. guided the OpenCap measurement and inverse dynamic simulations. Y.L., H.P. and Y.J. helped in the preparation of material synthesis and fabrication. D.Z., M.K. and Y.N. helped with microfabrication of measurement devices. Y.J. designed the low-impedance polymer electrode. K.K.K., T.J.Z., S.S., S.L.D. and Z.B. prepared the paper. Z.B. and S.L.D. directed the project.

Competing interests

Stanford University has filed a patent application related to this technology. The patent application number is PCT/US2024/056006. K.K.K. and Z.B. are listed as inventors in this application. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s44460-025-00002-2>.

Correspondence and requests for materials should be addressed to Zhenan Bao.

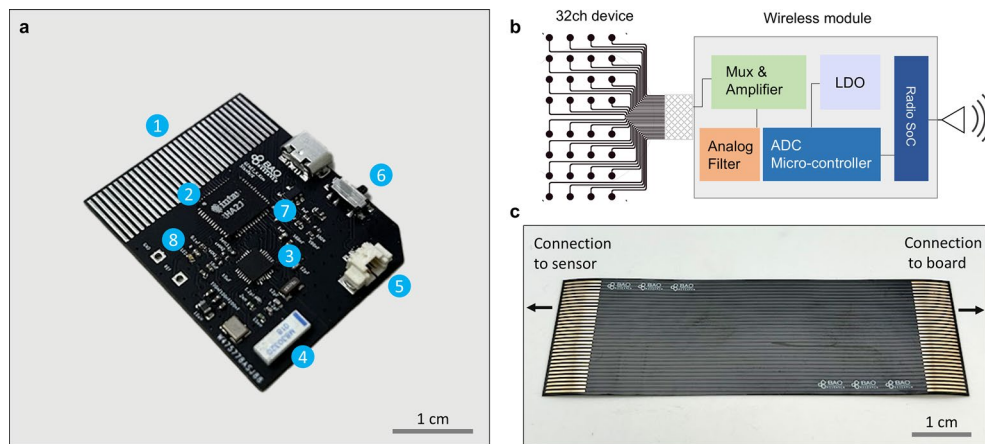
Peer review information *Nature Sensors* thanks Roozbeh Ghaffari and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

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

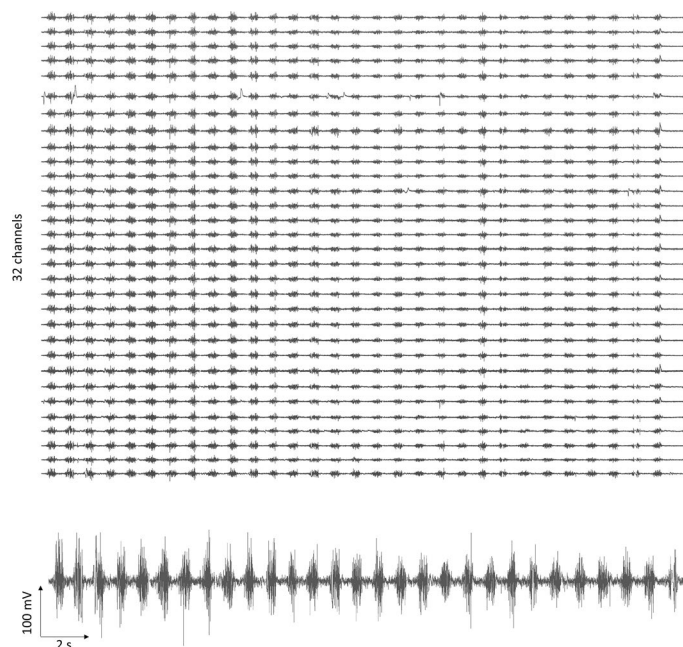
Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2025

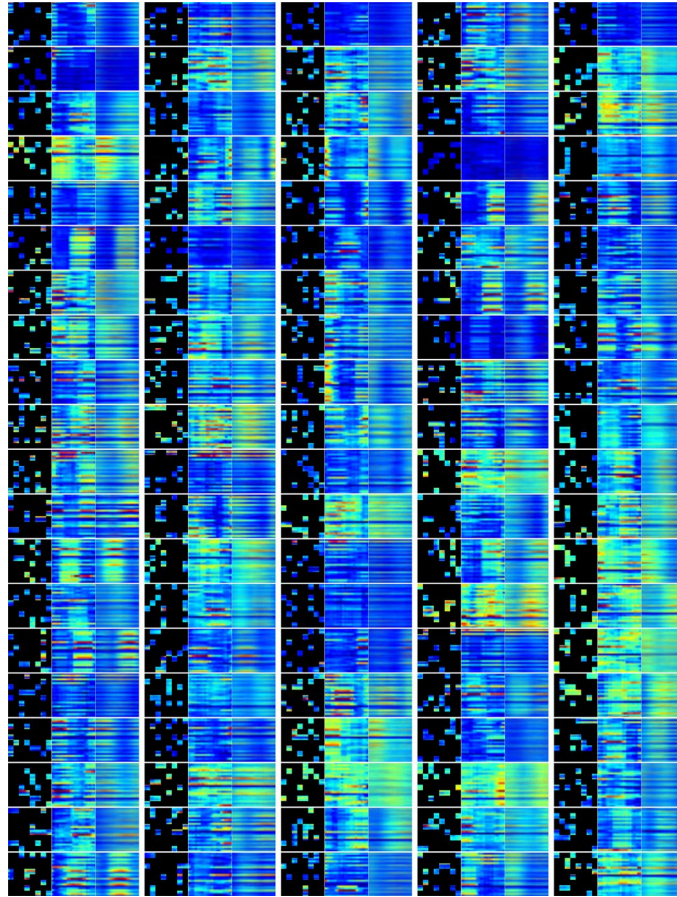


Extended Data Fig. 1 | Wireless sensing module for dataset collection. a, Wireless module including 1) FFC connector, 2) MUX/Amplifier, 3) Micro-controller, 4) Bluetooth Antenna, 5) Battery connector, 6) Power switch, 7)

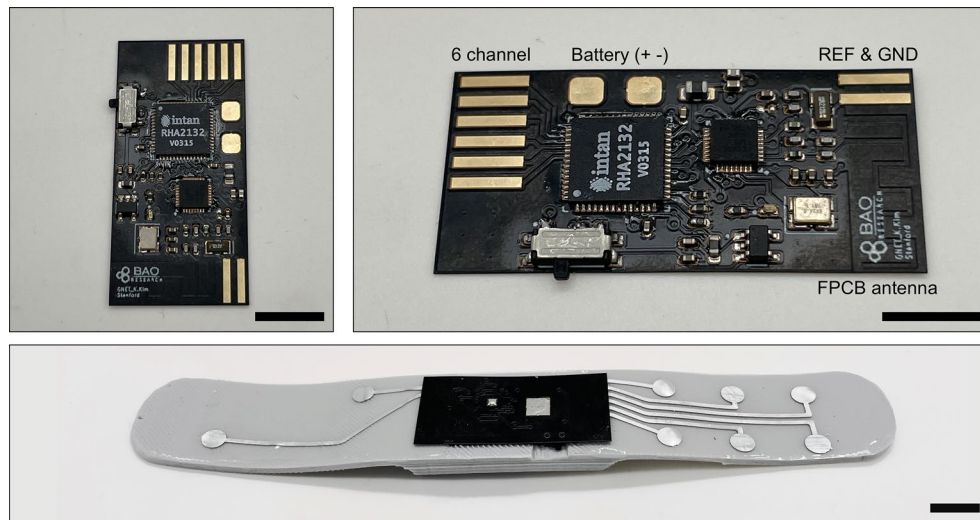
Low-dropout voltage regulator (LDO), and 8) Analog filter. **b,** Illustration of measurement setup where FFC connector bridges between wireless module and the 32-channel EMG device. **c,** Custom made FFC connector with 1 mm pitch.



Extended Data Fig. 2 | Representative 32-Channel EMG Signals Captured During Hand Grasp. Representative examples of 32-channel EMG signals collected during hand grasp movements from the training dataset. The signal from the first channel is magnified and displayed below for clarity.



Extended Data Fig. 3 | 100 samples of generated EMG sets using GenENet. Each set consists of three images: the first represents the masked signals, the second shows the original signals, and the third displays the generated signals.



Extended Data Fig. 4 | Wireless sensing module for a 6-channel system. The module features a 4-layer PCB integrated with a 6-channel electrode array attached to a silicone watch-style device, with extensions to external electrodes made with EGaIn.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
<input type="checkbox"/>	<input checked="" type="checkbox"/> The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
<input type="checkbox"/>	<input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
<input checked="" type="checkbox"/>	<input type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
<input checked="" type="checkbox"/>	<input type="checkbox"/> A description of all covariates tested
<input checked="" type="checkbox"/>	<input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
<input checked="" type="checkbox"/>	<input type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
<input checked="" type="checkbox"/>	<input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
<input checked="" type="checkbox"/>	<input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
<input checked="" type="checkbox"/>	<input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
<input checked="" type="checkbox"/>	<input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	A custom built SoC (CC2650, Texas Instruments) was used for the data collection and the firmware was programmed through Code Composer Studio (CCS).
Data analysis	Data was analyzed using Python. The source code is available at https://github.com/richkim92/GenENet

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The data that support the results within this study is available at <https://github.com/richkim92/GenENet>

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	All participants in the study were male. No gender-based analysis and study to compare between sexes were not performed. Study was not designed to compare between sexes.
Reporting on race, ethnicity, or other socially relevant groupings	Any socially relevant categorizations including race or ethnicity were not collected or considered in this study.
Population characteristics	The study included three male participants. No covariate-relevant population characteristics were collected or considered relevant for the purpose of this study.
Recruitment	Participants were recruited through direct invitation.
Ethics oversight	This protocol is approved through Stanford IRB-54795.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The study involved three participants selected internally from the research team. No formal sample size calculation was performed due to the technical nature of the study in electronics research.
Data exclusions	No data were excluded from the analyses.
Replication	All experimental results were reproducible and are presented in the manuscript.
Randomization	Randomization is not relevant to this study as it involved controlled technical measurements in a non-clinical electronics setting.
Blinding	Blinding was not performed.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.