Carbon Nanotube Sensor Array for the Label-Free Discrimination of Live and Dead Cells with Machine Learning

*Zhengru Liu,[†] Galina V. Shurin,[‡] Long Bian,[†] David White,[†] Michael R. Shurin,[‡§] and Alexander Star[†||]**

[†]Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

[‡]Department of Pathology, University of Pittsburgh Medical Center, 3550 Terrace Street, Pittsburgh, Pennsylvania 15261, USA

[§]Department of Immunology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania 15213, USA

[||]Department of Bioengineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA

Keywords: carbon nanotubes, field effect transistors, sensor array, machine learning, sensing mechanism, cell death

Abstract

Developing robust cell recognition strategies is important in biochemical research, but the lack of well-defined target molecules creates a bottleneck in some applications. In this

paper, a carbon nanotube sensor array was constructed for the label-free discrimination of live and dead mammal cells. Three types of carbon nanotube field-effect transistors (NTFET) were fabricated, and different features were extracted from the transfer characteristic curves for model training with linear discriminant analysis (LDA) and support-vector machine (SVM). Live and dead cells were accurately classified in more than 90% of samples in each sensor group using LDA as the algorithm. The recursive feature elimination with cross-validation (RFECV) method was applied to handle the overfitting and optimize the model, and cells could be successfully classified with as few as four features and higher validation accuracy (up to 97.9%) after model optimization. The RFECV method also revealed the crucial features in the classification, indicating the participation of different sensing mechanisms in the classification. Finally, the optimized LDA model was applied for the prediction of unknown samples with accuracy of 87.5% to 93.8%, indicating that live and dead cell samples could be well recognized with the constructed model.

Introduction

Carbon nanotubes (CNTs) have been widely used for the detection of biomolecules at trace levels and for whole-cell sensing by electrochemical[1-4] and optical[5-6] techniques. However, the extra labeling steps are needed for those cell recognition strategies which are primarily based on detection of specific target biomolecules. In addition, the cell recognition could also be hindered by the complexity of biological matrix and lack of well-defined target biomolecule. The introduction of machine learning methods provides new opportunities for the cell recognition applications.[7-9] Specifically, CNT-based electronic noses/tongues, which are based on sensor arrays fabricated from chemically functionalized CNTs, have shown the great potential in sensing applications with the help of machine learning.[10-13] By analyzing the changes in electronic properties of different CNT sensors in the array, the fingerprints of target analytes can be drawn, which are classified by statistical analysis using machine

learning algorithms such as linear discriminant analysis (LDA)[8, 14-15] and support-vector machine (SVM).[14, 16-17] Because of the high sensitivities, rapid responses and high integrations, the CNT-based electronic noses/tongues have been applied to the label-free detection of many analytes such as volatile organic compounds (VOCs),[11, 13, 18] odors from different food items,[19] and metal ions.[20] Liu et al.[11] developed the method for the classification of different VOCs with principal component analysis (PCA) and LDA. Different metalloporphyrin−CNT hybrid materials were selected for the fabrication of chemiresistor array. Schroeder et al.[19] constructed the chemiresistor sensor array for the classification of cheese, liquor, and edible oil based on their odor using a k-nearest neighbors model and a random forest model. The accuracy of prediction can be well improved by model training process. However, the loss of useful information in data collection may be a challenge for cell sensing applications using CNT-based electronic noses/tongues. Large amount of useful information was discarded in the feature extraction process as only few electronic properties, such as resistance and capacitance changes, were extracted as features in model training. We contend that using more features could potentially increase accuracy in whole cell recognition and provide invaluable information about the sensing mechanism when considering the feature importance, which in turn could provide a guidance for the design of new sensors for other applications.

Sensor arrays comprised carbon nanotube-based field-effect transistors (NTFET) have shown advantages in the number of features, when compared to other CNT device configurations. By measuring the transfer characteristics of NTFET, many parameters, such as source-drain current values at different voltages, transconductance, on/off-current ratio, and threshold voltage, can be extracted as features in model training, thus providing more useful information compared with chemiresistor- or capacitor-based methods. Moreover, the changes of transfer characteristics can provide useful information about the sensing mechanism, especially the interactions between the CNTs and biomolecules. In our previous works, we identified 11 different features from the NTFET characteristics for model training and realized the classification of nonmalignant and malignant cells with LDA.[21] The same strategy was also applied for the classification of different purine compounds.[22] From the

feature elimination process in model training, we evaluated dominant mechanisms in purine compounds sensing, which was further collaborated by DFT calculations.

In this work, we aim to develop the label-free strategy for the discrimination of live and dead cells. Although classic methods, such as cell staining with different dyes, have been widely used for dead cell recognition in routine biological tasks, it is still essential to develop non-optical methods for the discrimination of live and dead cells. Compared with live cells, dead cells exhibit changes in morphology, permeability, and metabolism, which we expect to affect their interactions with CNTs, resulting in different NTFET characteristics. For the construction of sensor array in this work, we chose three types of functional single-walled carbon nanotubes (SWCNTs). Commercial semiconductor enriched SWCNTs were decorated with gold nanoparticles, and gold nanoparticles were functionalized with self-assembled monolayers (SAM) of dodecanethiol (DD) and 11-mercaptoundecanoic acid (MUA) because of their respective hydrophobic and hydrophilic properties. We have collected 15 features for the model training with LDA and linear SVM for each sensor group. Our training model showed good validation accuracy (up to 91.7%) for the classification. Then, recursive feature elimination with cross-validation method was applied for the evaluation of different features. The optimized model showed that live and dead cells could be classified with only 4 to 5 features and higher validation accuracy (up to 97.9%). The optimized LDA model was also applied for the prediction of unknown samples with accuracy of 87.5% to 93.8%. Finally, we could conclude that the classification highly relied on several key features, and the several sensing mechanisms dominated in the classification.

Materials and Methods

Commercial semiconductor-enriched single-walled carbon nanotubes (IsoSol-S100, NanoIntegris) with 0.1% metallic and 99.9% semiconducting composition were used for the fabrication of sensors in this work. 1 mM chloroauric acid ($HAuCl_4 \cdot 3H_2O$, Alfa Aesar) was prepared in 0.1 M HCl solution for the gold nanoparticles decorations. 10 mM dodecanethiol

(Sigma-Aldrich) and 10 mM 11-mercaptoundecanoic acid (Sigma-Aldrich) were prepared in water/ethanol mixture for the functionalization of carbon nanotubes. Doxorubicin (Sigma-Aldrich) was used for the induction of cell death in the experiment.

Field-Effect Transistor Fabrication. The interdigitated gold electrodes were patterned on the 2 × 2 mm^2 Si/SiO_2 wafers using photolithography. The chips were then wire-bonded into standard 40-pin ceramic dual inline packages and secured with polydimethylsiloxane (PDMS) by heating at 80 ºC for 1 h. Subsequently, 3 μL of IsoSol-SWCNTs (0.02 g mL^{-1}, dispersed in toluene) were deposited between the gold electrodes through dielectrophoresis (DEP) using a Keithley 3390 Arbitrary Waveform (10 V_{PP}, 100 kHz for 2 min), then annealed at 120 ºC for 24 h in the oven to evaporate the organic solvent.

The gold nanoparticles were deposited on the carbon nanotubes by bulk electrolysis of $HAuCl_4$ solution using a 3-electrodes system. Ag/AgCl electrode/Pt electrode/carbon nanotubes were employed as reference/counter/working electrode, respectively. The electrolysis was finished at –0.2 V for 30 s. The gold nanoparticles were characterized by a scanning electron microscopy (ZEISS sigma500 VP) with an accelerating voltage of 3 kV. Dodecanethiol and 11-mercaptoundecanoic acid were decorated on gold nanoparticles as self-assembled monolayers of thiols. 200 μL of dodecanethiol solution (10 mM, $V_{ethanol}/V_{water}$ = 1:1) was incubated on the gold nanoparticles for 2 h. 11-mercaptoundecanoic acid solution (10 mM, $V_{ethanol}/V_{water}$ = 7:3) was incubated on the gold nanoparticles for 24 h (protected by ethanol bath). After the functionalization, the devices were rinsed with ethanol and deionized water consecutively. The devices decorated with 11-mercaptoundecanoic acid decoration were kept in the vacuum for another 1 h to remove the ethanol remained on the surface.

Electrical measurements. Two different Keithley SourceMeter units, 2400 and 2602B, were employed for electrical measurements. NTFET characteristics were measured in phosphorate buffer (pH = 7.4) with an Ag/AgCl electrode as the liquid gate electrode, and the gate voltage (V_g) was swept from +0.6 V to –0.6 V with a source-drain voltage (V_{sd}) of 0.05 V.

Cell Sensing on Sensor Array. The scheme of whole-cell detection on the fabricated liquid gated NTFET is shown in Figure 1. Mouse B16 melanoma cells were prepared for cell sensing. The apoptosis was induced with the addition of 10 μM of doxorubicin for 24 h. Both cell culture conditions and the procedure for induction of apoptosis are described in supporting information. MTT tests was also employed for the evaluation of cell death (Supporting Information). Before the cell sensing test, the devices were immersed in 0.01 M PBS to obtain constant NTFET characteristics. Then the devices were immersed in 400 μL of PBS for 1 h and NTFET characteristics were measured as the control group. Subsequently, the devices were incubated with 400 μL of cell suspension (10^6 cells mL^{-1}) for another 1 h. After the incubation, the cell suspension was discarded, and NTFET characteristics were measured as the test group. For each measurement, the device was rinsed with deionized water and immersed in another 400 μL of PBS for 2 min before data collection. The same protocol was employed for both live/dead cell sensing on the devices with different decorations.

Cell Discrimination with Machine Learning. The same protocol was employed for features extraction as previously reported.[21-23] The NTFET characteristics of each sample from test group and control group were utilized for the calculation features. The 15 features shown in Figure S1 were calculated according to the equations shown in Table 1 using Python 3.7. Linear discriminant analysis (LDA) and support-vector machine (SVM) were utilized as algorithms for the classification of live and dead cells using Python. The Past 3 software was utilized for the model visualization after LDA. Recursive feature elimination with cross-validation (k=5) method was applied for the elimination of the features.

Results and Discussion

Decoration of CNTs. Carbon nanotubes were deposited between the interdigitated gold electrodes using dielectrophoresis. Scanning electron microscopy (SEM) images revealed that carbon nanotubes were aligned between the fingers of interdigitated gold

electrodes and formed interconnected networks (Figure 2). Due to the heterogeneity of SWCNTs suspension, the carbon nanotubes were not uniformly distributed among different electrodes. In addition, we found that the amount of carbon nanotubes deposited should be well controlled to guarantee the high on/off ratio of NTFET characteristics. Gold nanoparticles decorated both SWCNTs and the surface of gold electrodes with a narrow size distribution (μ=73.8 nm, σ=15.3 nm, calculated using ImageJ software). The NTFET characteristics of bare SWCNT and Au-SWCNT are depicted in Figure S2. NTFET transfer characteristics are consistent with p-type semiconducting behavior of SWCNT. The devices became more conductive after the gold nanoparticles deposition. The statistics of the maximum of source-drain current in NTFET characteristics before and after gold nanoparticles deposition are shown in Figure S3a and S3c. The data collected from 375 devices shows that the maximum of drain current increased by 151% on average after gold nanoparticles deposition, resulting in higher transconductance. The minimum of source-drain current also increased on average from 0.93 μA to 9.3 μA after deposition with gold nanoparticles (Figure S3b and S3d).

Dodecanethiol (DD) and 11-mercaptoundecanoic acid (MUA) were selected because of their similar molecular size but different surface properties of the resulting self-assembled monolayers (SAM) on the gold surface.[24-26] The terminal alkyl chain of DD yields hydrophobic SAM, while SAM produced from MUA are hydrophilic due to the terminal carboxyl groups. Different protocols were chosen for the decoration of these two thiol molecules. The rapid kinetics of the self-assembled monolayers of thiols on gold substrates have been reported,[27] which was consistent with the our results in DD decoration. The conductance of devices dramatically decreased after incubation with DD for 2 h. However, we found the conductance only slightly decreased with the MUA incubation, indicating longer time to form SAM. Herein, 24 h was finally employed as the time span for the MUA decoration. The NTFET characteristics indicated the decrease of conductance after decoration. As shown in Figure S2, the conductance values decreased by about 90% after DD decoration, and about 60% after MUA decoration. Meanwhile, the shift of threshold voltage to the

negative direction was also observed, indicating that the carbon nanotubes were n-doped after SAM formation on gold nanoparticles.

Induction of cell death. We chose doxorubicin, a chemotherapy drug widely used in the induction of apoptosis, for the cell treatment in this project. The experiment started from 2 million cells, and the percentage of dead cells was evaluated by cell counting with trypan blue staining. As shown in Tables S1 and S2, the percentage of dead cells increased with the increase of the concentration of doxorubicin after 24 h induction. At the same time, less cells were observed after the treatment, indicating that the cell proliferation was inhibited by doxorubicin. After 48 h, the percentage of dead cells was around 60% in the three groups treated with different concentration of doxorubicin, and even less cells were observed. These results were confirmed by the data from MTT assay. As shown in Figure S4, the decrease of adsorbance in the experimental group indicated the high apoptotic rate, indicating that at least 60% of cells were dead after 10 μM of doxorubicin treatment for 24 h in both experimental groups. In addition, we also observed the obvious morphological changes of cells after 10 μM treatment for 24 h. Therefore, 10 μM of doxorubicin and 24 h treatment were considered effecitive for the cell death induction, which were employed in the following experiments.

Cell Sensing: NTFET Characteristics and Sensing Mechanisms. The NTFET characteristics were measured both before and after incubation with live cells on three different types of devices with gold nanoparticle-decorated SWCNTs before (i) and after functionalization with dodecanethiol (DD) (ii) and 11-mercaptoundecanoic acid (MUA) (iii) (Figure 3a-c, labeled as Au-NTFET, DD-NTFET and MUA-NTFET in the following discussion). Because incubation with cells was performed in phosphate buffer saline (PBS), the influence of PBS on the devices was also investigated. The NTFET characteristics of these three types of devices remained constant after incubation in PBS for 1 h (Figure S5). After cell incubation, the source-drain conductance values of the devices commonly decreased by 10% to 40% (when measured at –0.5 V gate voltage). The differences in the conductance decrease might be related to the batch-to-batch variations in different cell

suspensions. Meanwhile, the negative shift of threshold voltage could be observed on all the three types of sensors, indicating that the SWCNTs were n-doped after cell incubation. The permanent conductance decreases, and threshold voltage shifts after cell incubation showed that the irreversible non-specific cell adhesion might play a key role in the cell sensing. The cells could be adsorbed on the devices after incubation, which has been demonstrated by fluorescence microscopy in previous work.[21] The cell adsorption might change the NTFET characteristics in different ways. Firstly, the cell adhesion on the device might change the channel resistance, leading to the conductance changes. Secondly, the ions/biomolecules released from the absorbed live/dead cells might interact with CNT hybrids, which was shown as the doping behavior in the NTFET characteristics. In addition, the transconductance, which is defined by the slope value of the linear region in the transfer NTFET characteristic curve, also decreased, indicating that the change of NTFET characteristics might also be related to other mechanisms such as Schottky barrier modulation and carrier mobility changes.[28]

Similar NTFET characteristics were observed after dead cells incubation (Figure 3d-f). Since doxorubicin was applied for the induction of cell death, we also addressed the influence of doxorubicin residue. The NTFET characteristics of all three types of devices were measured after incubation with different concentrations of doxorubicin (10^{-9} M to 10^{-5} M). The results shown in Figure S6 indicated that doxorubicin almost had not influence on the NTFET characteristics of MUA-NTFET even at high concentrations. For Au-NTFET and DD-NTFET, the signal change could be obviously observed from 10^{-8} M and reached a plateau at 10^{-5} M, which might be because of the saturation of doxorubicin adsorption on the surface of chip. The signal changes indicated that rinsing step was essential to eliminate the influence of doxorubicin by diluting doxorubicin to 10^{-9} M. Since rinsing steps were repeated for several times in our protocol, the doxorubicin could be effectively removed from the cell suspension. The assumption was well supported by the NTFET characteristics after incubation with dead cells (Figure 3d-f), which didn't show an extra conductance decrease compared with the NTFET characteristics after incubation with live cells (Figure 3a-c). In

conclusion, the doxorubicin could be effectively removed after rinsing steps and is not associated with the change of NTFET characteristics.

The choice of cell concentration was also investigated in the protocol. The NTFET characteristics after incubation with different concentrations of cells (10^3 to 10^6 mL^{-1}) were measured, and conductance change at –0.5 V gate voltage was applied for the discussion of signal changes. The results in Figure S7 showed that conductance decrease could be observed even with 10^4 mL^{-1} of cells, which is consistent with our previous results where we found that the adsorption of several hundred cells might lead to the change of NTFET characteristics.[21] However, the signal changes were not uniform across all device types with cell suspensions at low concentrations (for example, 10^4 mL^{-1} of dead cells in DD-NTFET in Figure S7e and 10^5 mL^{-1} of live cells in MUA-NTFET in Figure S7c). To guarantee that enough cells could be adsorbed on the chip surface for the interaction with carbon nanotubes, we selected 10^6 mL^{-1} as the concentration for the cell incubation. The influence of different percentages of dead cells was also investigated by obtaining different percentages of dead cells after treating with different concentrations of doxorubicin. Features F1, F2, and F15 were selected for the comparison of NTFET characteristics due to their significance in the classification (*vide infra*). As shown in Figure S8, there is no trend for the change of these three features with the increase of percentage of dead cells. These similarities showed that the discrimination of live/dead cells could not be achieved by a single feature in the NTFET characteristics, indicating the necessity of machine learning method.

Classification of live/dead cells. To construct the data matrix, we have picked 15 different features from the transfer NTFET characteristic curves, which are summarized in Figure S1. Compared with our previous studies,[21-23] another 4 features related to the conductance changes at ± 0.5 V and ± 0.55 V were included because of the rapid conductance change in the characteristic curves. The choice of features was based on the changes of physical properties of transistors, such as the changes of conductance values at different voltages, shift of threshold voltage and transconductance change in the linear region, which

were related to different sensing mechanisms such as doping mechanism and Schottky barrier modulation.[28-32] The data range of each feature value is given in the Figure S9-S11 and the feature values of each sample in different models are provided in Supporting Material 2. Firstly, the classification of live and dead cells with each type of sensor was investigated based on all the 15 features. Both linear discriminant analysis (LDA) and linear support-vector machine (SVM) were applied as the algorithms for the classification with 15 features in each sensor group (Au-NTFET: 30 dead cell samples and 30 live cell samples; DD-NTFET: 24 dead cell samples and 24 live cell samples; MUA-NTFET: 30 dead cell samples and 30 live cell samples). The Python codes for model training are provided in the supporting information. As shown in Table S3 and Figure 4 (the canonical scores of samples in different LDA models are given in Supporting Material 3), good training accuracy values (85.4% to 100%) could be obtained in all the three LDA models and two SVM models (except MUA-NTFET with the training accuracy of 69.1%). Good validation accuracy values could also be obtained in all the three groups (90.0% in Au-NTFET, 91.7% in DD-NTFET, and 90.0% in MUA-NTFET) after model training with LDA, suggesting that live and dead cells could potentially be well classified with a LDA model. The normalized confusion matrices are presented in Table S4, which are calculated with the confusion matrices of five test sets in the cross validation (k=5). The validation accuracy values of the training models with linear SVM were much lower (78.3% in Au-NTFET, 72.9% in DD-NTFET, and 60.0% in MUA-NTFET, shown in Table S5). We have also explored the combinations of two different sensor groups (labeled as Au-NTFET+DD-NTFET, Au-NTFET+MUA-NTFET and DD-NTFET+MUA-NTFET in the following discussion, including 15×2=30 features) for the classification. For the LDA models, the results showed that the validation accuracy decreased in all the combinations of two sensor groups (79.2% in Au-NTFET+DD-NTFET, 88.3% in Au-NTFET+MUA-NTFET, and 88.3% in DD-NTFET+MUA-NTFET). The reason could be the statistical distributions of the features in sensor combinations are not consistent with each other regarding the classes. The validation accuracy was marginally improved after combinations of sensor groups in the SVM models (79.2% in Au-NTFET+DD-NTFET, 80.0%

in Au-NTFET+MUA-NTFET, and 75.0% in DD-NTFET+MUA-NTFET). However, the validation accuracy was much lower than the training accuracy after combinations of sensor groups in both LDA and SVM models (100% in all the LDA models, and 87.5% to 96.4% in the SVM models, Table S3), indicating that overfitting might happen with more features.

It should be noticed that some features in the above models are highly correlated, which could cause overfitting problems, and undermine the model's predictiability.[33] The correlations between different features were evaluated by Pearson correlation coefficient values, which are shown in the heatmaps (Figure S12-S14). We have found that features F4, F6, F8, F10, and F12 were highly correlated to each other in all the three sensor groups. It might be because all these 5 features were related to conductance changes in the linear region, which were also based on similar equations in the calculations of feature values. For the same reasons, features F5, F7, F9, F11, and F13 were highly correlated since all these features were related to the off-current changes at positive voltages in NTFET transfer characteristics, which is indicative of the sensing mechanisms based on doping mechanism and Schottky barrier modulation.[28] Features F4, F6, F8, F10 and F12 were negatively correlated to F5, F7, F9, F11, and F13 in the Au-NTFET, but relatively uncorrelated in DD-NTFET and MUA-NTFET. On the other hand, features F3 and F14 were also highly correlated to each other because the minimum of conductance was equal to conductance at 0.6 V in some samples (Figure S1). For the other three features, F2 was uncorrelated to other features in Au-NTFET but highly correlated in the other two groups. It should be noticed that features F1 and F15 were highly correlated in both Au-NTFET and DD-NTFET but the correlations with other features were still different. Herein, these two features were still discussed separately in the following discussion. One-way ANOVA was also run for the comparison of feature values between live and dead cells (Table S6). The results showed the means of features F1, F10 and F12 were statistically significantly different between live and dead cell samples in all the three sensor types. Otherwise, for several features (F9, F13 and F15), the distinct p-values among the three groups should be noticed, indicating that there is no statistically significant

difference for these three features between live and dead cells in several groups, and different features may dominate the classification in different sensor groups.

To explore the which features are essential in the cell classification, the recursive feature elimination with cross-validation (RFECV) method was introduced for the selection of features, which eliminated features based on a weighting scheme with the weights being initially assigned to the features from an estimator.[22] We performed RFECV on both LDA and SVM models. For the LDA models, the results in Figure S15 and Table 1 showed that 5 to 11 of 15 features were needed for the classification with each sensor group, and 7 to 23 of 30 features were needed for the sensor group combinations. The validation accuracy shown in Table 2 was improved with both single sensor group (up to 95.8% in DD-NTFET) and sensor group combinations (up to 97.9 % in Au-NTFET+DD-NTFET). Among these features, F1, F2, F12, and F15 participated in the classifications in all the combinations. These four features also represent 4 different parameters in the transfer characteristics, which are transconductance change (F1), threshold voltage shift (F2), on-current change (F15), and off-current change (F12). Conversely, features F3 and F14 never appeared in any combinations, showing that these two features might be least significant for the cell classification using a LDA model. The validation accuracy was also improved in both single sensor group (up to 80% in Au-NTFET with 4 features) and sensor group combinations (up to 87.5% in Au-NTFET+DD-NTFET with 9 features) in SVM models (Figure S16, Table S7-S8), which were also satisfactory for the classification. For the SVM models, it should be noticed that the classification was still not satisfactory for the MUA-NTFET (only 68.3% with Feature 1) after RFECV process. Features F1, F5, and F13 appeared in most combinations using SVM models, which were mainly related to transconductance change and on-current change. If we only considered about the two training models with the best validation accuracy values in SVM models (Au-NTFET+DD-NTFET and Au-NTFET+MUA-NTFET), we found that the off-current change still played an important role in the classification. The main difference between LDA models and SVM models is the participation of threshold voltage shift.

We also noticed that several highly correlated features were present in the optimized models after RFECV. The results in Figure S15 also showed that a high validation accuracy (higher than 90%) could still be obtained with even less features. Herein, the model was further simplified with less features by declining the correlated features. We chose the three single sensor groups in the LDA models for the simplification. For the highly correlated features which also have similar correlations with the other features, only the feature with lowest p-value in the Table S6 would be kept in model training. The final feature combinations for each sensor group were shown in Table S9. Good validation accuracy could be kept in both Au-NTFET (93.3% with 5 features, Table S10) and DD-NTFET (97.9% with 4 features, which was even higher than the combinations recommended by RFECV). The validation accuracy decreased to 85% in MUA-NTFET with only 5 features, which might explain why 11 features were needed in the optimized model after RFECV. In summary, the validation accuracy value was still satisfactory after model simplification in the three sensor groups, indicating that live and dead cells could be well classified with a simple LDA model using each type of sensor.

Prediction of unknown samples with optimized model. Although our optimized model showed good accuracy in the classification, the validation of model with unknown samples is still important to evaluate the model in the applications. Herein, an external validation set was constructed for the evaluation of the optimized model after RFECV method (the models with selected features are shown in Table 1 and Table S7, and the feature values of each sample are provided in Supporting Material 4), which contained 8 replicates of both live and dead cell samples for each type of NTFET. As shown in Table 3, the unknown samples were well predicted with the optimized LDA models (87.5% for Au-NTFET, 93.8% for DD-NTFET, and 87.5% for MUA-NTFET). The SVM models were also applied for the predictions. Compared with the validation accuracy of the optimized SVM model in Table S8, the prediction accuracy was higher for DD-NTFET (87.5% vs 79.2%), and almost identical for MUA-NTFET (68.8% versus 68.3%). A decrease of prediction accuracy was observed for

Au-NTFET (62.5% versus 80%) with SVM model, which was mainly due to the incorrect prediction of live cell samples. It should be noticed that the validation accuracy of optimized SVM model is at most 80% (Table S8), which may undermine its credibility in the prediction. The LDA model still showed higher accuracy in the prediction of unknown samples compared with SVM model, which was consistent with the conclusion from the training models. In summary, the prediction of unknown samples was still satisfactory with the optimized LDA model, showing that LDA might be more feasible for the classification of live and dead cells in this work.

Conclusion

In this work, a carbon nanotube field-effect transistor (NTFET) sensor array was constructed for the label-free whole-cell sensing. The NTFET characteristics in cell sensing showed the participation of different sensing mechanisms, such as doping and Schottky barrier modulations. With the extraction of features from the transfer characteristics, two different algorithms, linear discriminant analysis (LDA) and support-vector machine (SVM), were applied for the classification of live and dead cells. Live and dead cells were well classified with a good validation accuracy higher than 90% in each sensor group using a LDA model, and the accuracy could be further improved in each sensor group and sensor group combinations using the recursive feature elimination (RFECV) with cross-validation method in both LDA and SVM models. The RFECV results indicated that cell classification was highly relied on several crucial features. These crucial features reveled that the transconductance and on-current change were greatly significant in the classification. In addition, the threshold voltage shift and off-current change also played important roles in the classification, suggesting the participation of several different mechanisms. The simplified model showed that live and dead cells could be well classified with only 4 to 5 features using a LDA model, and live and dead cells as unknown samples could be well predicted. However, we still need to point out that there are more challenges for the applications in real samples

compared with this proof-of-concept study. The cell number may vary in a real sample, which may also be a mixture of different types of cells. The complexity of biological matrix may also challenge the application of sensors in some cases. In addition, the improvement in sensor fabrication should be made to guarantee the repeatability in the test with the same type of sensors. A more general and robust model will be continuously explored in our further study, which is expected for the classification of more target cells.

Supporting Information

Protocols of cell experiments; Results of cell death induction experiments; NTFET characteristics of different devices in different texts; Feature extraction methods; Results of statistical analysis; LDA and SVM plots; Confusion matrices (pdf)

Datasets for the feature values of each sample in different training models (xlsx)

Canonical scores of samples in different LDA models (xlsx)

Datasets for the feature values of each unknown sample in external validation sets (xlsx)

Acknowledgements

This work is supported by the National Science Foundation under grant no. 2003302.

Author Information

Corresponding Author

Alexander Star — *Department of Chemistry and Department of Bioengineering, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States; orcid.org/0000-0001-7863-5987; Email: astar@pitt.edu*

Authors

Zhengru Liu — *Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States*

Galina V. Shurin — *Department of Pathology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania 15261, United States*

Long Bian — *Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States*

David L. White — *Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States*

Michael R. Shurin — *Departments of Pathology and Immunology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania 15261, United States; orcid.org/0000-0002-6570-7395*

References

1. Zhang, R.; Rejeeth, C.; Xu, W.; Zhu, C. Y.; Liu, X. Y.; Wan, J. J.; Jiang, M. W.; Qian, K., Label-Free Electrochemical Sensor for CD44 by Ligand-Protein Interaction. *Anal. Chem.* **2019,** *91* (11), 7078-7085.
2. Gulati, P.; Kaur, P.; Rajam, M. V.; Srivastava, T.; Mishra, P.; Islam, S. S., Vertically aligned multi-walled carbon nanotubes based flexible immunosensor for extreme low level detection of multidrug resistant leukemia cells. *Sens. Actuators, B* **2019,** *301*, 127047.
3. Li, J.; Jiang, M.; Su, M.; Tian, L.; Shi, W.; Yu, C., Stretchable and Transparent Electrochemical Sensor Based on Nanostructured Au on Carbon Nanotube Networks for Real-Time Analysis of H_2O_2 Release from Cells. *Anal. Chem.* **2021,** *93* (17), 6723-6730.
4. Kim, J.; Park, G.; Lee, S.; Hwang, S. W.; Min, N.; Lee, K. M., Single wall carbon nanotube electrode system capable of quantitative detection of $CD4^+$ T cells. *Biosens. Bioelectron.* **2017,** *90*, 238-244.
5. Dinarvand, M.; Neubert, E.; Meyer, D.; Selvaggio, G.; Mann, F. A.; Erpenbeck, L.; Kruss, S., Near-Infrared Imaging of Serotonin Release from Cells with Fluorescent Nanosensors. *Nano Lett.* **2019,** *19* (9), 6604-6611.
6. Budhathoki-Uprety, J.; Langenbacher, R. E.; Jena, P. V.; Roxbury, D.; Heller, D. A., A Carbon Nanotube Optical Sensor Reports Nuclear Entry via a Noncanonical Pathway. *ACS Nano* **2017,** *11* (4), 3875-3882.
7. Becher, B.; Schlitzer, A.; Chen, J. M.; Mair, F.; Sumatoh, H. R.; Teng, K. W. W.; Low, D.; Ruedl, C.; Riccardi-Castagnoli, P.; Poidinger, M.; Greter, M.; Ginhoux, F.; Newell, E. W., High-dimensional analysis of the murine myeloid cell system. *Nat. Immunol.* **2014,** *15* (12), 1181-1189.
8. Das Saha, N.; Sasmal, R.; Meethal, S. K.; Vats, S.; Gopinathan, P. V.; Jash, O.; Manjithaya, R.; Gagey-Eilstein, N.; Agasti, S. S., Multichannel DNA Sensor Array Fingerprints Cell States and Identifies Pharmacological Effectors of Catabolic Processes. *ACS Sens.* **2019,** *4* (12), 3124-3132.
9. Erzina, M.; Trelin, A.; Guselnikova, O.; Dvorankova, B.; Strnadova, K.; Perminova, A.; Ulbrich, P.; Mares, D.; Jerabek, V.; Elashnikov, R.; Svorcik, V.; Lyutakov, O., Precise cancer detection via the combination of functionalized SERS surfaces and convolutional neural network with independent inputs. *Sens. Actuators, B* **2020,** *308*, 127660.

10. Schroeder, V.; Savagatrup, S.; He, M.; Ling, S. B.; Swager, T. M., Carbon Nanotube Chemical Sensors. *Chem. Rev.* **2019,** *119* (1), 599-663.
11. Liu, S. F.; Moh, L. C. H.; Swager, T. M., Single-Walled Carbon Nanotube-Metalloporphyrin Chemiresistive Gas Sensor Arrays for Volatile Organic Compounds. *Chem. Mater.* **2015,** *27* (10), 3560-3563.
12. Schnorr, J. M.; van der Zwaag, D.; Walish, J. J.; Weizmann, Y.; Swager, T. M., Sensory Arrays of Covalently Functionalized Single-Walled Carbon Nanotubes for Explosive Detection. *Adv. Funct. Mater.* **2013,** *23* (42), 5285-5291.
13. Wang, F.; Swager, T. M., Diverse Chemiresistors Based upon Covalently Modified Multiwalled Carbon Nanotubes. *J. Am. Chem. Soc.* **2011,** *133* (29), 11181-11193.
14. Huang, C. H.; Zeng, C.; Wang, Y. C.; Peng, H. Y.; Lin, C. S.; Chang, C. J.; Yang, H. Y., A Study of Diagnostic Accuracy Using a Chemical Sensor Array and a Machine Learning Technique to Detect Lung Cancer. *Sensors* **2018,** *18* (9), 2845.
15. Wang, J.; Zhu, L. Y.; Zhang, W. L.; Wei, Z. B., Application of the voltammetric electronic tongue based on nanocomposite modified electrodes for identifying rice wines of different geographical origins. *Anal. Chim. Acta* **2019,** *1050*, 60-70.
16. Hunter, R.; Sohi, A. N.; Khatoon, Z.; Berthiaume, V. R.; Alarcon, E. I.; Godin, M.; Anis, H., Optofluidic label-free SERS platform for rapid bacteria detection in serum. *Sens. Actuators, B* **2019,** *300*, 126907.
17. Chen, Q.; Zhang, J. L., Classification and Recognition of Ovarian Cells Based on Two-Dimensional Light Scattering Technology. *J. Med. Syst.* **2019,** *43* (5), 127.
18. Hwang, S. I.; Franconi, N. G.; Rothfuss, M. A.; Bocan, K. N.; Bian, L.; White, D. L.; Burkert, S. C.; Euler, R. W.; Sopher, B. J.; Vinay, M. L.; Sejdic, E.; Star, A., Tetrahydrocannabinol Detection Using Semiconductor-Enriched Single-Walled Carbon Nanotube Chemiresistors. *ACS Sens.* **2019,** *4* (8), 2084-2093.
19. Schroeder, V.; Evans, E. D.; Wu, Y. C. M.; Voll, C. C. A.; McDonald, B. R.; Savagatrup, S.; Swager, T. M., Chemiresistive Sensor Array and Machine Learning Classification of Food. *ACS Sens.* **2019,** *4* (8), 2101-2108.
20. da Silva, G. S.; de Oliveira, L. P.; Costa, G. F.; Giordano, G. F.; Nicoliche, C. Y. N.; da Silva, A. A.; Khan, L. U.; da Silva, G. H.; Gobbi, A. L.; Silveira, J. V.; Souza, A. G.; Schleder, G. R.; Fazzio, A.; Martinez, D. S. T.; Lima, R. S., Ordinary microfluidic electrodes combined with bulk nanoprobe produce multidimensional electric double-layer capacitances towards metal ion recognition. *Sens. Actuators, B* **2020,** *305*, 127482.
21. Silva, G. O.; Michael, Z. P.; Bian, L.; Shurin, G. V.; Mulato, M.; Shurin, M. R.; Star, A., Nanoelectronic Discrimination of Nonmalignant and Malignant Cells Using Nanotube Field-Effect Transistors. *ACS Sens.* **2017,** *2* (8), 1128-1132.
22. Bian, L.; Sorescu, D. C.; Chen, L.; White, D. L.; Burkert, S. C.; Khalifa, Y.; Zhang, Z.; Sejdic, E.; Star, A., Machine-Learning Identification of the Sensing Descriptors Relevant in Molecular Interactions with Metal Nanoparticle-Decorated Nanotube Field-Effect Transistors. *ACS Appl. Mater. Inter.* **2019,** *11* (1), 1219-1227.
23. Bian, L.; Wang, Z.; White, D. L.; Star, A., Machine learning-assisted calibration of Hg^{2+} sensors based on carbon nanotube field-effect transistors. *Biosens. Bioelectron.* **2021,** *180*, 113085.
24. Zhao, X. L.; Cai, Y. Q.; Wang, T.; Shi, Y. L.; Jiang, G. B., Preparation of Alkanethiolate-Functionalized Core/Shell Fe_3O_4@Au Nanoparticles and Its Interaction with Several Typical Target Molecules. *Anal. Chem.* **2008,** *80* (23), 9091-9096.
25. Rasch, M. R.; Rossinyol, E.; Hueso, J. L.; Goodfellow, B. W.; Arbiol, J.; Korgel, B. A., Hydrophobic Gold Nanoparticle Self-Assembly with Phosphatidylcholine Lipid: Membrane-Loaded and Janus Vesicles. *Nano Lett.* **2010,** *10* (9), 3733-3739.
26. Wiseman, M. E.; Frank, C. W., Antibody Adsorption and Orientation on Hydrophobic Surfaces. *Langmuir* **2012,** *28* (3), 1765-1774.
27. Love, J. C.; Estroff, L. A.; Kriebel, J. K.; Nuzzo, R. G.; Whitesides, G. M., Self-Assembled Monolayers of Thiolates on Metals as a Form of Nanotechnology. *Chem. Rev.* **2005,** *105* (4), 1103-1170.

28. Heller, I.; Janssens, A. M.; Mannik, J.; Minot, E. D.; Lemay, S. G.; Dekker, C., Identifying the mechanism of biosensing with carbon nanotube transistors. *Nano Lett.* **2008,** *8* (2), 591-595.
29. Boussaad, S.; Tao, N. J.; Zhang, R.; Hopson, T.; Nagahara, L. A., In situ detection of cytochrome c adsorption with single walled carbon nanotube device. *Chem. Commun.* **2003,** (13), 1502-1503.
30. Chen, R. J.; Choi, H. C.; Bangsaruntip, S.; Yenilmez, E.; Tang, X. W.; Wang, Q.; Chang, Y. L.; Dai, H. J., An investigation of the mechanisms of electronic sensing of protein adsorption on carbon nanotube devices. *J. Am. Chem. Soc.* **2004,** *126* (5), 1563-1568.
31. Artyukhin, A. B.; Stadermann, M.; Friddle, R. W.; Stroeve, P.; Bakajin, O.; Noy, A., Controlled electrostatic gating of carbon nanotube FET devices. *Nano Lett.* **2006,** *6* (9), 2080-2085.
32. Gui, E. L.; Li, L. J.; Zhang, K. K.; Xu, Y. P.; Dong, X. C.; Ho, X. N.; Lee, P. S.; Kasim, J.; Shen, Z. X.; Rogers, J. A.; Mhaisalkar, S. G., DNA sensing by field-effect transistors based on networks of carbon nanotubes. *J. Am. Chem. Soc.* **2007,** *129* (46), 14427-14432.
33. Strieth-Kalthoff, F.; Sandfort, F.; Segler, M. H. S.; Glorius, F., Machine learning the ropes: principles, applications and directions in synthetic chemistry. *Chem. Soc. Rev.* **2020,** *49* (17), 6154-6168.

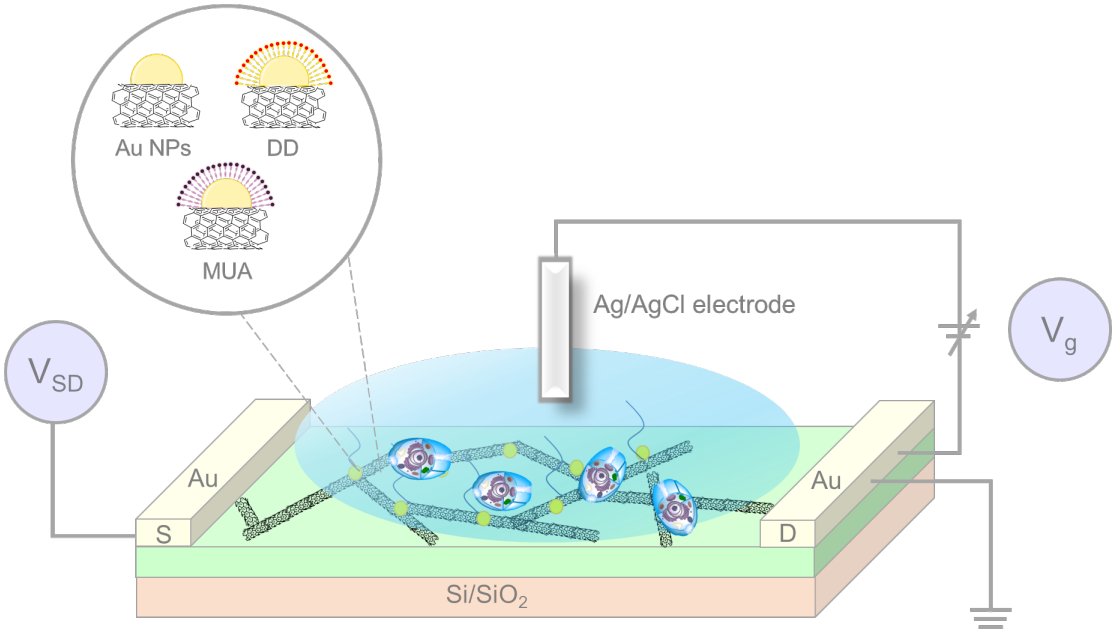

Figure 1 Schematic illustration of a liquid gated NTFET. A Ag/AgCl reference electrode was employed as liquid gate electrode (G) whereas interdigitated gold electrodes were employed as source (S) and drain (D) electrodes with a source–drain bias voltage as 50 mV. The inset shows the three diffrerent deocrations on the carbon nanotubes including bare gold nanoparticles, dodecanethiol (DD) and 11-mercaptoundecanoic acid (MUA).

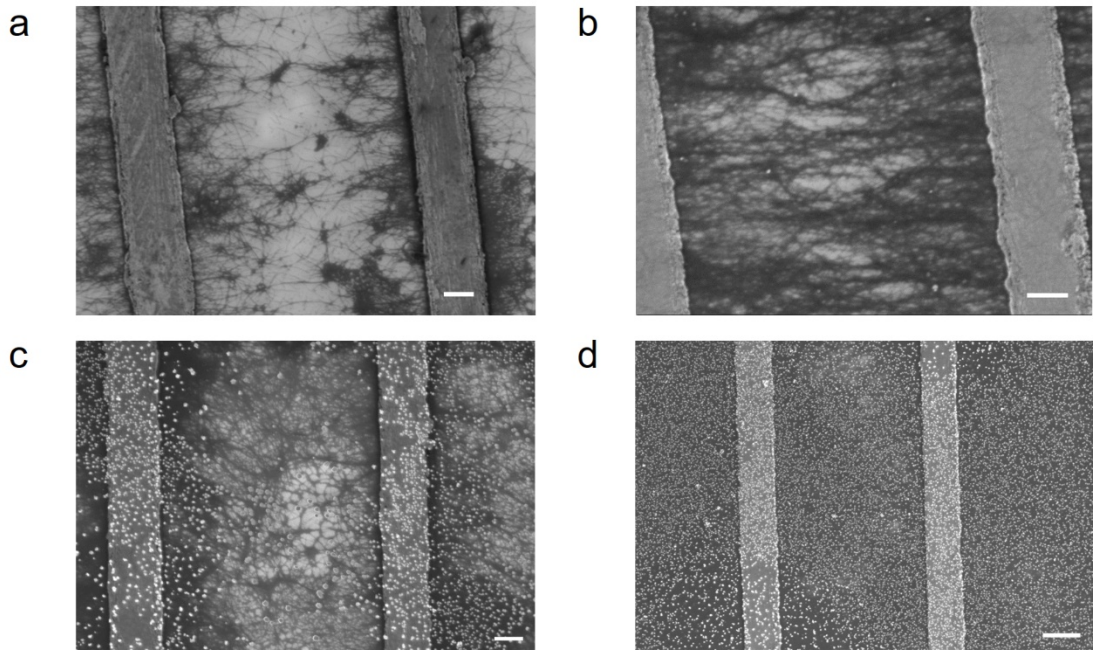

Figure 2 Scanning electron microscopy (SEM) images of SWCNTs hybrids. (a-b) SEM images of bare SWCNTs deposited between the interdigitated gold electrodes. (c-d) SEM images of gold nanoparticles decorated on SWCNTs. Scale bar is 1 μm for panel a-c and 2 μm for panel d.

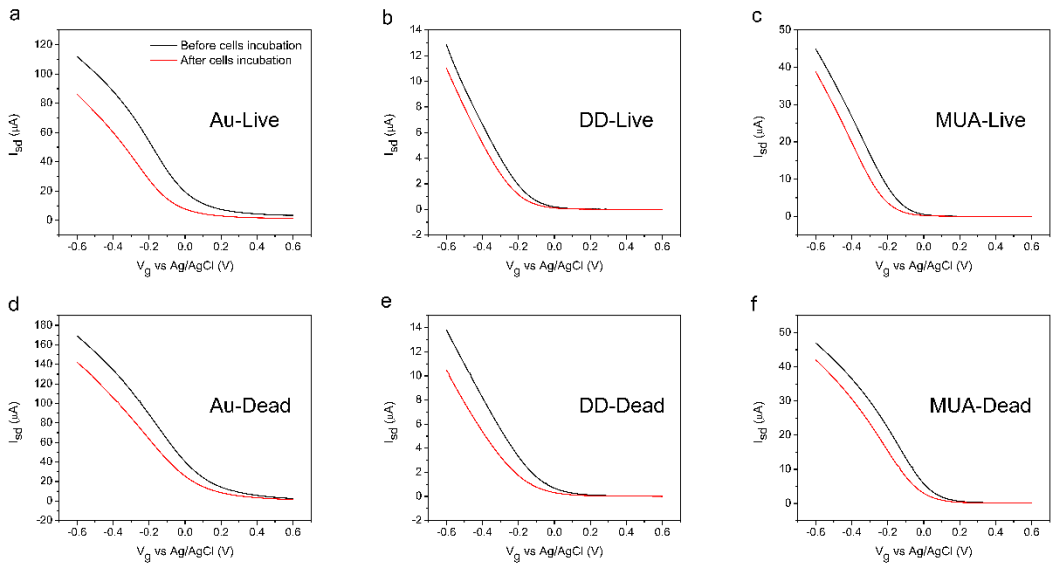

Figure 3 Transfer characteristics of NTFET, i.e, source-drain current (I_{sd}) versus applied liquid gate voltage, of NTFET with different decorations after cells incubation. (a-c) Transfer characteristics of NTFET with (a) bare gold nanoparticles, (b) dodecanethiol, and (c) 11-mercaptoundecanoic acid decoration after live cells incubation. (d-f) Transfer characteristics of NTFET with (d) bare gold nanoparticles, (e) dodecanethiol, and (f) 11-mercaptoundecanoic acid decoration after dead cells incubation. Devices were incubated with phosphate buffer for 1 h as the control and with cells for another 1 h before the measurement of transfer characteristics. Drain current was measured by sweeping gate voltage (V_g) from −0.6 V to 0.6 V. Source–drain bias voltage (V_{sd}) was 50 mV.

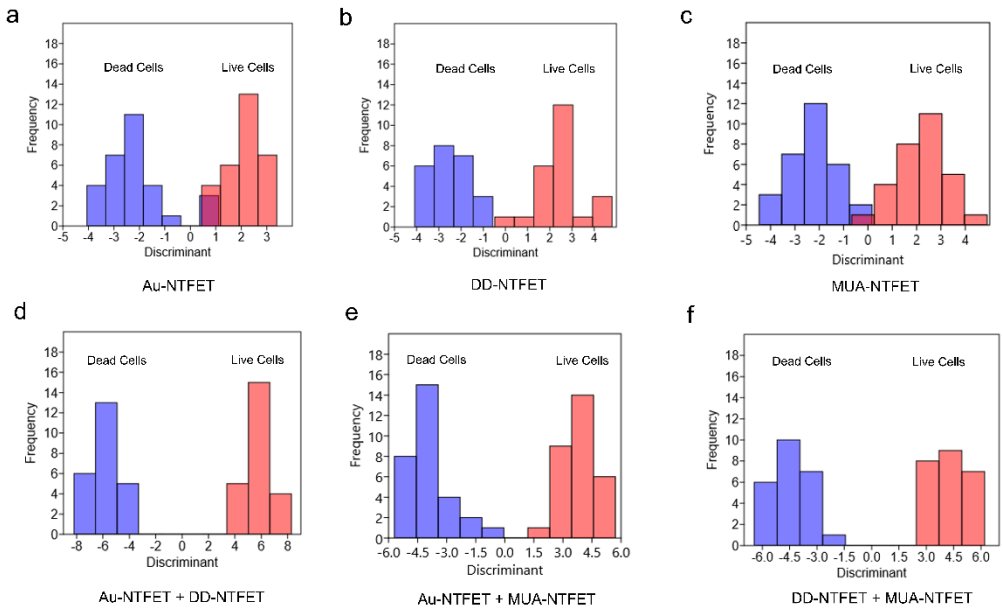

Figure 4 Visualization of training models of live/dead cells classification with diffrerent sensor groups and sensor group combinations. (a-c) Classification of live cells (red) and dead cells (blue) with (a) Au-NTFET, (b) DD-NTFET, and (c) MUA-NTFET. (d-f) Classification of live cells and dead cells with the combination of (d) Au-NTFET and DD-NTFET, (e) Au-NTFET and MUA-NTFET, and (f) DD-NTFET and MUA-NTFET. Linear discriminant analysis was applied as the algorithm for the model traning. 15 features were applied for each sensor group and thirty features (15×2) were applied for each sensor group combination

Table 1 Selection of Features and Prediction Accuray after Model Training with RFECV-LDA in Different Sensor Groups and Sensor Group Combinations. k=5 in the Cross-Validation Process.

NO.	Features	Au	DD	MUA	Au+DD	Au+MUA	DD+MUA
1	$(t_c-t_0)/t_0$	✓	✓	✓	✓✓	✓_	✓✓
2	$V_{th,c}-V_{th,0}$	✓	✓	✓	_✓	_✓	✓✓
3	$(G_{c,min}-G_{0,min})/G_{0,min}$						
4	$(G_c-G_0)/G_{@vth}$ (0.6V)				✓_		✓✓
5	$(G_c-G_0)/G_{@vth}$ (–0.6V)			✓			✓_
6	$(G_c-G_0)/G_{@vth}$ (0.55V)		✓	✓	_✓	_✓	✓✓
7	$(G_c-G_0)/G_{@vth}$ (–0.55V)			✓			✓✓
8	$(G_c-G_0)/G_{@vth}$ (0.5V)	✓		✓		_✓	_✓
9	$(G_c-G_0)/G_{@vth}$ (–0.5V)			✓			_✓
10	$(G_c-G_0)/G_{@vth}$ (0.4V)	✓			✓_	✓	✓✓
11	$(G_c-G_0)/G_{@vth}$ (–0.4V)			✓			✓✓
12	$(G_c-G_0)/G_{@vth}$ (0.3V)	✓	✓	✓	✓✓	_✓	✓✓
13	$(G_c-G_0)/G_{@vth}$ (–0.3V)			✓			✓✓
14	$(G_c-G_0)/G_0$ (0.6V)						
15	$(G_c-G_0)/G_0$ (–0.6V)	✓	✓	✓	✓✓	✓_	✓✓
Accuracy		93.3%	95.8%	95.0%	97.9%	96.7%	93.8%

Table 2 Normalized Confusion Matrices of Test Sets after Model Training with RFECV-LDA in Different Sensor Groups and Sensor Group Combinations. k=5 in the Cross-Validation Process.

	Dead Cells	Live Cells	Accuracy	Precision	Recall	F1 Score
Au-NTFET	0.9	0.1	0.933	0.964	0.9	0.931
	0.033	0.967				
DD-NTFET	0.958	0.042	0.958	0.958	0.958	0.958
	0.042	0.958				
MUA-NTFET	0.967	0.033	0.95	0.935	0.967	0.951
	0.067	0.933				
Au-NTFET+DD-NTFET	0.958	0.042	0.979	1	0.958	0.979
	0	1				
Au-NTFET+MUA-NTFET	0.933	0.067	0.967	1	0.933	0.966
	0	1				
DD-NTFET+MUA-NTFET	0.917	0.083	0.938	0.957	0.917	0.937
	0.042	0.958				

Table 3 Prediction Accuracy of Unknown Samples with the Optimzed Models after RFECV

	Dead Cells	Live Cells	Accuracy
Au-NTFET (LDA)	7	1	0.875
	1	7	
DD-NTFET (LDA)	8	0	0.938
	1	7	
MUA-NTFET (LDA)	6	2	0.875
	0	8	
Au-NTFET (SVM)	8	0	0.625
	6	2	
DD-NTFET (SVM)	7	1	0.875
	1	7	
MUA-NTFET (SVM)	7	1	0.688
	4	4	

TOC Graphic