

Machine learning-enhanced detection of minor radiation-induced defects in semiconductor materials using Raman spectroscopy

Cite as: J. Appl. Phys. **135**, 025701 (2024); doi: [10.1063/5.0179881](https://doi.org/10.1063/5.0179881)

Submitted: 5 October 2023 · Accepted: 8 December 2023 ·

Published Online: 8 January 2024



Jia Yi Chia,¹  Nuatawan Thamrongsiripak,²  Sornwit Thongphanit,²  and Noppadon Nuntawong^{1,a)} 

AFFILIATIONS

¹National Electronics and Computer Technology, National Science and Technology Development Agency, Pathum Thani 12120, Thailand

²Development and Service Section, Irradiation Center, Thailand Institute of Nuclear Technology (Public Organization), Nakhon Nayok 26120, Thailand

Note: This paper is part of the special topic, Defects in Semiconductors 2024.

a) Author to whom correspondence should be addressed: noppadon.nuntawong@nectec.or.th

ABSTRACT

Radiation damage in semiconductor materials is a crucial concern for electronic applications, especially in the fields of space, military, nuclear, and medical electronics. With the advancements in semiconductor fabrication techniques and the trend of miniaturization, the quality of semiconductor materials and their susceptibility to radiation-induced defects have become more important than ever. In this context, machine learning (ML) algorithms have emerged as a promising tool to study minor radiation-induced defects in semiconductor materials. In this study, we propose a sensitive non-destructive technique for investigating radiation-induced defects using multivariate statistical analyses combined with Raman spectroscopy. Raman spectroscopy is a contactless and non-destructive method widely used to characterize semiconductor materials and their defects. The multivariate statistical methods applied in analyzing the Raman spectra provide high sensitivity in detecting minor radiation-induced defects. The proposed technique was demonstrated by categorizing 100–500 kGy irradiated GaAs wafers into samples with low and high irradiation levels using linear discrimination analysis ML algorithms. Despite the high similarity in the obtained Raman spectra, the ML algorithms correctly predicted the blind testing samples, highlighting the effectiveness of ML in defect study. This study provides a promising approach for detecting minor radiation-induced defects in semiconductor materials and can be extended to other semiconductor materials and devices.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0179881>

I. INTRODUCTION

Radiation is a significant concern for electronic applications requiring a high level of reliability, such as space applications, military operations, nuclear facilities, and medical electronics. Studies of the formation mechanisms of radiation defects and their influence on semiconductor properties began around 1950 in connection with the availability of perfect single crystals.¹ Subsequently, the production of semiconductor devices, widely used in electronics, prompted extensive research into radiation damage in transistors and diodes.² In the 1970s, numerous studies were conducted to investigate radiation effects and radiation hardening on metal-oxide semiconductor (MOS) devices, charge-coupled devices

(CCDs), and optoelectronics.³ The introduction of complementary metal-oxide semiconductor (CMOS) devices in the modern fabrication techniques initiated active radiation effect studies as these techniques could introduce radiation damages as well.^{4,5}

As the semiconductor industry advances rapidly, there is an increasing need to study radiation-induced defects, combining their effects in new semiconductor materials and devices.^{6–9} The relentless trend of device miniaturization has driven developments across various technologies.^{10,11} As devices become smaller, the quality of semiconductor materials becomes more crucial than ever. Low-level radiation damage, often not considered seriously, could become a significant concern in future electronics. Machine learning (ML), an

14 January 2024 17:40:57

emerging field of algorithms and statistical models, could effectively address the study of these minor radiation defects.

ML has found wide application in the semiconductor industry, particularly, in wafer inspection and various fabrication processes, to improve manufacturing yields and predict device performance.^{12–17} In terms of radiation damage, numerous computational modeling works complement experimental observations in understanding and predicting this inherently complex phenomenon.^{18–21} To the best of our knowledge, there are no reports exploring the application of ML algorithms, particularly, multivariate statistical analyses combined with Raman spectroscopy, to establish a sensitive non-destructive technique in investigating radiation defects.

Raman spectroscopy, a vibrational spectroscopy method, has been long recognized as a versatile method to characterize semiconductor materials and crystal defects, providing valuable information about crystal quality, strain states, and sample composition in a nondestructive and contactless manner.^{22,23} For example, Raman spectroscopy has been used to probe radiation effects induced by gamma radiation,²⁴ protons,²⁵ neutrons,^{26,27} and other energetic ions.^{28–31} This is typically achieved by investigating changes in peak intensity, band position, and the full width at half maximum (FWHM) of a Raman band.³² However, limitations arise when the radiation-induced change in the Raman spectrum is too small to be detected using conventional peak analysis techniques, such as curve fitting and peak deconvolution. To address this, we apply multivariate statistical methods to analyze the acquired Raman spectra. The applied analysis method allows us to identify statistical correlations from extensive data without *a priori* assumptions and provides a level of sensitivity beyond conventional data analysis techniques. Notably, recent applications of ML methods in conjunction with Raman spectroscopy in the field of analytical sciences have enabled the extraction of chemical information with high accuracy, sensitivity, and selectivity.³³

II. MATERIALS AND METHODS

A. Experimental details

To demonstrate the proposed approach, semi-insulating gallium arsenide (SI-GaAs) was irradiated with an electron beam at absorbed doses of 100, 200, 300, 400, and 500 kGy, corresponding to electron fluences of 3.29×10^{15} , 6.59×10^{15} , 9.88×10^{15} , 1.32×10^{16} , and $1.65 \times 10^{16} \text{ cm}^{-2}$, respectively. GaAs was selected as the irradiation target due to its well-known high tolerance to radiation damage and its widespread use in radiation hardened devices.^{34–36} In this experiment, the $1 \times 1 \text{ cm}^2$ irradiated SI-GaAs samples were cleaved from the center area of a (100) VGF (vertical gradient freeze) grown undoped SI-GaAs single crystal wafer. The wafer was supplied from AXT-Tongmei with a diameter of $100.0 \pm 0.5 \text{ mm}$ and a thickness of $625 \pm 25 \mu\text{m}$. The electron beam was generated from a linear accelerator (MB10-50, Mevex Corporation Ltd) with a kinetic energy of 10 MeV. Due to its very low mass, the electron beam requires a very high energy to displace lattice atoms. Nevertheless, its radiation damage is less significant than that caused by neutrons and other ion sources because the energy transferred by an electron can only displace atoms without causing secondary displacements.³⁷

Following irradiation, the GaAs samples were characterized using an inVia confocal Raman microscope from Renishaw with a

laser excitation wavelength of 532 nm with a power of 4.3 mW and an exposure time of 10 s, employing a backscattering geometry. The laser light was focused on the sample surface and the scattered light was collected using a $50\times$ objective lens, resulting in a laser spot size of $1.3 \mu\text{m}$. Over 800 Raman spectra were collected for each irradiation dose and were processed using a ML algorithm developed in Python, a programming language supported by the Python Software Foundation.

B. Data analytics framework

A data analytics framework was developed to analyze the measured Raman spectra. As depicted in Fig. 1, the raw data were imported and processed before applying ML algorithms. The raw spectral data were extracted from the Raman spectroscopy surface mapping measurements, without considering acquisition position information. A total of 35 measurements, each comprising between 100 and 500 spectra, resulting in a cumulative 5900 spectra, were used in this study. The data were divided into training and testing sets, consisting of 29 measurements (5300 spectra) and six measurements (600 spectra), respectively.

Both datasets underwent several identical preprocessing steps, including interpolation, cosmic ray removal, outlier rejection, and baseline correction. In the initial step, all spectra were interpolated to achieve an equally spaced Raman shift frequency of 0.2 cm^{-1} between each data point. Cubic spline interpolation was selected due to its ability to eliminate background noise without requiring data smoothing³⁸ and its effectiveness in combining datasets with different sampling frequencies without sacrificing analytical accuracy.³⁹ For cosmic ray removal, the Whitaker–Hayes algorithm,⁴⁰ which identifies and removes anomalous spikes based on modified Z scores calculated from the once-differenced detrended spectrum, was employed. Next, the dataset of each mapping was projected into a high-dimensional space using principal component analysis (PCA), and the first five principal components were retained for subsequent outlier removal. Mahalanobis distance,⁴¹ measuring the distance of each spectrum from the center of the data cloud, was calculated for each measurement, and spectra outside the 95% confidence range were rejected. Subsequently, linear polynomial correction was applied to the spectra to remove the fluorescence background.

Concerning the deployment of the ML algorithm, PCA was performed to extract essential information from the training dataset and visualize the similarity among observations. The processed Raman spectra were organized in a $5300 \times 62\,500$ matrix containing signals recorded at Raman shift positions ranging from 150 to 1400 cm^{-1} , with a step of 0.2 cm^{-1} . Using scikit-learn,⁴² the input data were initially centered and then subjected to linear dimensionality reduction using singular value decomposition (SVD) to project it into a lower dimensional space, specifically its principal component (PC). Let B_{ij} represent the standardized dataset while X_{ij} and \bar{X}_j represents the matrices of input Raman spectra and the mean of the variable j , respectively,

$$B_{ij} = X_{ij} - \bar{X}_j, \quad (1)$$

$$T = BV, \quad (2)$$

14 January 2024 17:40:57

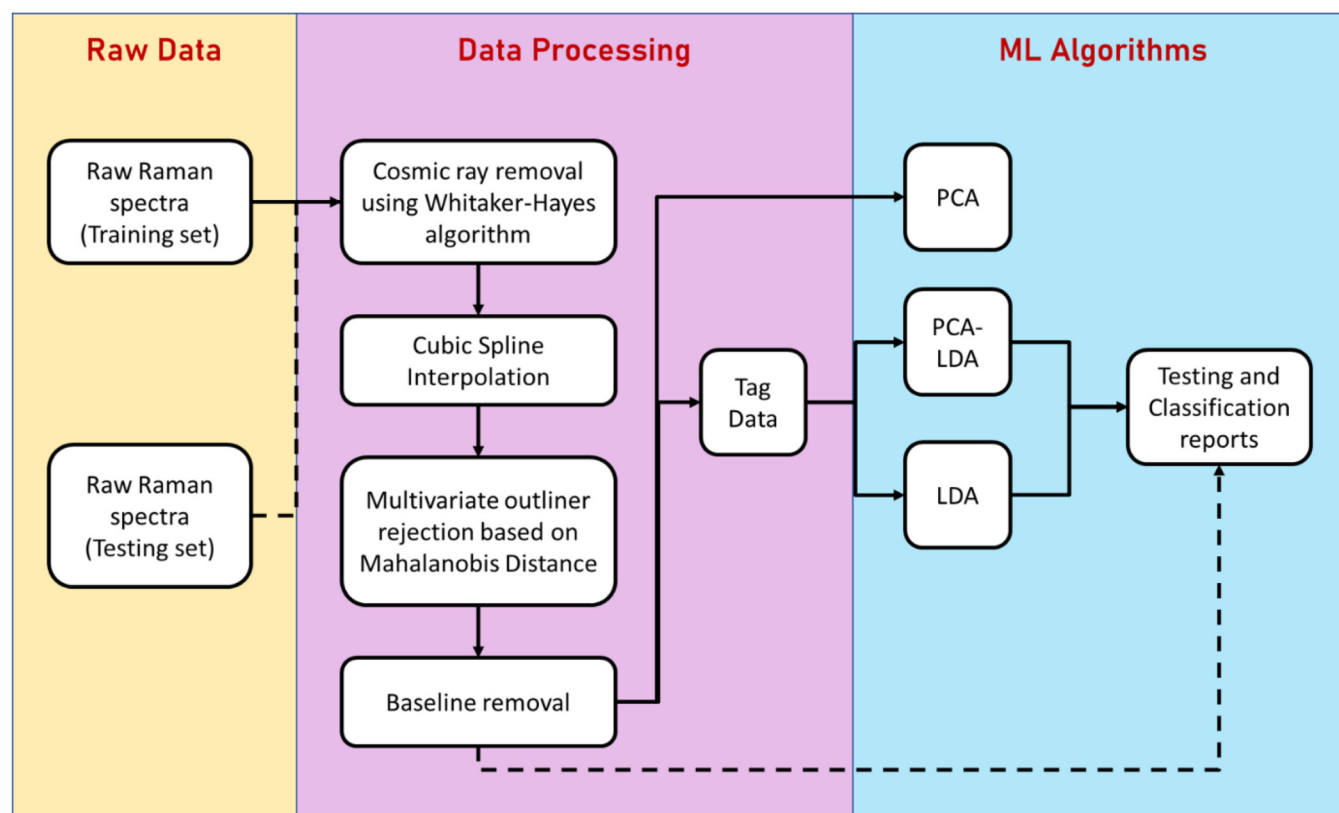


FIG. 1. The schematic of data analytic flow consists of three parts: raw data, data processing, and ML algorithms.

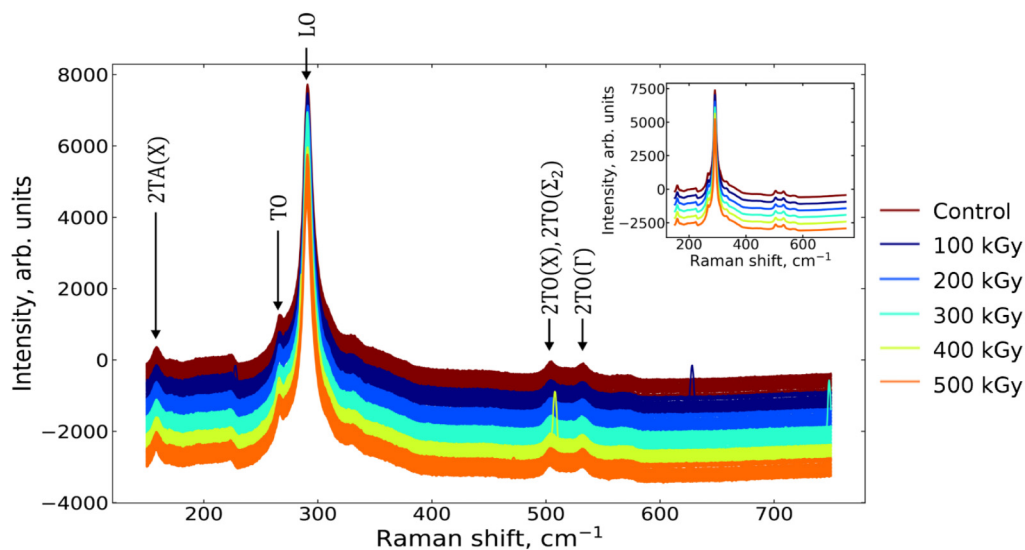


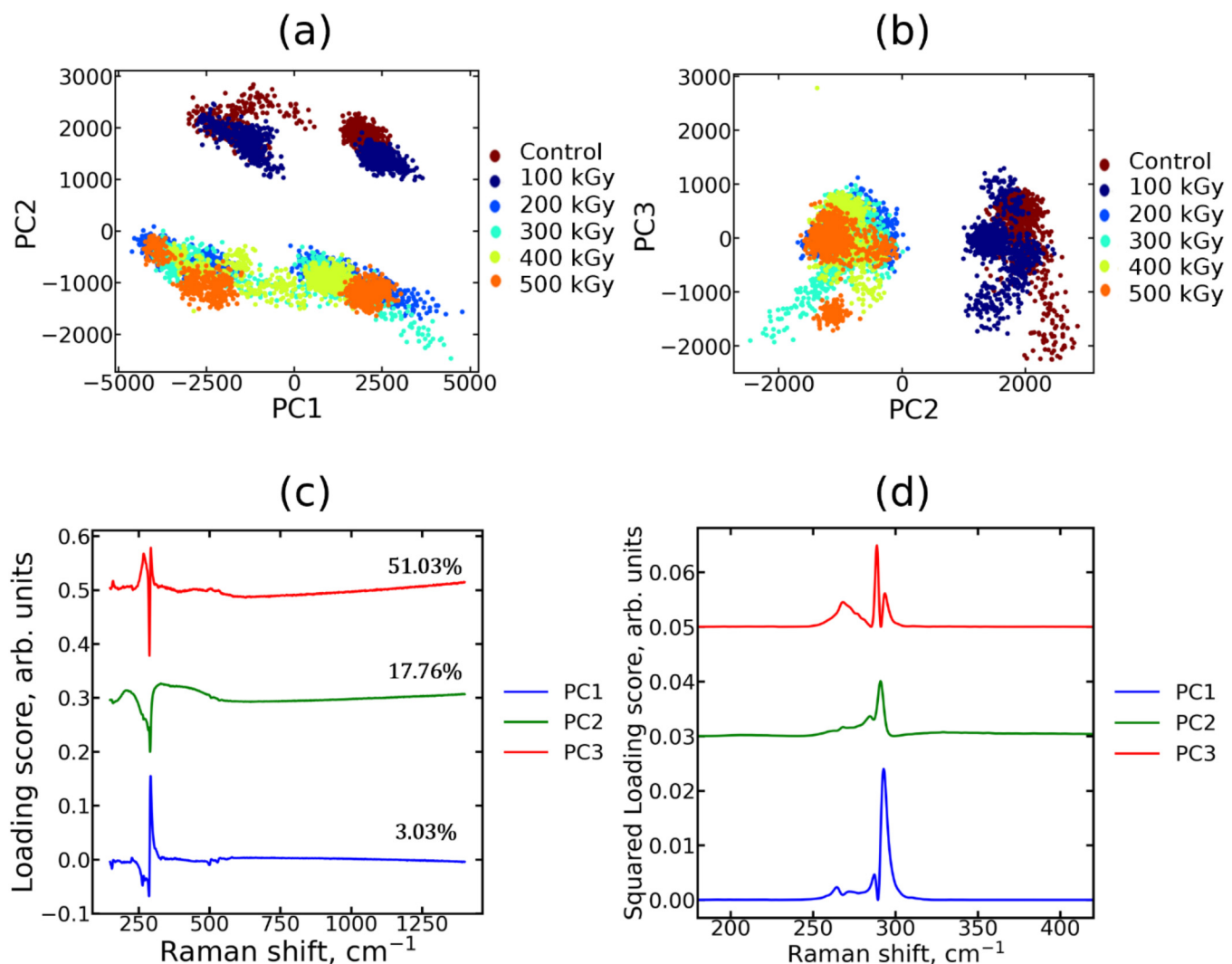
FIG. 2. Processed Raman spectra of the training dataset prior to feeding into ML algorithms. Each spectrum was stacked with an offset of 100 arbitrary units. The wider linewidth indicates a larger number of spectra in that particular dose. The inset shows the average spectrum of each radiation dose.

$$\mathbf{T} = \mathbf{U}\mathbf{\Sigma}, \quad (3)$$

where \mathbf{T} and \mathbf{V} are the matrices of PC and loading, representing Eigenvectors of the covariance matrix of $\mathbf{B}^T\mathbf{B}$, respectively. Equation (3) was derived from SVD, where \mathbf{U} comprises an orthogonal matrix consisting of the left singular Eigenvectors of $\mathbf{B}\mathbf{B}^T$ and $\mathbf{\Sigma}$ is a diagonal matrix consisting of the ranked singular (Eigen) values.

In addition to PCA, two classification models were utilized: a pure linear discrimination analysis (LDA) model and a model combining PCA and LDA (PCA-LDA). These models were employed to classify the training set data and predict absorbed doses in the testing dataset. Using the scikit-learn library, LDA was applied

for supervised dimensionality reduction, projecting the input data into a linear subspace that maximizes the separation between classes. LDA is a multiclass classifier with a linear decision boundary, generated by fitting class conditional densities to the data and using Bayes' rule. The model fits a Gaussian density to each class, assuming that all classes share the same covariance matrix. For the PCA-LDA model, both the training and testing datasets were initially processed with PCA to extract the first ten score vectors, accounting for 75.57% of the data's variation. The clustered data points in the reduced dimensions were then fed into an LDA classification model. A blind test report was generated to interpret the prediction and classification results of the testing datasets obtained from each model.



14 January 2024 17:40:57

FIG. 3. PCA data distribution. In (a) and (b), the distribution of data points in the PCA model is shown, plotting PC 2 against PC 1 and PC 3 against PC 2, respectively. (c) and (d) depict the loading scores and their squared values for each analyzed Raman shift feature.

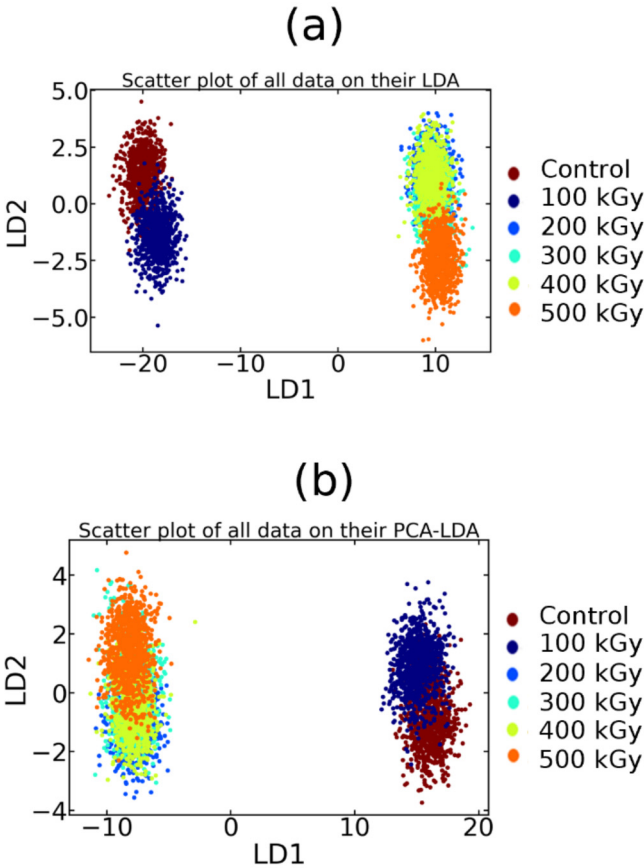


FIG. 4. Classification results of (a) LDA and (b) PCA-LDA models.

III. RESULTS AND DISCUSSIONS

In Fig. 2, we present the processed training spectra, comprising 5300 data points from six distinct data groups corresponding to absorbed doses of 0, 100, 200, 300, 400 and 500 kGy. These spectra exhibit a remarkable degree of similarity, except for sporadic cosmic ray spikes that occurred randomly. Due to their broad peak widths, these residue cosmic spikes remained undetected by the

algorithm. It is worth noting that a slight intensity offset has been applied to these highly overlapping spectra for better visualization. Raman spectra prominently feature a distinct and strong peak at the Raman shift position of 291 cm^{-1} , which corresponds to the first-order longitudinal optical phonon (LO) peak of GaAs. A less-defined peak is observed at a Raman shift of 268 cm^{-1} , corresponding to the first-order transverse optical phonon (TO) peak of GaAs. This observation is noteworthy because, according to the selection rules derived from the symmetry-related Raman dispersion tensor, the TO phonon is symmetry-forbidden in first-order Raman scattering of a (100) zinc-blende crystal, particularly, when measured in a backscattering geometry.⁴³ This phenomenon suggests the presence of native defects formed during the crystal growth process of the SI-GaAs wafer. Additionally, these spectra exhibit several overtones of higher-order scattering.^{44,45}

The collected Raman spectra show a noteworthy difference in the behavior of electron beam irradiated undoped SI-GaAs wafers compared to ion-implanted GaAs samples^{28,46,47} and Si doped GaAs film subjected to gamma irradiation.⁴⁸ Unlike these cases, the collected spectra of the electron beam irradiated undoped SI-GaAs wafer remain highly identical. This suggests that defect-rich GaAs is more susceptible to crystal deformation compared to an undoped perfect crystal. Previous research indicated that intense electron beam irradiation can degrade the crystallinity of SI-GaAs, as evidenced by the broadening of a Bragg profile measured by x-ray diffraction (XRD).⁴⁹ Figure S1 in the [supplementary material](#) shows high-resolution XRD spectra of the irradiated samples measured at a step size of 0.001° . Despite the high-resolution mode, none of significant changes could be observed based on the measurement result. The studied absorbed dose in this study was two orders lower than in the previous study, making changes in crystallinity less noticeable by XRD.

Figure 3 displays the results of PCA, focusing on the first three PCs that collectively account for 71.80% of the variance. Notably, data points from all six irradiation groups (0–500 kGy) are distributed across both positive and negative sides of PC 1 and PC 3. An interesting distinction emerges with PC 2, which effectively separates samples with low irradiation doses (0–100 kGy) from those with high irradiation doses (200–500 kGy). In Fig. 3(c), the loading scores for each Raman shift are illustrated. PC 1 predominantly captures the main peaks and overtones observed in Fig. 2. PC 2 exhibits a distinctive broad, asymmetry peak with negative scores spanning from a Raman shift of $234\text{--}299\text{ cm}^{-1}$. In PCA

TABLE I. The confusion matrix of the blind test obtained by the pure LDA model.

Actual value	Value predicted by the pure LDA model						Total
	Control	100 kGy	200 kGy	300 kGy	400 kGy	500 kGy	
Control	31	67	0	0	0	0	98
100 kGy	34	65	0	0	0	0	99
200 kGy	0	0	41	27	28	3	99
300 kGy	0	0	33	30	26	10	99
400 kGy	0	0	34	42	16	7	99
500 kGy	0	0	17	16	15	45	93

TABLE II. The confusion matrix of the blind test obtained by the PCA-LDA model.

Actual value	Value predicted by the PCA-LDA model						Total
	Control	100 kGy	200 kGy	300 kGy	400 kGy	500 kGy	
Control	12	86	0	0	0	0	98
100 kGy	11	88	0	0	0	0	99
200 kGy	0	0	71	4	21	3	99
300 kGy	0	0	33	31	30	1	99
400 kGy	0	0	28	39	20	12	99
500 kGy	0	0	10	11	16	56	93

interpretation, a negative loading signifies the absence of a particular characteristic in the associated principal component.⁵⁰ Notably, Fig. 3(d) displays squared loading scores, highlighting that the first three PCs are primarily influenced by the first-order optical phonon peaks of GaAs, falling within the range of 260–300 cm^{-1} . For example, the highest loading scores of PC 1, PC 2, and PC 3 correspond to features at 292.8, 291.0[−], and 288.8 cm^{-1} , respectively.

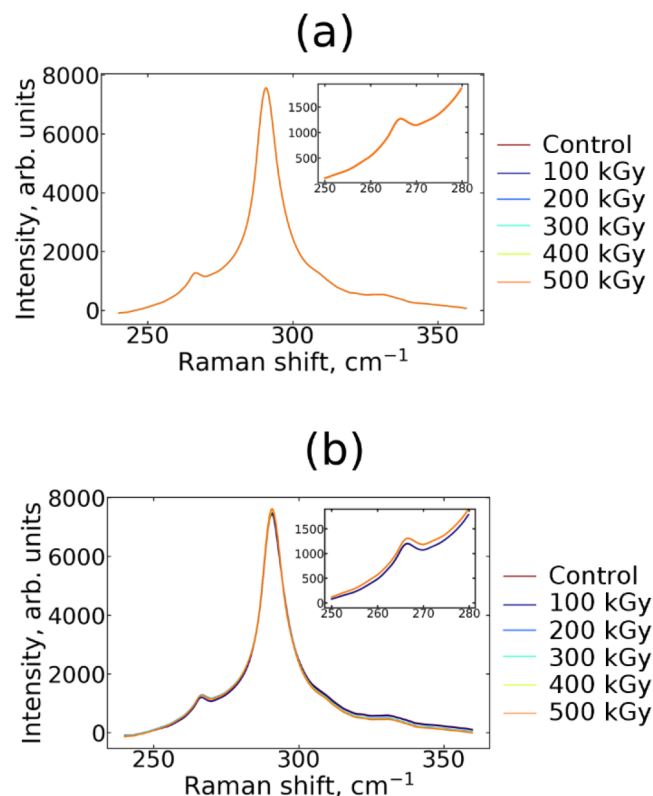
**FIG. 5.** The inverse spectra of (a) the LDA model and (b) the PCA-LDA model calculated based on the center data point of each group.

Figure 4 presents the computed results of both the LDA model (a) and the PCA-LDA model (b). Despite a sign flip in LD 1, both models exhibit remarkably similar outcomes. These classification models demonstrate their ability to effectively distinguish between samples subjected to low irradiation doses and those exposed to high irradiation doses through the first discriminant component, LD 1. Among the low irradiation dose samples, the models can further differentiate between 0 and 100 kGy using the second discriminant component, LD 2, to a moderate extent. However, for the high irradiation dose samples (200–500 kGy), their distributions significantly overlap, except for 500 kGy.

Tables I and II display the blind test results for the testing dataset using the pure LDA model and the PCA-LDA model, respectively. The calculated results indicate that the accuracy of predicting the classes in the testing dataset is 0.39 for the pure LDA model and 0.48 for the PCA-LDA model. These results suggest that the dataset can be effectively categorized into two classes, rather than the originally defined five different classes based on their irradiation doses. For both pure LDA model and PCA-LDA model, a prediction accuracy of 100% can be achieved by relabeling the training set data as either low irradiation samples (0–100 kGy) or high irradiation samples (200–500 kGy).

While both the pure LDA model and the PCA-LDA model exhibit similar data distribution patterns, the blind test result shows that the PCA-LDA model achieves better prediction accuracy. The improvement is attributed to the PCA's ability to preserve the main features of the original dataset to an optimum extent. Figure 5 illustrates the inverse spectra of the center data point of each group in the scatterplots presented in Fig. 4, obtained by performing the Moore-Penrose pseudo-inverse of the reduced matrices generated by the LDA models. Figure 5(a) shows that all the inverse spectra of the pure LDA model were the same because the information was lost due to the discarded components. On the other hand, the PCA-LDA model's inverse spectra depicted in Fig. 5(b) demonstrate a significant variation, particularly at the TO phonon peak of GaAs in the range of 250–280 cm^{-1} . This variation clearly distinguishes between low irradiation dose (0–100 kGy) and the high irradiation dose (200–500 kGy) samples, as shown in the insets of Fig. 5. The slight frequency shift and broadening of TO peaks in the high irradiation dose samples indicate a lattice change produced by the strains resulting from the displacement of atoms of the host lattice from their normal equilibrium positions. Compared to the previous work reported by Phakkhawan *et al.*, revealing the

14 January 2024 17:40:57

broadening of Raman TO peaks of the 10 MeV electron-irradiated SI-GaAs sample at 500 kGy, our analysis method provides a more sensitive detection of the lattice distortion caused by Ga or As vacancies, interstitial atoms, amorphous formation, and mosaic structure domain.⁵¹

IV. CONCLUSIONS

Using multivariate statistical analyses in combination with Raman spectroscopy, we have proposed a sensitive non-destructive technique for investigating radiation defects. The Raman measurement revealed nearly identical spectra for GaAs wafers irradiated at doses ranging from 100 to 500 kGy. Through the application of ML algorithms, specifically a pure LDA model and a PCA-LDA model, we successfully categorized the samples into two groups: low irradiation and high irradiation. Despite the high similarity in the obtained Raman spectra, the ML algorithm accurately predicted the blind test samples. These results underscore the formidable power of ML and its potential applicability in the study of various defects. In conclusion, our study demonstrates the effectiveness of combining multivariate statistical analyses with Raman spectroscopy for detecting radiation-induced defects. This approach holds promise for defects studies in other materials and applications.

SUPPLEMENTARY MATERIAL

See the supplementary material for the results of high-resolution x-ray diffraction measurement.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Thailand's Program Management Unit for Competitiveness (PMU-C) (Contract No. C10F650168), National Electronics and Computer Technology Center, and Thailand Institute of Nuclear Technology (TINT) for financial and instrumentation supports. The authors would also like to express a sincere gratitude to Ms. Supakan Kijamnajsuk from National Metal and Materials Technology Center in performing and analyzing HRXRD measurements.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Jia Yi Chia: Formal analysis (equal); Visualization (equal); Writing – original draft (equal). **Nuatawan Thamrongsiripak:** Methodology (equal); Validation (supporting). **Sornwit Thongphanit:** Methodology (equal); Validation (equal). **Noppadon Nuntawong:** Conceptualization (equal); Data curation (equal); Funding acquisition (equal); Methodology (equal); Writing – review & editing (equal).

DATA AVAILABILITY

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

REFERENCES

- ¹V. S. Vavilov, N. P. Kekelidze, and L. S. Smirnov, *Effects of Radiation on Semiconductors* (Springer, New York, 1965).
- ²V. S. Vavilov and N. A. Ukhin, *Radiation Effects in Semiconductors and Semiconductor Devices* (Consultants Bureau, New York, 1995).
- ³J. R. Srour, "Radiation effects R&D in the 1970s A retrospective view," *IEEE Trans. Nucl. Sci.* **41**(6), 2660–2665 (1994).
- ⁴K. Iniewski, *Radiation Effects in Semiconductors* (CRC Press, 2018).
- ⁵V. A. Kozlov and V. V. Kozlovski, "Doping of semiconductors using radiation defects produced by irradiation with protons and alpha particles," *Semiconductors* **35**, 735–761 (2001).
- ⁶Y. Gonzalez-Velo, H. J. Barnaby, and M. N. Kozicki, "Review of radiation effects on ReRAM devices and technology," *Semicond. Sci. Technol.* **32**(8), 083002 (2017).
- ⁷C. Y. Huang, "The effect of gamma irradiation on the stability of amorphous InGaZnO metal-semiconductor-metal UV photodetectors," *J. Non Cryst. Solids* **546**, 120292 (2020).
- ⁸Z. Orhan, E. Cinan, Z. Çaldıran, Y. Kurucu, and E. Daş, "Synthesis of CuO-graphene nanocomposite material and the effect of gamma radiation on CuO-graphene/p-Si junction diode," *J. Mater. Sci. Mater. Electron.* **31**, 12715–12724 (2020).
- ⁹J. F. Felix, A. F. Da Silva, S. W. Da Silva, F. Qu, B. Qiu, J. Ren, W. M. de Azevedo, M. Henini, and C.-C. Huang, "A comprehensive study on the effects of gamma radiation on the physical properties of a two-dimensional WS₂ monolayer semiconductor," *Nanoscale Horiz.* **5**(2), 259–267 (2020).
- ¹⁰H. Wong and H. Iwai, "The road to miniaturization," *Phys. World* **18**(9), 40 (2005).
- ¹¹M. J. Madou, *Fundamentals of Microfabrication: the Science of Miniaturization* (CRC Press, 2018).
- ¹²D. Stanisavljevic and M. Spitzer, A review of related work on machine learning in semiconductor manufacturing and assembly lines. In *SAMI 2016—Science, Application and Methods in Industry 4.0: Proceedings of the 1st International Workshop on Science, Application and Methods in Industry 4.0 co-located with (i-KNOW 2016)* (CEUR Workshop Proceedings, 2016).
- ¹³L. Lingitz, V. Gallina, F. Ansari, D. Gyulai, A. Pfeiffer, W. Sihm, and L. Monostori, "Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer," *Procedia CIRP* **72**, 1051–1056 (2018).
- ¹⁴G. Tello, O. Y. Al-Jarrah, P. D. Yoo, Y. Al-Hammadi, S. Muhaidat, and U. Lee, "Deep-structured machine learning model for the recognition of mixed-defect patterns in semiconductor fabrication processes," *IEEE Trans. Semicond. Mfg.* **31**(2), 315–322 (2018).
- ¹⁵C. Wright and S. J. Yang, "Deep learning for automated focus quality detection in wafer inspection," *Proc. SPIE* **11787**, 146–157 (2021).
- ¹⁶R. Ghoshhajra, K. Biswas, and A. Sarkar, "A review on machine learning approaches for predicting the effect of device parameters on performance of nanoscale MOSFETs," *2021 Devices for Integrated Circuit (DevIC)* (IEEE, 2021), pp. 489–493.
- ¹⁷A. K. Pimachev and S. Neogi, "First-principles prediction of electronic transport in fabricated semiconductor heterostructures via physics-aware machine learning," *NPJ Comput. Mater.* **7**(1), 93 (2021).
- ¹⁸K. Nordlund, S. J. Zinkle, A. E. Sand, F. Granberg, R. S. Averback, R. E. Stoller, T. Suzudo, L. Malerba, F. Banhart, W. J. Weber, F. Willaime, S. L. Dudarev, and D. Simeone, "Primary radiation damage: A review of current understanding and models," *J. Nucl. Mater.* **512**, 450–479 (2018).
- ¹⁹M. R. Tonks, A. Cheniour, and L. Aagesen, "How to apply the phase field method to model radiation damage," *Comput. Mater. Sci.* **147**, 353–362 (2018).
- ²⁰M. J. Caturla, "Object kinetic Monte Carlo methods applied to modeling radiation effects in materials," *Comput. Mater. Sci.* **156**, 452–459 (2019).
- ²¹J. A. Stewart, A. A. Kohnert, L. Capolungo, and R. Dingreville, "Design and analysis of forward and reverse models for predicting defect accumulation, defect energetics, and irradiation conditions," *Comput. Mater. Sci.* **148**, 272–285 (2018).

- ²²M. Kitajima, "Defects in crystals studied by Raman scattering," *Crit. Rev. Solid State Mater. Sci.* **22**(4), 275–349 (1997).
- ²³J. Ibáñez and R. Cuscó, "Raman spectroscopy of compound semiconductors," in *Semiconductor Research: Experimental Techniques*, edited by P. Amalia and B. Naci (Springer, Berlin, 2012), pp. 259–281.
- ²⁴N. Papež, A. Gajdoš, D. Sobola, R. Dallaev, R. Macků, P. Škarvada, and L. Grmela, "Effect of gamma radiation on properties and performance of GaAs based solar cells," *Appl. Surf. Sci.* **527**, 146766 (2020).
- ²⁵H. Y. Kim, J. A. Freitas, and J. Kim, "Probing proton irradiation effects in GaN by micro-Raman spectroscopy," *Europhys. Lett.* **96**(2), 26004 (2011).
- ²⁶T. Koyanagi, M. J. Lance, and Y. Katoh, "Quantification of irradiation defects in beta-silicon carbide using Raman spectroscopy," *Scr. Mater.* **125**, 58–62 (2016).
- ²⁷F. Cataldo and S. Iglesias-Groth, "Neutron damage of hexagonal boron nitride: h-BN," *J. Radioanal. Nucl. Chem.* **313**(1), 261–271 (2017).
- ²⁸J. He, Y. Shen, B. Li, X. Xiang, S. Li, X. Fang, H. Xiao, X. Zu, and L. Qiao, "A comparative study of the structural and optical properties of Si-doped GaAs under different ion irradiation," *Opt. Mater.* **111**, 110611 (2021).
- ²⁹S. Mishra, D. Kabiraj, A. Roy, and S. Ghosh, "Effect of high-energy light-ion irradiation on Si-GaAs and GaAs: Cr as observed by Raman spectroscopy," *J. Raman Spectrosc.* **43**(2), 344–350 (2012).
- ³⁰S. Agarwal, Q. Chen, T. Koyanagi, Y. Zhao, S. J. Zinkle, and W. J. Weber, "Revealing irradiation damage along with the entire damage range in ion-irradiated SiC/SiC composites using Raman spectroscopy," *J. Nucl. Mater.* **526**, 151778 (2019).
- ³¹S. Miro, G. Velisa, L. Thomé, P. Trocellier, Y. Serruys, A. Debelles, and F. Garrido, "Monitoring by Raman spectroscopy of the damage induced in the wake of energetic ions," *J. Raman Spectrosc.* **45**(6), 481–486 (2014).
- ³²Z. Xu, Z. He, Y. Song, X. Fu, M. Rommel, X. Luo, A. Hartmaier, J. Zhang, and F. Fang, "Topic review: Application of Raman spectroscopy characterization in micro/nano-machining," *Micromachines* **9**(7), 361 (2018).
- ³³F. Lussier, V. Thibault, B. Charron, G. Q. Wallace, and J. F. Masson, "Deep learning and artificial intelligence methods for Raman and surface-enhanced Raman scattering," *Trends Anal. Chem.* **124**, 115796 (2020).
- ³⁴C. Claeys and E. Simoen, *Radiation Effects in Advanced Semiconductor Materials and Devices* (Springer Science & Business Media, Berlin, 2002).
- ³⁵A. Šagátová, B. Zafko, F. Dubecký, T. L. Anh, V. Nečas, K. Sedláčková, M. Pavlovič, and M. Fülöp, "Radiation hardness of GaAs sensors against gamma-rays, neutrons and electrons," *Appl. Surf. Sci.* **395**, 66–71 (2017).
- ³⁶A. Sagatova, V. Krsjak, S. Sojak, O. Riabukhin, E. Kovacova, and B. Zafko, "Semi-insulating GaAs detectors degraded by 8 MeV electrons up to 1500 kGy," *J. Instrum.* **16**(12), C12032 (2021).
- ³⁷G. S. Was, *Fundamentals of Radiation Materials Science: Metals and Alloys* (Springer, Berlin, 2016).
- ³⁸B. Mortada, "Trend removal from Raman spectra With local variance estimation and cubic spline interpolation," *Circuits Syst. Int. J. (CSIJ)* **2**(1), 1–12 (2015).
- ³⁹Y. Xie, Q. You, P. Dai, S. Wang, P. Hong, G. Liu, J. Yu, X. Sun, and Y. Zeng, "How to achieve auto-identification in Raman analysis by spectral feature extraction & adaptive hypergraph," *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **222**, 117086 (2019).
- ⁴⁰D. A. Whitaker and K. Hayes, "A simple algorithm for despiking Raman spectra," *Chemometr. Intell. Lab. Syst.* **179**, 82–84 (2018).
- ⁴¹P. C. Mahalanobis, "On the generalized distance in statistics," *Sankhya Ser. A* **80**, S1–7 (2018).
- ⁴²F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, and J. Vanderplas, "Scikit-learn: Machine learning in python," *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
- ⁴³P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors* (Springer-Verlag, Berlin, 2005).
- ⁴⁴T. Sekine, K. Uchinokura, and E. Matsuura, "Two-phonon Raman scattering in GaAs," *J. Phys. Chem. Solids* **38**(9), 1091–1096 (1977).
- ⁴⁵C. Patel, T. J. Parker, H. Jamshidi, and W. F. Sherman, "Phonon frequencies in GaAs," *Phys. Status Solidi B* **122**, 461–467 (1984).
- ⁴⁶P. S. Peercy, "Raman scattering of ion-implanted GaAs," *Appl. Phys. Lett.* **18**(12), 574–576 (1971).
- ⁴⁷S. K. Mohanta, R. K. Soni, N. Gosvami, S. Tripathy, and D. Kanjilal, "Morphological and micro-Raman investigations on Ar⁺-ion irradiated nano-structured GaAs surface," *Appl. Surf. Sci.* **253**(10), 4531–4536 (2007).
- ⁴⁸Y. Shen, X. Fang, X. Ding, H. Xiao, X. Xiang, G. Yang, M. Jiang, X. Zu, and L. Qiao, "Structural features and photoelectric properties of Si-doped GaAs under gamma irradiation," *Nanomater* **10**(2), 340 (2020).
- ⁴⁹J. Y. Chia, Y. Zhang, K. Li, W. Kusolthossakul, A. Sathukarn, K. Tantiwanichapan, P. Rattanawan, R. Jintamethasawat, N. Thamrongsiripak, and N. Nuntawong, "Facile activation of a GaAs substrate with electron beam irradiation for THz photoconductive antenna," *Appl. Phys. Express* **15**(10), 107002 (2022).
- ⁵⁰I. Burstyn, "Principal component analysis is a powerful instrument in occupational hygiene inquiries," *Ann. Occup. Hyg.* **48**(8), 655–661 (2004).
- ⁵¹A. Phakkhawan, A. Sakulalavek, S. Buranurak, P. Klangtakai, K. Pangza, N. Jangsawang, S. Nasompag, M. Horprathum, S. Kijamnajsuk, and S. Sanorpim, "Investigation of radiation effect on structural and optical properties of GaAs under high-energy electron irradiation," *Materials* **15**(17), 5897 (2022).