Two-stage amplification of an ultrasensitive MXene-based intelligent artificial eardrum

Guang-Yang Gou1,2†, Xiao-Shi Li1,2†, Jin-Ming Jian1,2, He Tian1,2‡, Fan Wu1,2, Jie Ren1,2, Xiang-Shun Geng1,2, Jian-Dong Xu1,2, Yan-Cong Qiao1,2, Zhao-Yi Yan1,2, Guanhua Dun1,2, Chi Won Ahn3*, Yi Yang1,2*, Tian-Ling Ren1,2*

We report an artificial eardrum using an acoustic sensor based on two-dimensional MXene (Ti$_3$C$_2$T$_x$), which mimics the function of a human eardrum for realizing voice detection and recognition. Using MXene with a large interlayer distance and micropyramid polydimethylsiloxane arrays can enable a two-stage amplification of pressure and acoustic sensing. The MXene artificial eardrum shows an extremely high sensitivity of 62 kPa$^{-1}$ and a very low detection limit of 0.1 Pa. Notably, benefiting from the ultrasensitive MXene eardrum, the machine-learning algorithm for real-time voice classification can be realized with high accuracy. The 280 voice signals are successfully classified for seven categories, and a high accuracy of 96.4 and 95% can be achieved by the training dataset and the test dataset, respectively. The current results indicate that the MXene artificial intelligent eardrum shows great potential for applications in wearable acoustical health care devices.

INTRODUCTION

Nowadays, flexible pressure sensors have received much attention because of their excellent capability to sense mechanical stimuli in electronic skin (1–3) and intelligent robotics (4, 5). So far, many efforts have been devoted to developing high-performance pressure sensors based on various sensing mechanisms, such as piezoresistive (6), piezoelectric (7), and capacitive effects (8). Among various devices, piezoresistive sensors have the following advantages: outstanding sensing sensitivity, low-cost fabrication, and low-power consumption, which show promise in detecting small mechanical vibrations (9). In particular, the piezoresistive sensors with a higher sensitivity can be used for sound detection by sensing the resistance change relying on the vibration of the film. Among different forms of pressure, sound is a very gentle force with a dynamic frequency. Only very sensitive pressure sensors can be used for sound detection. To date, some sound/acoustic sensors based on resistive effect have realized the function of human voice detection and speech recognition (10, 11). Furthermore, the artificial eardrum with sound-sensing properties has attracted much interest because of its application in implantable acoustical health care devices (12, 13). Previously, Dinh et al. (10) reported a graphene-based acoustic sensor that couples hierarchically structured surface morphology and micro-racked conductive thin film for voice recognition. In addition, Gu et al. (14) demonstrated a flexible microphone based on single-walled carbon nanotubes for recording the human voice. Although these resistive devices have achieved the functionality available for voice real-time detection and recognition, it is still a great challenge for acoustic sensors to achieve high sensitivity, low detection limit, and wide spectrum response (10). The development of sensing materials with nanostructure plays a critical role in improving the response of pressure sensors.

MXenes are a new class of two-dimensional (2D) transition metal carbides and nitrides with a general chemical formula of $\text{M}_{x+1}\text{X}_n\text{T}_x$, where M is a transition metal (Ti, Mo, or Cr), X stands for C or N, $\text{T}_x$ is a surface-terminating functionality (-OH, -O, or -F), and $n = 1, 2$, or $3$ (15, 16). Since their discovery in 2011 (17), MXenes have attracted great attention because of their great promise in energy storage (18), electromagnetic shielding (19), electronics, and sensing applications (20–22). The 2D compound Ti$_3$C$_2$T$_x$ as the most extensively studied MXene exhibits many excellent properties, such as large specific surface area, metallic electrical conductivities, and high transparency (23–25). Combined with the above unique properties and its applications in lithium (Li)–ion batteries and supercapacitors, MXene has already proved to be a promising electrode material (26, 27). Compared with layered graphene or other material, MXene nanoflakes with greatly changed interlayer distances can achieve a highly mechanical response (28), which is suitable as a mechanical sensing material for building piezoresistive sensors to realize sound detection. However, the application of MXene in acoustic sensors for mimicking the function of the eardrum has been rarely studied using this special basic characteristic. Furthermore, speech recognition based on machine-learning methods needs to be further examined to obtain the optimized algorithm (29–31), which can expand the application of acoustic sensors in the field of intelligent artificial eardrums.

Here, we propose an MXene-based artificial eardrum that can be applied to detect human voice. The high-quality MXene nanoflakes with lateral sizes up to several micrometers are prepared by selectively etching the Ti$_3$AlC$_2$ precursors. We designed a polydimethylsiloxane (PDMS) substrate with micropyramid arrays by replicating from a silicon mold to improve the sensitivity to detect voice. The MXene eardrum exhibits an excellent acoustic sensing capability due to a two-stage enhancement of sound detection combining MXene with micropyramid substrates. The machine-learning model was adopted to realize the training and testing of large amounts of data recorded by the MXene device, indicating that the MXene artificial eardrum can be used for real-time voice classification.
Scientists have developed a new type of artificial eardrum that could be used in human-machine interaction systems. The new eardrum is designed to improve sound detection and is made from MXene, a type of two-dimensional material. The MXene artificial eardrum is composed of two stages: a piezoresistive sensor and a flexible substrate. The sensor layer is made from MXene nanoflakes, which can change their resistance in response to external pressure. The substrate layer is a micropyramid flexible substrate that can deform more easily than a cylindrical pillar of the same weight and height. The MXene artificial eardrum combines the two-stage amplification, which can significantly improve the acoustic response capability.

**RESULTS**

**Assembly of the MXene artificial eardrum**

Figure 1A shows the sensing mechanism of our MXene eardrum. Briefly, the piezoresistive sensor is composed of an electrode, a sensing layer, and a substrate. In this approach, at the sensing-layer level, the interlayer spacing of Ti$_3$C$_2$ nanoflakes (1.31 nm) is much higher than that of graphene (0.34 nm) and can be greatly changed under an external pressure, indicating that MXene can offer a sensitive mechanical response. Hence, the MXene can be used as a sensing layer to achieve a first-stage enhancement of sound sensing. At the substrate level, many efforts prove that a micropyramid flexible substrate can have a very small shape factor, which allows a pyramid to deform more easily than a cylindrical pillar with the same weight and height (32, 33). Applying micropyramid PDMS arrays to pressure sensors can promote a second-stage enhancement of sound sensing, which effectively improves the sensitivity and detection limit. Therefore, compared with a conventional sensor, our MXene artificial eardrum can realize the integration of two-stage amplification, which can significantly improve the acoustic response capability.
wire was bonded to form an electrode by using silver paste. Last, double-layer MPP film was laminated together face to face to obtain the eardrum device. Figure 1F shows the scanning electron microscope (SEM) image of the obtained PDMS substrate with pyramid arrays, indicating good surface shape. Figure 1G shows the SEM image of spin-coating MXene on the PDMS film. The MXene nanoflakes are tightly and uniformly stuck around the pyramid arrays to form a conductive network.

The mechanisms and performance of device
To investigate the real-time sound detection of the device, a custom-designed test platform was built, including a digital source meter and a speaker, as shown in Fig. S2A. When the sound pressure produced by the speaker spread to the surface of our artificial eardrum, the film vibration causes its resistance to change, which can be recorded by the digital source meter with a high sampling rate. As the layer distance between two neighboring interlayers in MXene nanoflakes can be changed under external pressure, the piezoresistive sensor shows excellent mechanical sensitivity, as shown in Fig. 2A. To study the resistive response of the artificial eardrum, a simple circuit model can be demonstrated in Fig. 2B and fig. S3. The total resistance around the pyramid arrays mainly consists of three components, namely, resistance on the surface of the pyramid ($R_s$), the resistances at the top of the pyramid ($R_t$), and the resistances without the pyramid ($R_w$), and can be written as (32)

$$R_{total} = R_s + R_t + R_w$$ (1)

By rewriting $R_s$, $R_t$, and $R_w$ in terms of their resistivities ($\rho$), contact area ($A_t$), perimeter ($P_s$), length ($L$), equivalent width ($E_w$), and the thickness of MXene ($D$), the equation can be given as

$$R_{total} = \frac{\rho L_s}{D_s P_s} + \frac{\rho L_t}{A_t} + \frac{\rho L_w}{D_w E_w}$$ (2)

After applying pressure, the sensor response to sound pressure is reflected in its resistance change with respect to the initial state, as expressed in Eq. 3

$$\Delta R = R' - R_0 = \rho \left( \frac{L_s}{D_s P_s} - \frac{L_s}{D_s P_s} \right) + \rho \frac{L_t}{A_t} \left( \frac{1}{A_t} - \frac{1}{A_t} \right) + \frac{\rho}{D_w E_w}$$ (3)

The device bends under sound pressure, which will produce a lateral strain, and the pyramid is compressed. It can be found that adjacent pyramids are squeezed and deformed under sound pressure, resulting in a decrease in the thickness of the MXene on the surface of the pyramid ($D_t$). Meanwhile, lateral strain can also stretch the PDMS film, which can increase the side length of the pyramid ($L_t$) and the length of the part without the pyramid ($L_w$). Hence, our MXene artificial eardrum can realize the integration of two-stage amplification, which can significantly improve the acoustic response capability. To compare with the MXene artificial eardrum, we fabricated the graphene-based device by spin coating the graphene solution on the PDMS-PE substrate with micropyramid arrays. The prepared process is the same as that of the MXene-based artificial eardrum. Figure 2C shows the variations in the resistance of both devices with a decreasing sound pressure level (SPL) at a frequency of 300 Hz to determine their detection limit. The resistance change can be defined as $\Delta R/R_0$, where $\Delta R$ is defined as $(R - R_0)$; $R_0$ and $R$ represent the initial resistance and changed resistance under sound pressure, respectively. It can be seen that the resistance response of the MXene-based device is higher than that of the graphene device at the same SPL, and a stronger value can be obtained as 8.5%. In the sound domain, the signal-to-noise ratio (SNR) can be used to evaluate the degree of voice recognition susceptible to background noises. Figure 2D shows the SNR of the MXene and graphene device at different sound frequencies. The MXene eardrum device can maintain a higher SNR of 50 dB from 200 to 900 Hz and remains 40 dB at higher frequencies up to 2.5 kHz. This result means that the signal detection of our MXene device is ~10,000 times higher than the noise level. Furthermore, the SNR of the graphene device is significantly lower than that of MXene devices at the same frequency. This result can be attributed to the advantage of larger interlayer distances of MXene, which shows good mechanical sensitivity. Hence, our MXene eardrum can exhibit an excellent sensing response to acoustic waves over the wide audible frequency range. To further study the response mechanism, the force equilibrium model of the device is established in fig. S4. Because the double-layer MPP film vibrates under the action of sound waves, the elastic force perpendicular to the film direction can cause the pyramid structure to deform easily. Hence, the MXene eardrum device shows excellent vibrational sensitivity. Figure 2E shows the theoretical prediction SNR of the MXene device at different frequencies. As the frequency of the sound wave increases, the mechanical response of the film gradually decreases.
resulting in a slight change in resistance. Figure 2F depicts the resistance response of various acoustic detectors at different frequencies. It can be seen that the MXene device on the PDMS substrate with a pyramid-patterned structure has a higher resistance response of 5.9% compared to the other devices (10, 14, 34). The highest response originates from the microstructure of the substrate and the excellent mechanical sensitivity of MXene.

To evaluate the sensitivity of the MXene device in response to voice detection, two types of device structures were prepared, namely, an MXene eardrum using a double-layer MPP film and an MXene eardrum using a single-layer MPP film. Figure 3A shows the resistance variations of two types of devices at the same SPL with different frequencies. The eardrum device based on the double-layer MPP film presents a narrow and sharp output response peak across a frequency range of 150 Hz to 3 kHz, with the strongest response being obtained as 5.9% at 300 Hz, and then the response decreases gradually at higher frequencies. Notably, the eardrum device produces a much higher acoustic response than the single-layer MPP film across the overall frequency range. This result can be attributed to the excellent advantage of vibrational sensitivity of the MXene eardrum. Figure S2B shows the time response of the MXene eardrum at an SPL of 93 dB, revealing excellent stability. Figure 2 (B and C) shows the enlarged section of fig. S2B, exhibiting an ultrafast response time and recovery time of 15 and 25 ms, respectively. Sensitivity is one of the most critical parameters for judging the performance of acoustic/pressure sensors. The sensitivity of a pressure sensor can be defined as $S = (\Delta R/R_0) / \Delta P$, where $\Delta R$ is the relative change in resistance, $R_0$ is the pristine resistance of the device under no pressure, and $\Delta P$ is the change of applied force. As shown in Fig. 3D, the sensitivity of the flexible MXene eardrum can reach 62 kPa$^{-1}$ in the pressure range of <1.25 Pa, and the detection limit can reach 0.1 Pa. Figure 3E shows a sensing performance comparison between our MXene eardrum and other pressure sensors reported recently (35–43). The sensing materials and microstructures are summarized in fig. S5. It can be found that our MXene electronic eardrum shows an incomparably high sensitivity in the low-pressure region and an ultralow detection limit, outperforming existing pressure sensors reported in the literature. Although microstructure, such as hollow-sphere structure (36), silk model (37), fractured structure (39), and fiber structure (40, 43), can significantly improve the performance of the devices, the reported devices only realize the first-level amplification of the mechanical signal and cannot achieve a higher sensitivity in the low-pressure region. The reason for the high sensitivity of our MXene electronic eardrum in the low-pressure region is that the micro-structured pyramid arrays on the PDMS substrate and a large interlayer distance of MXene can realize the integration of two-stage amplification of sound sensing. Hence, our MXene artificial eardrum can enhance the coupling of electrical and mechanical properties, which significantly promotes acoustic response capability.

**Voice detection and recognition**

To further demonstrate the applicability of the MXene artificial eardrum, the function to detect sound and recognize a human voice was investigated. We recorded the time-dependent variation in the resistance waveforms produced by the loudspeaker. Three sets of words—“sensor,” “excellent,” and “One World, One Dream”—were played by loudspeaker, and the acoustic waveforms were recorded by our MXene eardrum device. Although the center frequency of our

![Fig. 3. Resistance response and sensitivity of the MXene artificial eardrum.](image-url)
The eardrum is around 3 kHz, the response response to a different word is well synchronous to original audio signals, as shown in Fig. 4A and fig. S6. The characteristic peaks of each word recorded by our device are retained and reflected with fidelity. Furthermore, the recorded signals are analyzed by short-time Fourier transform (STFT) in Fig. 4B, and the corresponding spectrograms indicate that our MXene device exhibits a good frequency response. To verify the ability of our device for long-time voice recording, a speech entitled “Night Watcher Swear” was played on the MXene eardrum and iPhone. Figure S7 shows the corresponding resistance response of the device. The circuit was designed to filter the noise signal below 50 Hz caused by power frequency interference. Figure 4C shows three waveforms, including origin audio, waveforms recorded by iPhone, and waveforms converted from the MXene device. It can be seen that the time-dependent waveform of our MXene device is in agreement with the origin audio waveform of a sentence, indicating excellent acoustic sensing capability. Hence, our flexible MXene eardrum is comparable to that of a commercial voice recorder.

Speech recognition has attracted significant attention because of its potential applications in human-machine interactions due to the convenient bilateral communications (29–31). The machine-learning methods have been widely developed to analyze the signals for speech process under a complex environment. To realize the recognition of data recorded by our MXene eardrum, the voice recognition algorithm based on k-means clustering algorithm (k-means) is our ideal choice owing to its good clustering effect, local optimization, and low complexity. The resistance response waveform of seven kinds of words, namely, “sensor,” “excellent,” “hello,” “happy,” “beautiful,” “science,” and “national,” was recorded by our MXene eardrum, respectively (40 sets per word and a total of 280 sets of data). To allow the response of the device at different SPLs for the same voice, the obtained data are first normalized before being applied as the nominal information for training and testing. Figure 4 (D and E) shows the normalized response waveform of seven kinds of words recorded for the 1st and 40th sets by the MXene eardrum, respectively, indicating the unique characteristic peaks of each

Fig. 4. Application of the MXene artificial eardrum for voice detection and recognition. (A) The voice detection waveform of “One World, One Dream” recorded by the MXene eardrum. The illustration is the corresponding original signal. (B) The spectrogram of “One World, One Dream” analyzed by short-time Fourier transforms (FFTs). (C) Three acoustic signal waveforms including original audio (“Night Watcher Swear”), waveforms recorded by iPhone, and waveforms converted from the MXene device. (D and E) The normalized response waveform of seven kinds of words recorded for the 1st and 40th sets by the MXene eardrum, respectively. (F) The recognition flow diagram of the pronunciation of different voices. (G) Visualizing the pronunciation information of voice within 280 voices adopting t-distributed stochastic neighbor embedding (t-SNE) dimensionality reduction. (H) Confusion matrix of the voice’s prediction versus the test dataset.
voice. Furthermore, the waveform of each voice recorded for the 40th set is almost the same as the waveform recorded for the first time, which shows that our device has a good stability. Figure 4F shows the flow diagram of k-means-based machine learning to process the pronunciation of different words. In machine learning, the 140 sets of data are used for training data, and the remaining 140 sets of data are applied to testing data. The training dataset based on k-means method can be obtained by deeply analyzing the characteristics of normalized data. Then, the obtained algorithm model can be applied to test the dataset. Eventually, a high accuracy of 96.4 and 95% for the training dataset and the test dataset can be realized, respectively. The reason for a higher accuracy can be attributed to the high sensitivity of the device in the low-pressure region, which can enable the consistency of the device for each voice. Hence, the amount of information obtained from the same sound signal is very similar, which can achieve a high accuracy through machine-learning algorithms. Figure S8 shows the average similarity of the seven sets of voices recorded by our MXene electronic eardrum. It can be seen that our MXene electronic eardrum exhibits an excellent average similarity, indicating outstanding consistency for each voice. Therefore, the algorithm models can be easily trained, and the recognition ability of different types of signals has been improved. The t-distributed stochastic neighbor embedding can be used to visualize the high-dimensional data of similar characteristics into a low-dimensional space (44). As shown in Fig. 4G, the 280 sound sets are visualized to form seven clusters of different colors, and each point on the graph represents voice information. It can be seen that the points with the same color are very close, indicating excellent recognition and classification capabilities. The confusion matrices are applied to display the discordance between the voice’s prediction and test dataset, as shown in Fig. 4H. The results indicate that our optimized model exhibits a higher accuracy for each voice, which also reflects the high stability of our MXene device. Hence, the MXene eardrum based on our machine-learning algorithm shows great potential in the biomedical field for helping hearing-impaired people.

DISCUSSION
In summary, we have fabricated the high-quality MXene nanoflakes by selectively etching the Ti3AlC2 precursors. To achieve an ultrasensitive detection of sound, the PDMS-PE substrate with micropyrramid arrays was also prepared by replicating from a silicon mold at the same time. A high-performance MXene artificial eardrum was realized on the prepared PDMS-PE substrate based on the obtained Ti3C2 MXene nanoflakes. Because of the two-stage amplification of sound pressure, the MXene artificial eardrum shows an unprecedented sensitivity of 62 kPa−1 and a very low detection limit of 0.1 Pa. Hence, the as-fabricated MXene eardrum exhibits an acoustic sensing capability. Last, the machine-learning model was successfully applied to realize the training and testing of data recorded by the MXene eardrum. A high accuracy of 96.4 and 95% can be obtained from the training dataset and the test dataset, respectively. These results indicate that our MXene artificial eardrum demonstrates excellent stability, which shows potential in building a human-machine interaction system.

MATERIALS AND METHODS
Synthesis of Ti3C2 MXene solution
Ti3C2 MXene was synthesized following our previous method (45). The etching solution was prepared by adding 2 g of LiF (>99.98%,<100 μm) powder to 20 ml of 6 M HCl solution. Then, 2 g of Ti3AlC2 powder was slowly immersed in the etchant to selectively etch the Al layers at room temperature and stirred for 18 hours. The acidic suspension was then washed with deionized (DI) water until pH > 6 via centrifugation at 5000 rpm for 10 min. The prepared Ti3C2 powder was dispersed in DI water by ultrasonication for 1 hour. The suspension containing Ti3C2 nanoflakes was collected after centrifuging at 3500 rpm for 5 min.

Fabricating of Si mold
A 〈100〉 Si wafer with a 300-nm SiO2 layer was patterned to 10 μm by 10 μm by photolithography. Then, the patterns were etched in 300 nm SiO2 by using the reactive ion etching. As the KOH solution has anisotropic characteristics in wet-etching Si, the substrate was then etched with a 40% KOH solution for 10 min at 80°C. Last, the Si mold with a pyramid structure was prepared. The height of the pyramid structure is about 7 μm, and the spacing is 10 μm.

Preparation of flexible MXene artificial eardrum
To easily peel off the PDMS from the Si mold, the obtained Si mold was first treated with a release agent (chlorotrimethylsilane) in an oven at 90°C for 30 min. A 10:1 mixture of PDMS was spin-coated on the Si mold at a speed of 5000 rpm to prepare the PDMS substrate with a pyramid pattern. Then, the Si mold with PDMS was precured at 70°C for 30 min, and the 10-μm PE film was laminated on top of the PDMS surface. Subsequently, the PE-PDMS film was peeled off from the Si mold. The Ti3C2 MXene solution was synthesized through etching Ti3AlC2 powder by using the HCl solution with LiF powder. The obtained MXene nanoflake solution was spin-coated on the prepared PDMS-PE substrate with a pyramid structure to form an MPP conductive film. The contact electrodes were formed by using copper wire glued through silver paste. Last, two MPP-conductive films were laminated face to face to form an MXene eardrum.

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abn2156

REFERENCES AND NOTES
1. X. Zhao, Z. Zhang, Q. Liao, X. Xun, F. Gao, L. Xu, Z. Kang, Y. Zhang, Self-powered user-interactive electronic skin for programmable touch operation platform. Sci. Adv. 6, 4294 (2020).
2. M. Zhu, M. Lou, J. Yu, Z. Li, B. Ding, Energy autonomous hybrid electronic skin with multi-modal sensing capabilities. Nano Energy 78, 105208 (2020).
3. Y. Cai, J. Shen, C. W. Yang, Y. Wan, H. L. Tang, A. A. Albabr, C. Chen, J. H. Fu, X. Wei, K. W. Huang, Y. Han, S. Jonas, X. Dong, V. Tung, Mixed-dimensional MXene-hydrogel heterostructures for electronic skin sensors with ultrabroad working range. Sci. Adv. 6, eabb5367 (2020).
4. S. Lee, A. Reuveny, J. Reeder, S. Lee, H. Jin, Q. Liu, Y. Yokota, T. Sekitani, T. Isayama, Y. Abe, Z. Suo, T. Someya, A transparent bending-insensitive pressure sensor. Nat. Nanotechnol. 11, 472–478 (2016).
5. J. C. Yang, J. Mun, S. Y. Kwon, S. Park, Z. Bao, S. Park, Electronic skin: Recent progress and future prospects for skin-attachable devices for health monitoring, robotics, and prosthetics. Adv. Mater. 31, 1904765 (2019).
6. Y. Cheng, Y. Ma, L. Li, M. Zhu, Y. Yue, W. Liu, L. Wang, S. Jia, C. Li, T. Qi, J. Wang, Y. Gao, Bioinspired microspines for a high-performance spray Ti3C2Tx MXene-based piezoresistive sensor. ACS Nano 14, 2145–2155 (2020).
7. G. Tian, W. Deng, Y. Gao, D. Xiong, C. Yan, X. He, T. Yang, L. Jin, X. Chu, H. Zhang, W. Yan, W. Yang, Rich lamellar crystal baklava-structured PZT/PVDF piezoelectric sensor toward individual table tennis training. Nano Energy 59, 574–581 (2019).
8. N. Bai, L. Wang, Q. Wang, J. Deng, Y. Wang, P. Lu, J. Huang, G. Li, Y. Zhang, J. Yang, K. Xie, X. Zhao, C. Guo, Graded intrafolial architecture-based piezometric pressure sensor with ultra-broad-range high sensitivity. Nat. Commun. 11, 209 (2020).
11. S. Kang, S. Cho, R. Shanker, H. Lee, J. Park, D. S. Um, Y. Lee, H. Ko, Transparent...

9. C. Wang, J. Zhao, C. Ma, J. Sun, L. Tian, X. Li, F. Li, X. Han, C. Liu, C. Shen, L. Dong, J. Yang...

26. Y. Dong, H. Shi, Z. S. Wu, Recent advances and promise of MXene-based nanomaterials...