

ARTICLE OPEN



Multiple machine learning approach to characterize two-dimensional nanoelectronic devices via featurization of charge fluctuation

Kookjin Lee^{1,2,3,11}, Sangjin Nam^{4,11}, Hyunjin Ji⁵, Junhee Choi⁶, Jun-Eon Jin⁷, Yeonsu Kim³, Junhong Na⁸, Min-Yeul Ryu⁸, Young-Hoon Cho⁷, Hyebin Lee⁷, Jaewoo Lee³, Min-Kyu Joo^{9,10} and Gyu-Tae Kim³

Two-dimensional (2D) layered materials such as graphene, molybdenum disulfide (MoS₂), tungsten disulfide (WSe₂), and black phosphorus (BP) provide unique opportunities to identify the origin of current fluctuation, mainly arising from their large surface areas compared with those of their bulk counterparts. Among numerous material characterization techniques, nondestructive low-frequency (LF) noise measurement has received significant attention as an ideal tool to identify a dominant scattering origin such as imperfect crystallinity, phonon vibration, interlayer resistance, the Schottky barrier inhomogeneity, and traps and/or defects inside the materials and dielectrics. Despite the benefits of LF noise analysis, however, the large amount of time-resolved current data and the subsequent data fitting process required generally cause difficulty in interpreting LF noise data, thereby limiting its availability and feasibility, particularly for 2D layered van der Waals hetero-structures. Here, we present several model algorithms, which enables the classification of important device information such as the type of channel materials, gate dielectrics, contact metals, and the presence of chemical and electron beam doping using more than 100 LF noise data sets under 32 conditions. Furthermore, we provide insights about the device performance by quantifying the interface trap density and Coulomb scattering parameters. Consequently, the pre-processed 2D array of Mel-frequency cepstral coefficients, converted from the LF noise data of devices undergoing the test, leads to superior efficiency and accuracy compared with that of previous approaches.

npj 2D Materials and Applications (2021)5:4; <https://doi.org/10.1038/s41699-020-00186-w>

INTRODUCTION

Low-frequency (LF) 1/f noise spectroscopy is a nondestructive defect diagnosis tool, which identifies dominant scattering origins. Such scattering origins are caused by imperfect crystallinity, lattice vibration, surface trap distribution, and channel and dielectric defects, in addition to the Schottky barrier inhomogeneity at the metal-semiconductor interface in semiconductor devices^{1–6}. However, as the size of the channel material decreases, particularly in the case of two-dimensional (2D) layered materials, their atomically thin nature with a large surface-to-volume ratio makes it significantly difficult to investigate them using LF 1/f noise analysis as compared with their bulk silicon counterparts^{7,8}. Conventionally, the time-resolved current (*I*) variation in electronic devices has been ascribed to the carrier number and/or mobility fluctuation⁹; $\Delta I(t) \propto q\mu(\Delta N) + q(\Delta\mu)N$, where *q*, μ , and *N* denote the elementary unit charge, carrier mobility, and number of charge carriers, respectively. However, since the inherent vulnerability of 2D materials to surrounding interfaces considerably influences the charge fluctuation, this high sensitivity of LF noise features would reflect the individual effects of both channel and dielectric materials in addition to the presence of chemical/electrical doping^{10,11}.

Thus far, numerous LF noise features have been reported on 2D materials, such as the presence of electron-hole puddle induced charge scattering on graphene¹², the Coulomb scattering

suppression via high- κ passivation of black phosphorus (BP)¹³, the promotion of charge fluctuation in molybdenum disulfide (MoS₂) due to the height and inhomogeneity of the Schottky barrier¹¹, the anisotropic LF noise feature of rhenium disulfide (ReS₂)¹⁴, and the thickness-dependent Coulomb scattering parameter of molybdenum ditelluride (MoTe₂)¹⁵. These studies indicate the high feasibility of LF noise spectroscopy as a tool to classify the material and device properties. Nevertheless, the origin of carrier fluctuation, occurring either in the 2D layered material itself or at the interface between the 2D layered material and the gate dielectric, has not been identified clearly. Moreover, it is significantly difficult to identify an individual noise source from the LF noise data without appropriate data processing for the model-dependent LF noise analysis.

Most recently, the combination of artificial-intelligence (AI) based approach and scientific data analysis has been widely considered in various applications such as healthcare^{16,17}, image recognition^{18,19}, voice search²⁰, and molecular/material science^{21,22}. Further, it has also been determined that these combined techniques are suitable for solving the problems associated with non-linear processes or enormous combinatorial spaces with high efficiency^{16,22–24}. This clearly indicates that the machine learning (ML) and deep learning (DL) approaches can provide a better optimization and decision-making by converging the scientific data and extracting interpretable models from these

¹imec, Leuven, Belgium. ²Department of Materials Science, KU Leuven, Leuven, Belgium. ³School of Electrical Engineering, Korea University, Seoul, Republic of Korea. ⁴Department of Computer Science and Engineering, Korea University, Seoul, Republic of Korea. ⁵School of Electrical Engineering, University of Ulsan, Ulsan, Republic of Korea. ⁶Defense agency for technology and quality, Jinju-si, Gyeongsangnam-do, Republic of Korea. ⁷Samsung Electronics Co. Ltd, 1 Samsung-ro, Yongin-si, Gyeonggi-do, Republic of Korea. ⁸Samsung Display Co. Ltd, 1 Samsung-ro, Yongin-si, Gyeonggi-do 17113, Republic of Korea. ⁹Department of Applied Physics, Sookmyung Women's University, Seoul, Republic of Korea. ¹⁰Institute of Advanced Materials and Systems, Sookmyung Women's University, Seoul 04310, Republic of Korea. ¹¹These authors contributed equally: Kookjin Lee, Sangjin Nam. ✉email: mkjoo@sookmyung.ac.kr; gtkim@korea.ac.kr

data automatically^{22,25}. Recently, studies on applying ML or DL to analysis of 2D layered materials have been widely conducted^{26–29}.

In this study, we introduce an effective technique to classify and infer the characterization of current fluctuation with high efficiency and precision by combining AI and LF noise spectroscopy. Due to the similarities of the fabrication process, geometry, bandgap, and mobility of 2D FETs, classifying only using fundamental DC analysis of transfer curves and output characteristics is very difficult. On the other hand, since LF noise data measure the tiny fluctuations of carriers in channel according to time, characterizing an own FET is easily explained. Based on the time-resolved $\Delta I(t)$ measured from various 2D material-based field-effect transistors (FETs), 2D arrays of the Mel-frequency cepstral coefficient (MFCC) for several electronic properties were considered, and the corresponding features were obtained via a hidden Markov model (HMM). HMM has disadvantages that it must have a relatively large amount of data and hardly express dependencies between hidden states. However, HMM is suitable for processing a large amount of LF noise data based on the advantages of having a strong statistical foundation and enabling efficient learning from raw sequence data. This approach allows us to automatically identify essential device information such as the type of 2D channel materials and gate dielectrics, interface trap density (N_{it}), Coulomb scattering parameter (a_{SC}), and the presence of chemical and electron beam doping. Therefore, the combination of factors such as channel material, gate dielectric, contact metal, and electron beam irradiation significantly affects carrier fluctuations as a function of time. This combination, which has more than 100 LF noise data sets under 32 conditions, becomes a catalyst for machine learning that automatically and effectively classify the characteristics of various nanoelectronic devices. In addition, the obtained LF noise spectroscopy data are highly interpretable via machine learning techniques, thereby identifying the contribution of engineered features in characterizing the device information and performance.

RESULTS

Workflow for audio and current signal classification

The decimal data type, measured in the time domain, has been used generally for ML and DL in data science; however, the Fourier transform (FT) of this data are frequently employed in ML algorithms to improve data interpretation^{30,31}. The process of transforming raw data into a suitable representation for a learning algorithm is often called featurization. For instance, in speech recognition, proper methodology has been widely studied to convert a signal from the time domain to the frequency domain for more accurate classification and analysis^{31–33}. A typical data demonstration method that extracts the characteristics of the original audio signal through the Mel-frequency cepstral coefficient is illustrated in Fig. 1a. Each speech frame of the time domain signal is first obtained through the pre-emphasis, framing, window, and other processing of the original audio signal as expressed in schematic (i) of Fig. 1a. Subsequently, the speech signals comprising a 30 ms frame window are Fast Fourier transformed (FFTed) with a Hamming window. Further, each spectrum signal is processed by Mel filters (26 filters) to obtain the corresponding Mel-frequency spectrum. Finally, the Mel-frequency spectrums are processed using discrete cosine transform (DCT) to acquire the MFCCs in the cepstral domain as shown in schematic (ii) of Fig. 1a.

We employ this data processing algorithm for a number of $\Delta I(t)$ data obtained from various 2D material based FETs which have been fabricated and analyzed under various experimental conditions such as different gate dielectrics^{11,13,34,35}, temperatures^{11,34,36}, channel materials^{10,11,13,34,35,37–39}, chemical/electron beam doping^{40,41}, and source/drain contact metals^{11,34} (see

Fig. 1b). More than 100 LF noise data sets of various 2D layered FETs were considered in this study under 32 different conditions at a particular gate (V_G) and drain (V_D) bias condition. In contrast to the audio signal shown in Fig. 1a, after performing the additional signal normalization process, each MFCC of the current signal in the cepstral domain is consequently determined via FFT and DCT as displayed in schematic (i) and (ii) of Fig. 1b. The MFCCs of the audio and current signals, which comprise the 2D array, are respectively used in speech recognition and device classification (materials/characteristics) through the inference process using ML with the optimized algorithm (see Fig. 1c). The conditions of the device that ML trained and learned in this algorithm distinguishes are as follows (see Fig. 1d): BP, graphene, MoS₂, ReS₂, MoTe₂, and tungsten diselenide (WSe₂) were used as channel materials; h-BN and SiO₂ were employed as gate oxides; Ti, Au, Pt, and Cr were used as the contact metals; and passivation, temperature variations, triethanolamine (TEOA) doping, and electron beam irradiation were considered as the different external factors.

Process flowchart for learning and classifying 2D transistor

The $\Delta I(t)$ of 2D material-based FETs under several conditions were measured at a particular V_G and V_D (see Fig. 2a) in a shielding metal box (see Fig. S2 and Note 2 in the Supplementary Materials for details of the LF noise measurement system)⁴². The drain current I_D can be defined as the sum of the average statistic (DC) drain current (\bar{I}_D) and low-noise current fluctuations (ΔI_D); $I_D = \bar{I}_D + \Delta I_D$ ^{3–6,9}. Since the amplitude of ΔI_D is substantially smaller than I_D , ΔI_D is generally converted to the voltage signal using the low-noise current-to-voltage preamplifier, as depicted in Fig. 2b. The amplified noise signal was considered as the input $\Delta I_D(t)$ data used in Python, where the amplitude normalization and pre-emphasis processes were performed, as presented in Fig. 2c. Subsequently, the preprocessed $\Delta I_D(t)$ data were separated into specific frames with respect to the time domain, and FFT was performed on these data. The transformed data produced by each frame were expressed as power spectral density (S_i) in the frequency domain, and all S_i were filtered onto the Mel scale. This transformation of specific frames into S_i allowed the evaluation of periodic spectra, and the amount of spectral energy between frequencies could then be obtained by combining the respective frames. It was observed that the Mel scale filter interval was directly proportional to the frequencies i.e., narrow around low frequencies and became wider at the higher the frequencies indicating that the Mel-scale filter amplified the amount of energy around low frequencies (see Fig. S4 and Note 3 in the Supplementary Materials)^{30,43–47}.

The obtained data, which are called Mel-frequency spectrums and mainly used for learning, were consequently more sensitive to the low frequency values, allowing a precise carrier scattering analysis in the devices. Subsequently, the Mel-frequency spectrums were transformed through the DCT and extracted to a finite data point sequence, composed of the current MFCCs in the cepstral domain^{43–45}. Further, all S_i filtered by the Mel scale were overlapped, indicating the existence of correlations between the spectral densities. These correlations could be separated using the DCT method. The current MFCCs transformed by the DCT was expressed as the change in filter energy, and a part of them was extracted to store data as 2D arrays, as demonstrated in the schematic in Fig. 2c^{30,47}. Based on the research conducted thus far, the engineered current MFCC features were characterized into 2D arrays with a number between 200 and 1000 for each device class.

Every engineered current MFCCs feature was stored by class, based on the device conditions, and they were learned and classified using ML with an HMM algorithm and DL with an NN, as illustrated in Fig. 2d, e. The HMM based on the Markov chain^{30,31,48,49} in Fig. 2d was the first algorithm model used for

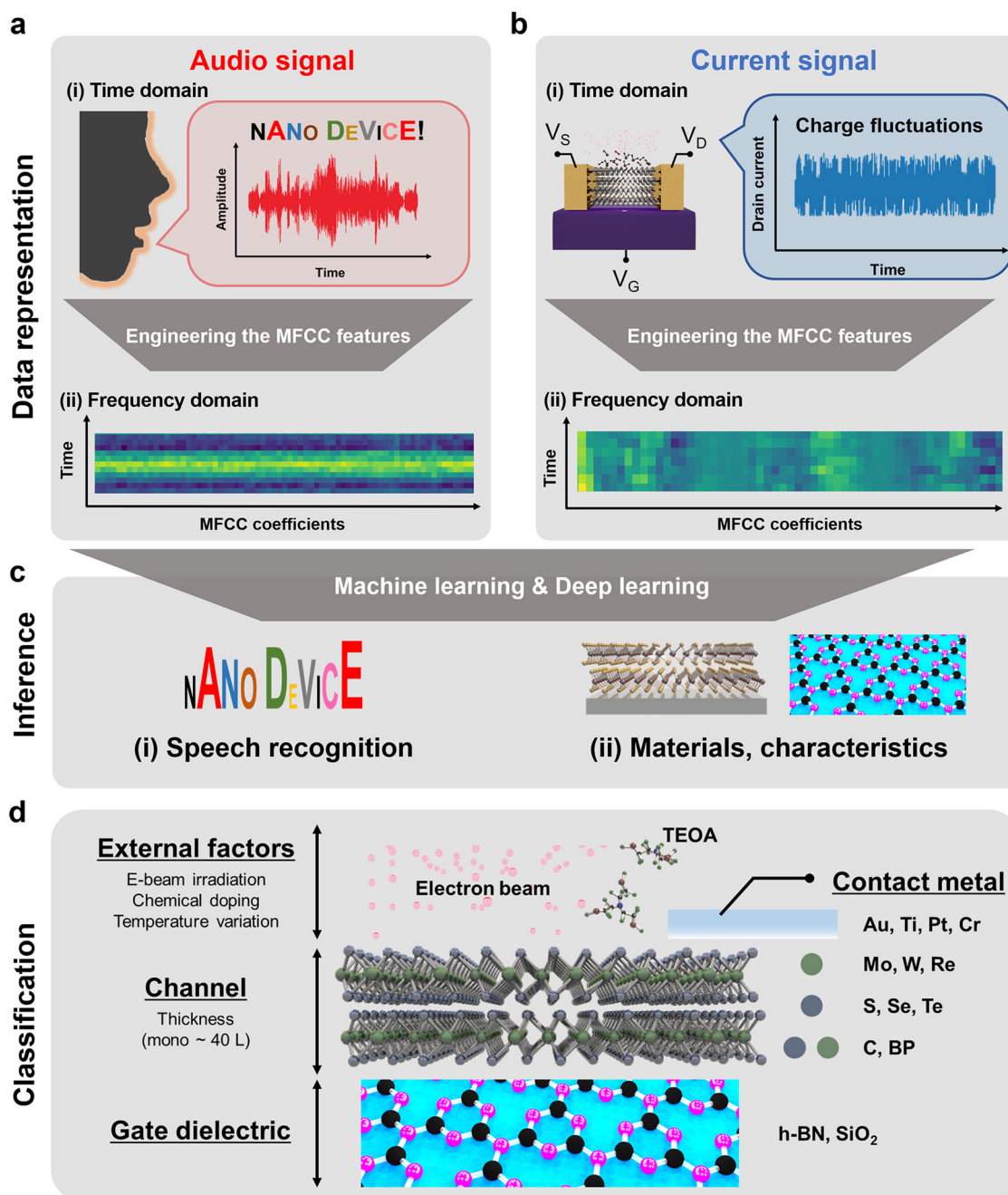


Fig. 1 Workflow comparison between speech and device characteristic classification. The initial steps of the data representation for training: **a** the audio signal and **b** the current signal (i) in the time domain and (ii) in the frequency domain (the darker color, the smaller value). **c** The inference steps after ML for (i) speech recognition and (ii) materials/characteristics classification. **d** Side views of the device structures, which are fabricated using various materials after being subjected to external factors such as e-beam irradiation, triethanolamine (TEOA) chemical doping, and temperature variations; Au, Ti, Pt, and Cr are used as the source/drain contact metals; silicon dioxide (SiO₂) and hexagonal boron nitride (h-BN) are used as the gate dielectrics; the channels are composed of a combination of various atoms such as Mo, W, S, Se, Te, C, and black phosphorus (BP); MoS₂, MoTe₂, WSe₂, ReS₂, graphene, and BP have thicknesses varying from monolayer through to 40 layers.

learning and classifying the data in this study. In HMM, the unaligned training sequences are processed by iteratively evaluating the data stored as current MFCCs. For all the training parameters, the estimates with prior probability distributions are assumed using a maximum *a posteriori* approach^{30,50,51}. The scores for the 32 classes were calculated as \mathbf{Y} , as shown in Fig. 2d. Further, the class with the highest score was determined, and could be used to infer the device conditions as shown in Fig. 2f.

The second method used was the NN^{30,52,53}, which is one of the DL methods. In this algorithm, the \mathbf{Y} values of classes, which had been calculated by HMM, were classified by performing one additional learning step. In this method, the input score vectors, \mathbf{Y} , were transferred to the first layer (layer-1) with 32 perceptrons, which is the number of classes, and were then transferred to the second layer (layer-2) by employing a rectified linear unit (ReLU) function as the activation function^{54,55}. Instead of the widely used

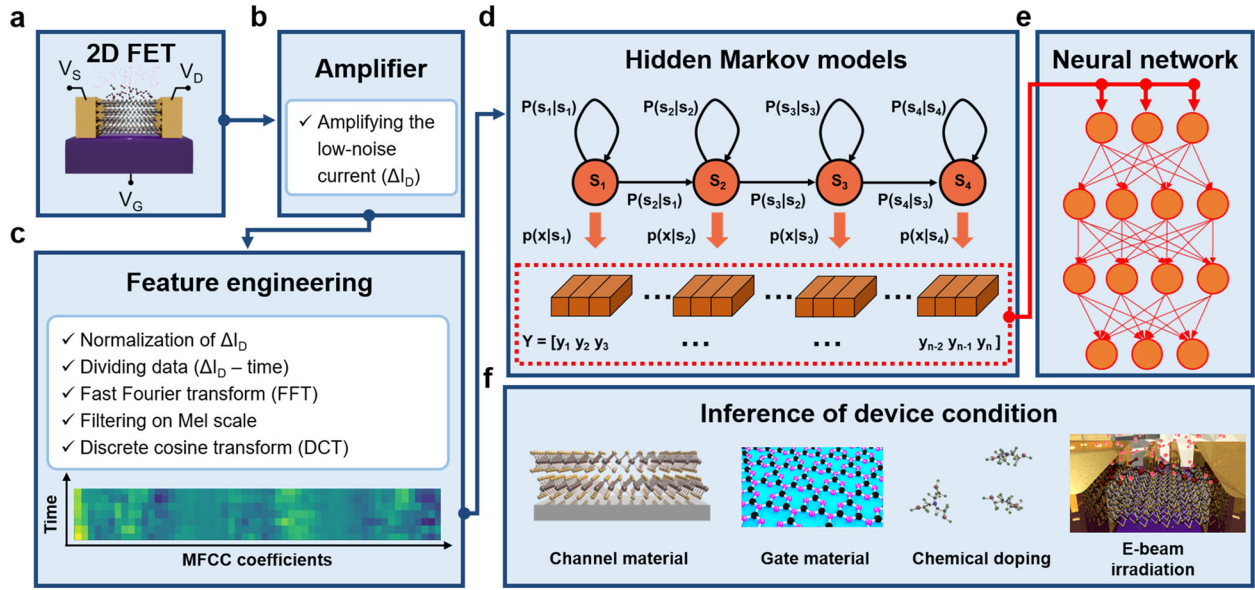


Fig. 2 Flowchart for learning and classifying characteristics of 2D transistors. **a** Schematic of a 2D layered FET, which was measured at a given V_D and V_G under various conditions in the shielded state; **b** amplification of the measured current signal using a low-noise current amplifier; **c** process of feature engineering the input data into a suitable representation (current MFCCs) through MFCC (the darker color, the smaller value); **d** ML with HMM algorithms using the current MFCCs, which comprises the 2D array; **e** deep learning process to re-learn into neural network (NN) through the score vector (\mathbf{Y}) extracted via ML with HMM the algorithm using current MFCCs; and **f** inference steps of device conditions (channel material, gate material, chemical doping, and e-beam irradiation) through ML with the HMM algorithm and deep learning with the NN algorithm.

sigmoid function, we considered a ReLU function here as the activation function because of its sparse activation property, which could be partially activated by providing zero as an output against a negative input^{30,54,55}. Subsequently, the score vector data were classified using the softmax function, which is used for classification in layer-2, and the probability of a specific class was calculated and classified, demonstrating a normalization effect. The softmax function was obtained by dividing the sigmoid value of each class by the sum of sigmoid values of all classes as described below^{30,54}:

$$S(y_i) = \frac{e^{y_i}}{\sum_{j=1}^i e^{y_j}} \quad (1)$$

Compared to the HMM method, the second method had the advantage of classifying the score via repetitive training and learning, which was performed to determine the maximum value of the scores, \mathbf{Y} , obtained through HMM. Finally, the device conditions could be inferred as indicated in Fig. 2f.

Data featurization

Numerous electrical properties of FETs, such as the carrier type (electrons or holes), field-effect mobility, subthreshold swing, and current on/off ratio of the device-under-test (DUT) can be determined from the $I_D - V_G$ transfer characteristics of 2D layered FETs (see Fig. 3a and Supplementary Materials Fig. S1). However, a precise classification of the 2D FETs with fundamental DC analysis is significantly challenging, due to the similarities of mobility, bandgap, geometry, and fabrication process of 2D FETs, except under a few specific conditions such as graphene FET. For instance, $\Delta I_D(t)$ can be measured during 0.5 s at a particular V_G and V_D in a device belonging to a specific class (condition) after excluding \bar{I}_D as illustrated in Fig. 3b. It is noteworthy that we only considered $\Delta I_D(t)$ data where \bar{I}_D was larger than 100 nA to avoid a possible error caused by the minimum detection limit of our system. Subsequently, the current normalization process was performed for $\Delta I_D(t)$ and divided into 11 frames with a 200 ms window. The S_i of each frame was converted using FFT as shown

in Fig. 3c and converted into a vector, x_{n_i} , possessing 100 current MFCC elements, a_{nm} , via Mel-scale filtering and DCT. The x_n for each frame was concatenated to create the current MFCC 2D array of each class (condition), $X_{(class)_i}$ as indicated in Fig. 3d.

$$X_{(class)_i} = [x_1 \ x_2 \ \dots \ x_n \ \dots \ x_{10} \ x_{11}] \quad (2)$$

$$x_n = [a_{n1} \ a_{n2} \ \dots \ a_{nm} \ \dots \ a_{n99} \ a_{n100}] \quad (3)$$

where i depends on the specific voltage applied to the device belonging to the class (condition).

In order to examine the high feasibility of our approach, we considered the carrier number fluctuation-correlated mobility fluctuation (CNF-CMF) model to interpret our $\Delta I_D(t)$ data (see the detailed LF noise theory in Supplementary Materials Note 2). This CNF-CMF model ascribes $\Delta I_D(t)$ to the carrier number fluctuation (CNF) caused by trapping/detrapping phenomena in the interface traps between the channel and gate dielectric in addition to the correlated mobility fluctuation. More specifically, $\Delta I_D(t)$ data can be influenced by many factors such as the carrier type of channel, interface quality and condition between the gate oxide and channel, and the presence of doping (see Fig. 3e). According to the CNF-CMF model, the drain current normalized S_i can be expressed as follows^{5,6,9,56}:

$$\frac{S_i}{\bar{I}_D^2} = \frac{q^2 k T N_{it}}{f \gamma W L C_{ox}^2} \left(1 + \frac{\alpha_{SC} \mu_{eff} C_{ox} \bar{I}_D}{g_m} \right)^2 \left(\frac{g_m}{\bar{I}_D} \right)^2 \quad (4)$$

where q is the carrier charge, k is the Boltzmann constant, T is the absolute temperature, f is the frequency, γ is the frequency exponent, C_{ox} is the dielectric capacitance per unit area, g_m is the transconductance ($=\Delta \bar{I}_D / \Delta V_G$), S_{vfb} is the flat-band voltage spectral density, and μ_{eff} is the effective mobility. The trapped carriers near the channel-gate dielectric interface not only cause variations in S_{vfb} , but also degrade electron mobility, resulting in modulation of the carrier density.

Figure 3f shows the representative S_i of each frame in the frequency domain among the fabricated DUTs. The observation of certain harmonics in the S_i could be attributed to the carrier

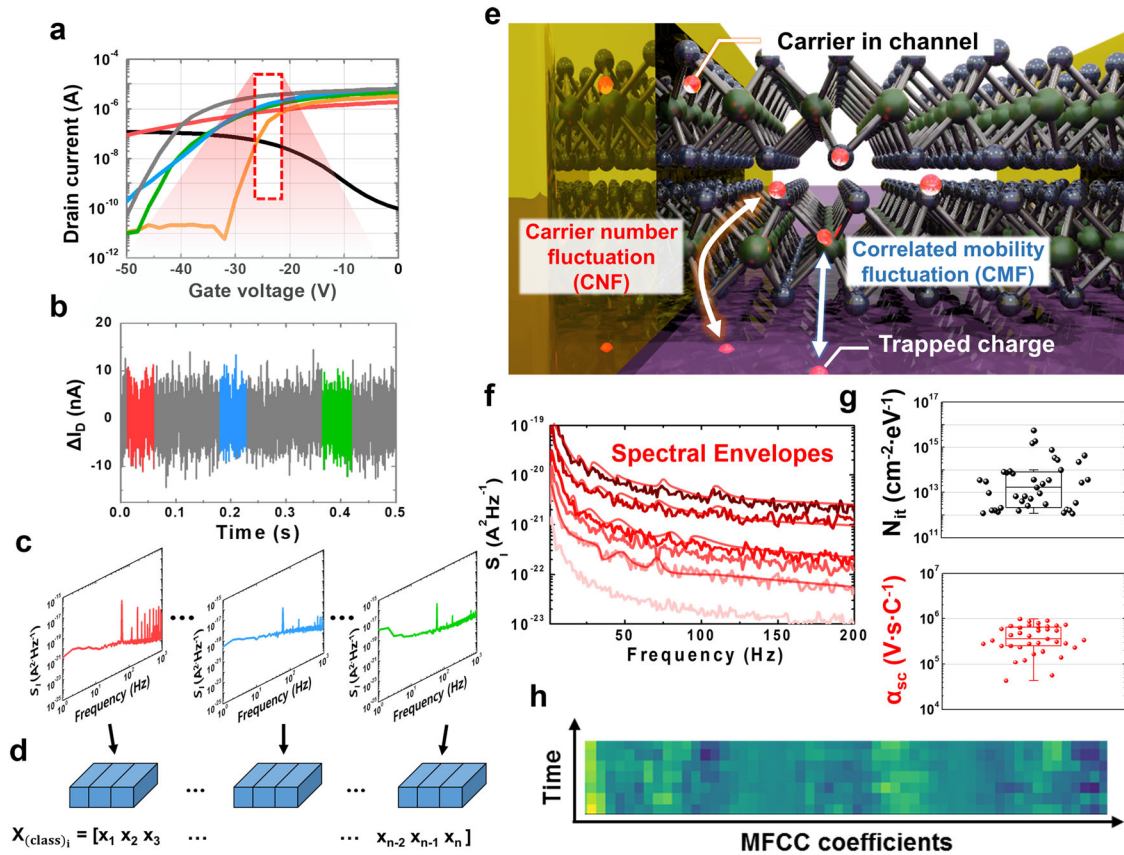


Fig. 3 Detailed flowchart of ΔI_D featurization. **a** Transfer characteristics ($I_D - V_G$) of 2D layered FETs measured under various conditions; **b** $\Delta I_D(t)$ data measured at a particular V_D and V_G and then divided into specific frames with sampling period; **c** current power spectral densities (S_I) in the frequency domain of each frame converted through FFT; **d** current MFCC comprising 2D array, $X_{(class)_i}$ obtained by concatenating current MFCC vector, x , processed through Mel filter and DCT; **e** sectional illustration of carrier behavior in a 2D layered FET; **f** $S_I(f)$ at a particular V_D for several V_G and spectral envelopes (the darker color, the larger V_G); **g** N_{it} and a_{SC} distributions, which were calculated by carrier number fluctuation-correlated mobility fluctuation (CNF-CMF) model of each class (the box plots are defined by 25th and 75th percentile); **h** engineered current MFCC 2D array, which contains carrier behaviors (the darker color, the smaller value).

trapping/de-trapping process in the gate oxide trap sites. In fact, these harmonics assisted in understanding the characteristics of each device and expressed these characteristics as spectral envelopes with specific peaks^{32,33,57,58}. Therefore, N_{it} and a_{SC} between the gate dielectric and the channel of each class (condition) have a significant effect on the spectral envelopes and unique characteristics of device (see Fig. 3g). As a result, the current MFCC 2D array comprises power spectral sequences for each frequency (amplified for the low frequency region), as demonstrated in Fig. 3h. The HMM algorithm that learns the previous state and infers the next state is efficient for learning the current MFCC 2D array that contains N_{it} , a_{SC} , and γ information according to the frequency sequence.

Current MFCC and classification accuracy

LF MFCC element parts (from a_{n1} to a_{n40}) of engineered current MFCC at the same $\bar{I}_D \approx 1 \mu A$ in each class ((i) graphene on trench structure, (ii) MoS₂ on SiO₂, (iii) MoS₂ on SiO₂ after e-beam irradiation, and MoS₂ on h-BN at (iv) $T = 25$ K, (v) $T = 100$ K, and (vi) $T = 200$ K) are directly compared in Fig. 4a (see also Supplementary Materials Figs. S5–6 and Note 3). In all DUTs, we consistently observe the $1/f$ noise tendency. The current MFCC elements in the LF regime were considered to significantly contribute to learning and classification. Except for the graphene case (see (i) in Fig. 4a), the S_I/\bar{I}_D^2 curves for all 2D materials in this study fit well to the CNF-CMF model, implying the engineered current MFCCs of graphene would have a different image than those of the other

2D layered materials. The effects of e-beam irradiating the monolayer MoS₂ FET on the SiO₂ substrate are compared in schematic (ii) and (iii) of Fig. 4a, b. The obtained N_{it} increases by a factor of 10 after electron beam irradiation. Moreover, the engineered current MFCC for a_{SC} as a function of T in the monolayer MoS₂ FET on h-BN is also demonstrated (see (iv) to (vi) in Fig. 4a, b). a_{SC} increases with increasing T from 3.23×10^4 ($T = 25$ K) to 3.08×10^5 V s C⁻¹ ($T = 200$ K)^{11,36}.

The frequency distributions are presented in a histogram with 20 intervals, as shown in Fig. 4a, using the normalized elements of LF MFCC of classes (i)–(vi) (see Fig. 4b). As N_{it} increases from condition (ii) to (iii), the highest frequency of the histogram shifts to the positive direction. A similar positive frequency shift is observed in cases (iv)–(vi) with the increasing T . Referring to Eq. (4), the S_I varies as a function of N_{it} and a_{SC} , and the corresponding current MFCCs can be extracted via featurization, consequently enabling the representation of a specific histogram tendency.

The HMM algorithm, which learns considering the correlation between the previous state and the next state, progresses under the following two learning conditions. The first learning condition is that the specific current fluctuation of each device in a specific class is due to N_{it} and a_{SC} , and the current MFCC contributes to learning by considering the above information. The second learning condition is that the HMM algorithm is learned by considering the correlation between the MFCC of the previous frequency and the MFCC of the next frequency. Thus, in Eq. (4), the exponent γ , which reflects the trap distribution, also influences the learning process with the HMM

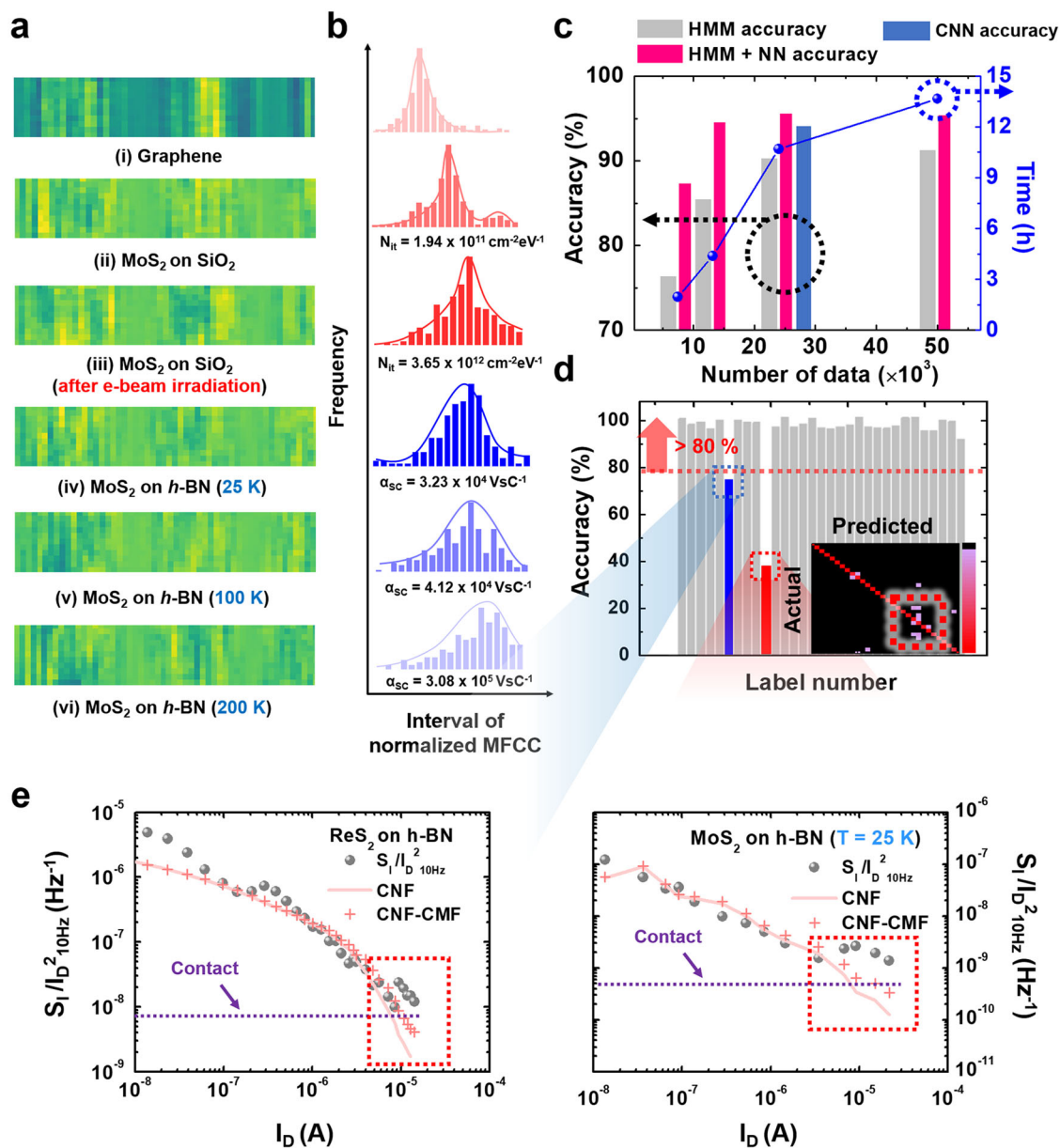


Fig. 4 Engineered current MFCC study and classification accuracy. **a** LF part of current MFCC (the darker color, the smaller value) converted at the same $I_D \approx 1 \mu\text{A}$ in each class ((i) graphene on trench structure, (ii) MoS₂ on SiO₂, (iii) MoS₂ on SiO₂ after e-beam irradiation, and MoS₂ on h-BN at (iv) $T = 25 \text{ K}$, (v) $T = 100 \text{ K}$, and (vi) $T = 200 \text{ K}$, respectively); **b** frequency distribution for a specific interval by normalizing the LF part of current MFCC; **c** classification accuracies and processing time obtained using the three learning and classification methods based on the number of data; **d** classification accuracies according to each class (the inset shows a confusion matrix and the darker color, the smaller value); **e** variations in the measured normalized drain current spectral densities for ReS₂ and MoS₂ on h-BN at $f = 10 \text{ Hz}$ and data fitted using the CNF and CNF-CMF models.

algorithm, with N_{it} and α_{sc} . Figure 4c displays the classification accuracies and processing time obtained using the HMM algorithm and the HMM score vector learning method employing the NN for a number of data. When the number of data was 7800, the HMM classification accuracy was 76.3% with f1-score and AUC value of <0.78 using fourfold cross-validation (see Supplementary Materials Fig. S3 and Note 3.11)^{30,59,60}. The HMM+NN classification accuracy was 85.2% with f1-score of 0.86, AUC value of 0.83, and processing time of 3 h for each fourfold cross-validation. For 22,100 data points, the HMM+NN classification accuracy increased to 95.5% with f1-score of 0.93, AUC value of 0.91, and processing time of 11 h for each fourfold cross-validation. However, for 48,800 data points, classification accuracies exhibited no further improvement, and only the processing time increased to 15 h. Moreover, the classification

accuracy learned by the convolution neural network (CNN) algorithm using current MFCCs as image data not through HMM architecture reached 93.6% with f1-score of 0.82, AUC value of 0.87 as good as the performance of HMM+NN. (see Supplementary Materials Fig. S7 and Note 4). On the other hand, the logistic regression model achieved only a classification accuracy of 88.8% with f1-score of 0.75, AUC value of 0.89. Therefore, provided that the performance of CNN architecture for classifying by learning perceptrons of each layer is acceptable, a transfer learning for any other channel or gate oxide materials can be possible⁶¹.

Most of the classes (or labels) were in good agreement with the CNF-CMF model with high averaged cross-validation accuracies of over 90% with f1-score of over 0.86 and AUC value of over 0.84, as presented in Fig. 4d. However, two exceptional classes, i.e., ReS₂

(blue bar) and MoS₂ (red bar) FETs fabricated on h-BN, are present in this figure with low classification accuracies of 74.2% (with f1-score of 0.79 and AUC of 0.75) and 38% (with f1-score of 0.49 and AUC of 0.51). This indicates that the current MFCCs for these classes were misinterpreted in the high current region. To interpret this miscalculation clearly, the corresponding S_I/\bar{I}_D^2 curves at $f=10$ Hz for both cases are displayed in Fig. 4e. Although they are well fitted to the CNF-CMF model in most of the current regions, the additional contact resistance (R_{CT}) contributing towards the total LF noise behavior in the high current regions curtails the accuracies in particular¹¹. The inset in Fig. 4d shows the confusion matrix of the HMM+NN architecture. Interestingly, some classes, which should consider the effects on additional contact resistance such as ReS₂ and MoS₂ FETs using h-BN as gate dielectric, WSe₂ FETs using Au as contact metal, and monolayered MoS₂ FETs, are sometimes confused with each other.

DISCUSSION

Combining the LF noise spectroscopy with machine learning algorithms provides an efficient and precise approach to characterize and classify 2D layered FETs. Through the use of an NN based on the hidden Markov model algorithm, we demonstrate that MFCCs, which were converted from the LF noise data of DUTs, can be predicted more precisely than the limits of fundamental measurements. Importantly, this method of applying only a specific voltage can be considered advantageous in both classifying device information and characterization of device performances. The combination of factors such as channel material, gate dielectric, contact metal, and electron beam irradiation have a profound effect on carrier fluctuations, enabling effective learning and training. Further, the learning models using LF noise spectroscopy presented herein are highly interpretable, and aid in identifying how engineered features, including the behaviors between carriers and traps, contribute to characterizing device information and performance. Therefore, the considerable flexibility of this approach makes it adaptable in distinguishing the degree of degradation and reliability of device and to modeling optimized fabrication conditions and device structures. The carrier transport direction, stacking order and orientation in 2D heterostructures would be a critical factor that influences significantly on charge fluctuation, expecting to enable the improved interpretation in the future via this approach. Moreover, the inference of engineered current MFCC features that currently lack sufficient noise data, combined with the CNF-CMF and additional contact noise approaches, and an improved ability to build models from limited experimental data should be possible using the developed model.

METHODS

Sample fabrication

An appropriately selected chemical vapor deposited monolayer MoS₂ and mechanically exfoliated 2D multilayer materials such as MoS₂, BP, ReS₂, MoTe₂, WSe₂, and h-BN were transferred onto high-quality 300 nm-thick SiO₂/p⁺-Si substrates. To make source and drain metal electrodes on them, standard electron beam lithography was used, and 80 nm-thick Au, Ti, Pt, and Cr were deposited using an electron-beam evaporation system. To suppress the contact resistance effect at the metal-semiconductor interface, all the fabricated devices were annealed under a high vacuum condition for 2 h at 473 K. The trenched graphene FETs in this study was fabricated on a pre-patterned parallel grid structure made of spin-coated poly(Methyl Methacrylate) A2 via conventional dry transfer methods³⁹. The Al₂O₃ passivation layer was deposited on 2D materials using an atomic layer deposition system.

In-situ measurement with e-beam irradiation

Electron-beam irradiation was conducted under high vacuum conditions ($\sim 10^{-6}$ Torr) at 300 K using a scanning electron microscope (SEM) (Quanta 3D FEG) chamber with a nano-manipulator for multilayer MoS₂, WSe₂, and

monolayer MoS₂ for 30 s with 30 kV and 50 pA. Four tungsten probes installed on the nano-manipulator system were electrically connected to a semiconductor parameter analyzer.

Electrical transport measurement

All the devices, except the Al₂O₃ passivated MoS₂ FETs, were characterized in a high vacuum-probe station system¹⁰. Fundamental electrical transport characterizations were performed using semiconductor analyzers (Keithley 4200, Agilent B1500A) with a temperature controllable probe system (335, Lake-Shore). Low-frequency noise characteristics were obtained from a home-made noise measurement system (the system details are presented in Fig. S2 in the Supplementary Materials), consisting of a home-made battery box, a low noise current-to-voltage pre-amplifier (SR570, Stanford Research Systems), and a data acquisition system (DAQ-4431, National Instruments)⁴².

Data processing and training

We used the Python speech features library in Github (https://github.com/jameslyons/python_speech_features) for processing of LF noise data into MFCC parameters. We only considered data where \bar{I}_D was larger than 100 nA to avoid a possible error. The optimized combination of hyperparameters was based on the previous studied LF noise analysis, narrowed the range, and found the best result by iterating through for loop. After Augmenting training MFCCs dataset using Gaussian noise, we used `hmmlearn` (<https://github.com/hmmlearn/hmmlearn>) library in Github for using HMM trainer function with training MFCCs data. Through HMM training, trained data generated for each class were converted into score vectors, and these vectors were trained by neural network based on the Tensorflow keras (<https://www.tensorflow.org/guide/keras>). Finally, We learned and trained current MFCC data directly using CNN also based on the Tensorflow keras (<https://www.tensorflow.org/guide/keras>).

Model validation

We used the 4-fold cross-validation method to train our MFCCs dataset, training MFCCs dataset was divided into 4 subsets having equal sizes randomly. Of the 4 subsets, a single subset was retained for the test data for evaluating the model, and the remaining three subsets were used as training. Our cross-validation process is repeated 4 times, with each of the four subsets used once for test. The remained test MFCC datasets were converted into score vectors to evaluate the model with training data learned through the HMM+NN architecture. We obtained not only the accuracy, but also confusion matrix, receiver operating characteristic (ROC) curves, area under the curve (AUC) value, and f1-score to evaluate the model performance accurately with imbalance of the data (<https://scikit-learn.org/stable/modules/classes.html#module-sklearn.metrics>).

DATA AVAILABILITY

Some of LF noise data that support the findings of this study are available from Github (<https://github.com/Kookjin-Lee/Kookjin.Sangjin.noiseML>) and the test LF noise data are uploaded in the subfolder with each label name in the folder (Test data_noise). All LF noise data are available from the corresponding author(s) upon reasonable request.

CODE AVAILABILITY

All the codes used to train, evaluate, calculate the presented results in this study are available from Github (<https://github.com/Kookjin-Lee/Kookjin.Sangjin.noiseML>) using the available Python file (LFnoise_ML_Classification_SJ_KJ.py).

Received: 12 August 2020; Accepted: 18 November 2020;

Published online: 04 January 2021

REFERENCES

1. Kirton, M. J. & Uren, M. J. Noise in solid-state microstructures: a new perspective on individual defects, interface states and low-frequency (1/f) noise. *Adv. Phys.* **38**, 367–468 (1989).
2. Rogers, C. T. & Buhrman, R. A. Nature of single-localized-electron states derived from tunneling measurements. *Phys. Rev. Lett.* **55**, 859 (1985).

3. Lee, K. et al. Understanding of aging pattern in quantum dot light-emitting diodes by low-frequency noise. *Nanoscale* **12**, 15888–15895 (2020).
4. Rogers, C. T. & Buhrman, R. A. Composition of 1/f noise in metal-insulator-metal tunnel junctions. *Phys. Rev. Lett.* **53**, 1272 (1984).
5. Schroder, D. K. *Semiconductor Material and Device Characterization* (John Wiley & Sons, Ulsan, 2006).
6. Hooge, F. N. 1/f noise sources. *IEEE Trans. Electron Devices* **41**, 1926–1935 (1994).
7. Song, S. H., Joo, M. K., Neumann, M., Kim, H. & Lee, Y. H. Probing defect dynamics in monolayer MoS₂ via noise nanospectroscopy. *Nat. Commun.* **8**, 1–5 (2017).
8. Balandin, A. A. Low-frequency 1/f noise in graphene devices. *Nat. Nanotechnol.* **8**, 549–555 (2013).
9. Von Haartman, M. & Östling, M. *Low-Frequency Noise in Advanced MOS Devices* (Springer Science & Business Media, 2007).
10. Na, J. et al. Low-frequency noise in multilayer MoS₂ field-effect transistors: the effect of high-k passivation. *Nanoscale* **6**, 433–441 (2014).
11. Joo, M. K. et al. Understanding coulomb scattering mechanism in monolayer MoS₂ channel in the presence of h-BN buffer layer. *ACS Appl. Mater. Interfaces* **9**, 5006–5013 (2017).
12. Xu, G. et al. Effect of spatial charge inhomogeneity on 1/f noise behavior in graphene. *Nano Lett.* **10**, 3312–3317 (2010).
13. Na, J. et al. Few-layer black phosphorus field-effect transistors with reduced current fluctuation. *ACS Nano* **8**, 11753–11762 (2014).
14. Mitra, R., Jariwala, B. & Bhattacharya, A. Probing in-plane anisotropy in few-layer ReS₂ using low frequency noise measurement. *Nanotechnology* **29**, 145706 (2018).
15. Ji, H. et al. Thickness-dependent carrier mobility of ambipolar MoTe₂: interplay between interface trap and Coulomb scattering. *Appl. Phys. Lett.* **110**, 183501 (2017).
16. Esteva, A. et al. A guide to deep learning in healthcare. *Nat. Med.* **25**, 24–29 (2019).
17. Gopnik, A. Making AI more human. *Sci. Am.* **316**, 60–65 (2017).
18. Lawrence, S., Giles, C. L., Tsoi, A. C. & Back, A. D. Face recognition: a convolutional neural-network approach. *IEEE Trans. Neural Netw.* **8**, 98–113 (1997).
19. de Albuquerque, V. H. C., Cortez, P. C., de Alexandria, A. R. & Tavares, J. M. R. S. A new solution for automatic microstructures analysis from images based on a backpropagation artificial neural network. *Nondestruct. Test. Eval.* **23**, 273–283 (2008).
20. Vignal, C., Mathevon, N. & Mottin, S. Audience drives male songbird response to partner's voice. *Nature* **430**, 448–451 (2004).
21. Butler, K. T. et al. Machine learning for molecular and materials science. *Nature* **559**, 547–555 (2018).
22. Lee, K. et al. Detection and Accurate Classification of Mixed Gases using Machine Learning with Impedance Data. *Adv. Theory Simul.* **3**, 2000012 (2020).
23. Ghahramani, Z. Probabilistic machine learning and artificial intelligence. *Nature* **521**, 452–459 (2015).
24. Jordan, M. I. & Mitchell, T. M. Machine learning: trends, perspectives, and prospects. *Science* **349**, 255–260 (2015).
25. Kireeva, N. et al. Generative topographic mapping (GTM): universal tool for data visualization, structure-activity modeling and dataset comparison. *Mol. Inform.* **31**, 301–312 (2012).
26. Mortazavi, B. et al. Exploring phononic properties of two-dimensional materials using machine learning interatomic potentials. *Appl. Mater. Today* **20**, 100685 (2020).
27. Das, S. et al. Machine learning in materials modeling—fundamentals and the opportunities in 2D materials. *Synth. Model. Charact. 2D Mater. their Heterostruct.* 445–468 <https://doi.org/10.1016/B978-0-12-818475-2.00019-2>. (2020).
28. Brown, K. A., Brittan, S., Maccaferri, N., Jariwala, D. & Celano, U. Machine learning in nanoscience: big data at small scales. *Nano Lett.* **20**, 2–10 (2020).
29. Mortazavi, B. et al. Machine-learning interatomic potentials enable first-principles multiscale modeling of lattice thermal conductivity in graphene/borophene heterostructures. *Mater. Horiz.* **7**, 2359–2367 (2020).
30. Borovcnik, M., Bentz, H.-J. & Kapadia, R. A *Probabilistic Perspective*. *Chance Encounters: Probability in Education* (Springer, Dordrecht, 1991).
31. Fine, S., Singer, Y. & Tishby, N. The hierarchical hidden Markov model: analysis and applications. *Mach. Learn.* **32**, 41–62 (1998).
32. Wu, J. & Yu, J. An improved arithmetic of MFCC in speech recognition system. *2011 Int. Conf. Electron. Commun. Control. ICECC 2011-Proc.* 719–722 (2011).
33. Shao, X. & Milner, B. MAP prediction of pitch from mfcc vectors for speech reconstruction. *8th Int. Conf. Spok. Lang. Process. ICSLP 2004* **1**, 2425–2428 (2004).
34. Ji, H. et al. Suppression of Interfacial Current Fluctuation in MoTe₂ Transistors with Different Dielectrics. *ACS Appl. Mater. Interfaces* **8**, 19092–19099 (2016).
35. Ko, S. P. et al. Low frequency noise reduction in multilayer WSe₂ field effect transistors. *IEEE-NANO 2015 - 15th Int. Conf. Nanotechnol.* 1118–1121 (2015).
36. Joo, M. et al. Electron excess doping and effective schottky barrier reduction on the MoS₂/h-BN heterostructure. *Nano Lett.* **16**, 6383–6389 (2016).
37. Cho, Y. H. et al. Soft-type trap-induced degradation of MoS₂ field effect transistors. *Nanotechnology* **29**, 1–8 (2018).
38. Jin, J. E. et al. Catalytic etching of monolayer graphene at low temperature via carbon oxidation. *Phys. Chem. Chem. Phys.* **18**, 101–109 (2016).
39. Jin, J. E. et al. Surface modulation of graphene field effect transistors on periodic trench structure. *ACS Appl. Mater. Interfaces* **8**, 18513–18518 (2016).
40. Ryu, M. Y. et al. Triethanolamine doped multilayer MoS₂ field effect transistors. *Phys. Chem. Chem. Phys.* **19**, 13133–13139 (2017).
41. Lee, K. et al. Real-time effect of electron beam on MoS₂ field-effect transistors. *Nanotechnology* **31**, 455202 (2020).
42. Joo, M. K. et al. A dual analyzer for real-time impedance and noise spectroscopy of nanoscale devices. *Rev. Sci. Instrum.* **82**, 034702 (2011).
43. Sigurdsson, S., Petersen, K. B. & Lehn-Schiöler, T. Mel frequency cepstral coefficients: An evaluation of robustness of MP3 encoded music. *ISMIR 2006 - 7th Int. Conf. Music Inf. Retr.* 286–289 (2006).
44. Hasan, R., Jamil, M., Rabbani, G. & Rahman, S. Speaker Identification Using Mel Frequency Cepstral Coefficients. *3rd Int. Conf. Electr. Comput. Eng. ICECE 2004* 28–30 (2004).
45. Farooq, O. & Datta, S. Mel filter-like admissible wavelet packet structure for speech recognition. *IEEE Signal Process. Lett.* **8**, 196–198 (2001).
46. Molau, S., Pitz, M., Schlüter, R. & Ney, H. Computing mel-frequency cepstral coefficients on the power spectrum. *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.* **1**, 73–76 (2001).
47. Churi, A., Bhat, A., Mohite, R. & Churi, P. P. E-zip: An electronic lock for secured system. *2016 IEEE Int. Conf. Adv. Electron. Commun. Comput. Technol. ICAECC 2016* **2**, 45–49 (2017).
48. Beal, M. J. et al. The Infinite Hidden Markov Model. *NIPS* **14**, 577–584 (2012).
49. Schuller, B. et al. Hidden Markov model-based speech emotion recognition. *Proc. - IEEE Int. Conf. Multimed. Expo.* **1**, 1401–1404 (2003).
50. Sonnhammer, E. L., von Heijne, G. & Krogh, A. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **6**, 175–182 (1998).
51. Zhang, J. & Sclaroff, S. Saliency detection: A boolean map approach. *Proc. IEEE Int. Conf. Comput. Vis.* 153–160 (2013).
52. Rahman, M. A. & Hoque, M. A. Online adaptive artificial neural network based vector control of permanent magnet synchronous motors. *IEEE Power Eng. Rev.* **17**, 28 (1997).
53. Lin, C. F. & Wang, S. De. Fuzzy support vector machines. *IEEE Trans. Neural Netw.* **13**, 464–471 (2002).
54. Agarap, A. F. Deep Learning using Rectified Linear Units (ReLU). *arxiv* 2–8. Preprint at <https://arxiv.org/abs/1803.08375> (2018).
55. Hara, K., Saito, D. & Shouno, H. Analysis of function of rectified linear unit used in deep learning. *Proc. Int. Jt. Conf. Neural Networks* 1–8 (2015).
56. Ghibaudo, G., Roux, O., Nguyen-Duc, C., Balestra, F. & Brini, J. Improved analysis of low frequency noise in field-effect MOS transistors. *Phys. Status Solidi* **124**, 571–581 (1991).
57. Juvela, L. et al. Speech waveform synthesis from MFCC sequences with generative adversarial networks. *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process* 5679–5683 (2018).
58. Jensen, J. H. et al. Evaluation of MFCC estimation techniques for music similarity. *European Signal Processing Conference* 2219–2491 (2006).
59. Schaffer, C. Technical Note: Selecting a Classification Method by Cross-Validation. *Mach. Learn.* **13**, 135–143 (1993).
60. Moore, A. W. & Lee, M. S. *Efficient Algorithms for Minimizing Cross Validation Error*. *Machine Learning Proceedings 1994* (Morgan Kaufmann Publishers, Inc., 1994).
61. Deepak, S. & Ameer, P. M. Brain tumor classification using deep CNN features via transfer learning. *Comput. Biol. Med.* **111**, 103345 (2019).

ACKNOWLEDGEMENTS

This research was supported by Nano-Material Technology Development Program through the National Research Foundation of Korea (NRF) funded by Ministry of Science and ICT and also supported by the Future Semiconductor Device Technology Development Program funded by Ministry of Trade, Industry & Energy (MOTIE) and Korea Semiconductor Research Consortium (KSRC) (NRF-2017M3A7B4049119 & Grant 10067739, G.-T.K.) supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation). Further, M.-K.J. also wishes to acknowledge the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (NRF-2019R1C1C1003467 & NRF-2019K2A9A1A06083674). We appreciate Hyunwoo J. Kim for helping us design and develop the code.

AUTHOR CONTRIBUTIONS

K.L. and S.N. initiated and developed the work under the supervision of M.-K.J. and G.-T.K. K.L., and S.N. designed and developed the code, and conducted the analysis. K.L., H.J., J.C., J.-E.J., Y.K., J.N., M.-Y.R., Y.-H.C., H.L., J.L., and M.-K.J. fabricated the samples and measured low-frequency noise data. K.L., S.N., M.-K.J., and G.-T.K. wrote the manuscript and discussed with all the authors.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at <https://doi.org/10.1038/s41699-020-00186-w>.

Correspondence and requests for materials should be addressed to M.-K.J. or G.-T.K.

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

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



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

© The Author(s) 2021