Perspective

Machine learning and its applications for plasmonics in biology

Gwiyeong Moon,1 Jongha Lee,1 Hyunwoong Lee,1 Hajun Yoo,1 Kwanhwi Ko,1 Seongmin Im,1 and Donghyun Kim1,*

SUMMARY

Machine learning (ML) has drawn tremendous interest for its capacity to extract useful information that may be overlooked with conventional analysis techniques and for its versatility in a wide range of research domains, including biomedical sensing and imaging. In this perspective, we provide an overview focused on the uses and benefits of ML in areas of plasmonics in biology. ML methodologies for processing data from plasmonic biosensing and imaging systems by supervised and unsupervised learning to achieve enhanced detection and quantification of target analytes are described. In addition, deep learning-based approaches to improve the design of plasmonic structures are presented. Data analysis based on ML for classification, regression, and clustering by dimension reduction is presented. We also discuss ML-based prediction and design of plasmonic structures and sensors using discriminative and generative models. Challenges and the outlook for ML for plasmonics in biology are summarized. Based on these insights, we are convinced that ML can add value to plasmonics techniques in biological sensing and imaging applications to make them more powerful with improved detection and resolution.

INTRODUCTION

Surface plasmon (SP) refers to electron density waves formed at the interface between metal and dielectric material. SP can be spatially localized with confined volume and produce enhanced electromagnetic (EM) fields, leading to diverse applications such as single-molecule detection,1,2 photovoltaic devices,3,4 and light-emitting devices.5 Resonance of SP shifts depends on the ambient conditions, which can be used to measure biochemical events such as bioaffinity interactions and quantify biomolecular concentration and kinetics.6–8 For this reason, many novel concepts of sensing and imaging techniques have appeared in biology and biomedical applications; e.g., SP resonance (SPR) biosensing,9–11 surface-enhanced Raman spectroscopy (SERS),12–14 and surface-enhanced infrared absorption spectroscopy (SEIRA).15–17

SP can be used to image molecular events based on metal-enhanced fluorescence with directional radiation of fluorophores.18–20 In labeled imaging techniques, an enhanced light signal has been achieved with super-resolution by localization of SP using plasmonic nanostructures, such as nanoholes, disks, and random arrays,21–25 and this may modulate optical characteristics for applications in imaging cell internalization of a virus, bacterial motility, intracellular mitochondrial movement, and structured illumination microscopy.26–28 Label-free imaging has been
attempted; e.g., in SPR microscopy (SPRM) for detection of extracellular vesicles,\textsuperscript{29} imaging of neurons,\textsuperscript{30} and quantification of cell adhesion.\textsuperscript{31} SP localization has also been used for super-resolved Raman microscopy.\textsuperscript{32,33} However, optical techniques using SP have their own issues. Imaging with labels suffers from limits because of fluorescence quenching when fluorophores are in proximity to the metallic surface. Label-free microscopy has relatively poor image quality because of the large propagation length of SP.

Ultimately, the issues that have plagued detection and imaging techniques based on SP affect signal-to-noise ratios (SNRs) and resolution. One of many approaches proposed to address these issues is to employ machine learning (ML) for improved application of SP techniques. ML refers to analysis of machine-intensive information that uses data-driven methods to create and train models without explicit programming as one of the applications of artificial intelligence. These models can be used to perform specific tasks; e.g., predicting outputs of data points and determining their class. Interpretation of data becomes more difficult as the data become more complicated and massive. ML is a novel method for developing machines that can recognize patterns and trends independent of human intervention and may use diverse commercial and non-commercial tools built on programming languages such as Python, R, Java, Julia, and MATLAB. Each language has its own advantages for performing certain tasks; e.g., Python and MATLAB are overall good candidates for handling images. ML can be adopted to analyze data to extract meaning and implications and design plasmonic structures with desired EM responses, where data may be preprocessed; e.g., by filtration, purification, and normalization.

ML-based approaches substantially improve data processing of signals/images measured from plasmonic biosensors and imaging systems by analyzing the relationship between data and target information, allowing detection and quantification of target analytes. ML does not require manual modeling to analyze data, and it is potent for high-dimensional data. These advantages make ML a powerful analytical tool for plasmonic systems. Figure 1 presents how ML can be useful in plasmonic biosensing and imaging, which can be divided into two main categories.
First, ML can provide a myriad of methodologies to analyze results obtained from plasmonic sensing for data processing into classification, regression, and dimension reduction/clustering (Figure 2). The three components do not work independently and can be employed together for effective data analysis.

Second, ML can be utilized to investigate the relationship between plasmonic nanostructure and EM responses (Figure 3). EM simulation can be conducted to obtain various optical characteristics of plasmonic structures. However, conventional methods require a long computation time and high memory usage, especially for calculation of the 3D structure. It needs a one-time training procedure, after which it can predict EM characteristics with much faster computation and less memory, although ML may not always be efficient, depending on applications, because generation of training data and optimization of an ML model can take a very long time. In addition to estimation of EM response, structures and material for plasmonic nanostructure with desired characteristics may be designed. This process, which may be referred to as an inverse design, is important to achieve ultrahigh sensitivity by amplifying light-matter interaction.34,35 Deep learning in particular has opened up a new way for plasmonic structure design with its distinct advantages of extracting useful information from massive datasets. Various approaches based on modification of neural network architecture and algorithms have been established, surpassing earlier design methods.

In this perspective, we first introduce ML algorithms that include supervised, unsupervised, and deep learning. Diverse ML approaches to process data from plasmonic sensors and imaging systems are described based on classification,
regression, and dimension reduction/clustering. We then discuss methodologies used to investigate the relationship between plasmonic structure and its optical responses. The methodologies include prediction of optical responses associated with plasmonic structures; for instance, reflection, transmission, absorption, and scattering intensity spectrum. Methods to design plasmonic structures to fit desired optical responses are introduced, with an emphasis on neural network-based approaches. This is followed by challenges for use of ML and the outlook of these approaches in the future. This perspective can benefit readers and nurture advanced applications of ML for plasmonic sensors and imaging systems.

**ML ALGORITHMS**

**Supervised/unsupervised learning**

Supervised learning provides a labeled dataset to the machine that maps the relationship between input and output data. Linear regression, k-nearest neighbor (k-NN) and support vector machines (SVMs) are representative examples of supervised learning. Linear regression, which is relatively simple to comprehend and implement, identifies a linear approximate relation between different variables. k-NN is an instance-based model that finds the most similar $k$ data points to estimate the output of data points of interest. It can be operated without training, but its performance suffers when few data are provided. SVM performs classification by calculating the hyperplane that maximizes the distance between the plane and the closest data. SVM works effectively when the number of samples is relatively small compared with the number of features.

In unsupervised learning, the machine finds hidden patterns in unlabeled data for classification and clustering. Principal-component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE) are two example algorithms of unsupervised learning. PCA is a linear transformation that can be used for dimension reduction and feature extraction by determining the direction in which the variation in data is at a maximum. PCA, along with many other techniques for cluster analysis and regression, started from the category of statistical analysis despite being often classified as an ML technique. t-SNE is a non-linear dimensionality reduction method that provides efficient visualization of data while maintaining dataset cluster distributions.
Deep learning

Artificial neural networks (ANNs) were inspired by biological neural networks and consist of a collection of interconnected nodes that are typically organized into multiple layers. The connections are used to transfer signals from one node to another. Deep learning models are ANNs that have multiple layers between input and output, and they can be classified into two types: discriminative and generative models. Discriminative models are used for classification or regression, whereas generative models are used to create new data points that are indistinguishable from the original dataset. Multilayer perceptron (MLP) is a basic form of discriminative ANN that consists of at least three layers, including input, hidden, and output layers. Convolutional neural networks (CNNs) are discriminative models frequently used for image analysis. Local spatial features of data can be extracted by convolution of shared-weight filters that can be optimized during the training procedure. Generative models, such as generative adversarial networks (GANs) and variable autoencoders (VAEs), approximate specific probability distributions and create new data samples based on the underlying distribution. VAEs are autoencoders that are trained under regularization to make latent space generate new data efficiently. On the other hand, GANs typically consist of two parts: a generating part, which is trained to produce new data samples that are indistinguishable from the original dataset, and the discriminating part, which is trained to identify whether the data come from the original dataset or the generative model. The two parts of GANs develop by competing against each other and attempting to reach a Nash equilibrium, where the discriminator cannot differentiate data from real data, while diverse approaches are introduced to improve the stability of GAN learning.

ML-BASED DATA ANALYSIS

Classification

The basic form of classification is a binary classification that divides data into two categories effective for a wide range of applications. For example, the concentration of cytokines has been analyzed with a CNN by analyzing the dark-field image of silver nanocubes conjugated with a cytokine. The scattering spots of nanocubes were detected with a CNN that classified the pixels of images into two categories, existence and non-existence, by yielding the binary output of the neural network as 1 and 0, respectively (Figure 4A). Binary classification has also been conducted to evaluate exposure to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) on plasmonic genosensors using several architectures of CNNs that were pre-trained with an ImageNet dataset to estimate the presence of neurotransmitters and identify the bacterial strains.

There is a need for classification into multiple categories beyond binary classification. For instance, the number of nanoparticles has been estimated by analyzing SPRM with a CNN. The CNN classified images into several categories based on the number of nanoparticles, allowing quantification of analytes in highly complex SPRM images with overlapping parabolic point spread functions. A CNN has been used for assessment of the degree of DNA damage, detection of cancer cells and liver diseases (Figure 4B), and determination of virus size and extracellular vesicles by classifying them into several categories. Although CNNs were originally developed for two-dimensional data, it may be used to analyze one-dimensional data as well, with proper conversion to two-dimensional data. However, achieving generalization capability is challenging. Thus, various approaches based on one-dimensional CNNs have been proposed for some applications and employed to recognize different substances from SERS spectra.
Figure 4. ML implementation for analysis of data from plasmonic sensing and imaging

(A) CNN-based analysis for counting silver nanocubes from dark-field images. Adapted by permission from Gao et al.43 Copyright 2021 American Chemical Society. 

(B) CNN-based analysis of serum: SERS spectra for serological detection of liver diseases. Reprinted from Cheng et al.47 Copyright 2021 with permission from Elsevier.

(C) Schematic of the development of an SVM model for analysis of SERS spectra of bacterial strains. Adapted with permission from Rahman et al.48 Copyright 2021 American Chemical Society.

(D) Schematic of extracellular vesicle detection by analyzing SPRM images with a CNN. Adapted with permission from Yang et al.49 Copyright 2021 American Chemical Society.

(E) Workflow of neural network-based spectral reconstruction from the transmission image of a plasmonic encoder. Adapted with permission from Brown et al.50 Copyright 2021 American Chemical Society.

(F) Workflow of dimension reduction and clustering analysis for reflection curves of meta-plasmonic structures. Reprinted from Moon et al.51 Copyright 2020 with permission from Elsevier.
In addition to CNNs, k-NN has been used to classify SPR sensorgrams into several groups in terms of the type and number of substances; the classified result was processed with further diagnosis protocols. k-NN can be used for identification of specific/cross-reactive binding of an antibody from SPR images and characterization of bacterial species (Escherichia coli) by classifying SERS spectra based on strains, culture conditions, and purification methods. On the other hand, SVM has been used for cell cytometry by classifying lens-free holograms aided by gold and silver nanoparticles and for differentiation of cancer cell-derived exosomes and various bacterial species from SERS spectra (Figure 4C). SVM was originally designed to be suited for binary classification, but multiclass classification can be carried out by splitting into multiple binary classifications with one versus all.

**Regression**

Continuous value can be predicted by regression; e.g., the concentration and size of analytes. For instance, various approaches, including MLP, CNNs, and Gaussian process, were applied to estimate the concentration of glucose from SEIRA spectra and drug molecules, including methylxanthine, E. coli O 157:H7, and rhodamine of a single molecule from SERS spectra. These approaches have yielded reliable results even when the concentration of analyte and data from the sensors were non-linearly related. The size as well as concentration of extracellular vesicles has been estimated by applying CNNs to SPRM images (Figure 4D).

A regression can be performed for analysis of multiple parameters as well as a single parameter. For instance, scattering spectra captured by plasmonic sensors are affected by several environmental parameters (e.g., the thickness and refractive index of a sample layer) and may cause confusion in sensing experiments. For this issue, MLP can efficiently extract relevant information from scattering patterns in the non-resonant wavelength bands and estimate the value of multiple parameters that affect spectra. MLP also allows reconstruction of unknown spectra from encoded information that was obtained by measuring the transmitted intensity of a plasmonic nanostructure and achieved a compact and lightweight solution compared with traditional spectroscopy (Figure 4E).

Two-dimensional data, such as images, can be a candidate output for regression because an image represents a set of values in the two-dimensional space. For example, the low SNR in a dark-field scattering image of gold nanorods in a living cell was converted to an image of scattering light spots, which facilitates clear identification of scattering signals. This is achieved with U-net, an encoder-decoder-based neural network that is well suited for image-to-image conversion. A phase profile of SPR sensor data was estimated from the amplitude of a single back focal plane image of SPRM. The recovered phase profile was applied to refractive sensing to achieve robustness to noise and an enhanced detection limit compared with conventional SPR intensity measurements. Generative models can also be used for regression; for instance, converting out-focused SPRM images into focused ones by training images measured at different focal distances.

**Dimension reduction/clustering**

Dimension reduction can be employed for feature extraction and efficient visualization of data. The extracted feature by dimension reduction can be used for training data of a classification and regression model, overcoming the curse of dimensionality, which refers to the issues related to analysis of data in a high-dimensional space. For example, SERS spectra of single-stranded DNA (ssDNA) sequences were processed with PCA. The dimension-reduced data were fed into a regression model for estimation of chemical
composition, yielding results superior to those from the regression model alone. Features of holographic images of the cell were extracted by dimension reduction with PCA and fed into SVM for classification of cell types.

Dimension reduction can be utilized for visualization of high-dimensional data with an additional clustering algorithm, e.g., k-means clustering. Clustering refers to the process of partitioning data into several groups so that data points in the same group are more comparable than those in other groups. Typically, high-dimensional data are usually reduced to two or three dimensions for plotting patterns more explicitly. For example, dimension reduction of high-dimensional data and clustering of reduced-dimension data have been used for analysis of optical characteristics of a meta-plasmonic sensor (Figure 4F). In the dimension reduction process, 197 dimension data of reflectance were reduced into 8 dimensions. The dimension-reduced data were subsequently mapped into 2 dimensions with t-SNE. The dimension of 197 corresponds to the number of measurement points with an incident angle in the range of 40°–89° at an interval of 0.25°. Finally, the 2-dimensional embedded data were grouped with k-means clustering and achieved efficient visualization of optical characteristics of plasmonic sensing and eventual identification of suitable candidates for practical sensor environment. t-SNE and k-means clustering were then used for visualization of scanning electron images of a plasmonic genosensor. UMAP, a methods for non-linear dimensional reduction, has been adopted for improved data visualization of SERS spectra of E. coli. Dimension reduction can be utilized to enhance weak signals from a sensor. PCA has been applied to enhance the weak signal of SERS, and the signal was reconstructed with 3 times higher intensity by retaining the 3 most significant principal components while eliminating the remaining components.

ML-BASED PREDICTION AND DESIGN
Prediction of EM responses from plasmonic structure
Among traditional ML algorithms, regression, which is used to uncover correlation between variables and predict continuous-output variables, has been demonstrated to estimate EM responses of plasmonic structures given their geometry. In particular, regularization is usually adopted for regression to prevent overfitting by altering the cost function with a penalty for coefficients. Ridge regression, which is also referred to as L2 regularization, has been used to predict the resonance wavelength of gold concave nanocubes with k-fold cross-validation (Figure 5A). Various optimization methods have been employed; for instance, genetic algorithm, particle swarm optimization, and adjoint methods. These traditional regression methods are generally easier to implement in a model and allow easier interpretation of the relationship between variables compared with ANNs (to be described subsequently). Therefore, traditional regression models may be appropriate when they yield results comparable with ANNs, and/or the relationship between structural parameters and EM responses is straightforward.

MLP is widely used to predict various EM responses from the geometry of plasmonic nanostructures. For this task, the input for MLP is set to be the geometry of plasmonic nanostructures, such as size, gap, periods, and rotation angles, and the output to be EM responses. For instance, MLP has been utilized to estimate reflection spectra of gold nanodisk arrays, wavelengths where SPR of gold concave nanocubes occurs, and resonance wavelengths of reflection for gold nanograting.

A CNN can also be employed to predict EM responses. It helps to circumvent the limitation of MLP, which predicts EM responses for constrained structural geometry;
for instance, a dimer structure with varying widths, lengths, and gaps. Two or higher multidimensional data that can represent an arbitrary shape of a structure can be fed into a CNN as an input. Then the features of data can be extracted by various trained kernels with an additional fully connected layer. For instance, a CNN has been utilized to estimate absorption spectra (Figure 5B)\(^87\) and scattering cross-section\(^90\) given an arbitrary shape of plasmonic structure.

The ANNet model, which can train a multilayer neural network without gradient computation, has been employed to circumvent the difficulties of acquiring a large training dataset and achieved much reduced time for training.\(^91\) ANNet successfully predicted leakage radiation characteristics from bright-field images of random nanostructures and showed fewer failure instances than U-net.\(^92\)

**Design of structure/material of plasmonic sensors**

Multiple structures, however, may correspond to one EM response, making the design of a structure ill-posed and the overall training process extremely challenging.\(^93\) Novel approaches, rather than a basic form of the neural network,\(^94\) were therefore introduced to circumvent this issue. A bidirectional neural network has been proposed for design of a plasmonic structure with desired transmission spectra by forming two consecutive networks, a geometry-predicting network (GPN) and a spectrum-predicting network (SPN), as presented in Figure 6A.\(^95\) A GPN was built to predict the geometry of a structure’s given transmission spectrum and an SPN for prediction of the transmission spectrum of a structure. The two networks were processed with simultaneous training, where the spectrum in the training data was fed into the GPN, and the output of the GPN was fed into the SPN. The weights of networks were then updated by back-propagation methods with a loss function, which includes the errors between the output of two networks and the ground truth. This approach has been shown to provide far better performance than separately trained GPNs and SPNs. Similar methods, with some alteration,
were also applied to the design of plasmonic nanoparticles and dimers given spectral cross-section or scattering\(^9^6,9^7\) and gold chiral structure given circular dichroism (Figure 6B).\(^9^8,9^9\) In addition to the bidirectional neural network, an iterative neural network was introduced, in which a forward network was trained to predict EM responses given the geometry of a structure, and then the input parameters (structure geometry) were optimized by minimizing the error between the output of the forward network and desired EM response.\(^1^0^0,1^0^1\)

Generative models, such as GANs and VAEs, have been employed to address issues of discriminative models; e.g., the difficulty of dealing with enormous degrees of freedom in structure and the fixed outcomes given specific input data. GANs, which consist of a generator, critic, and simulator, have been utilized to design unit cell patterns of gold structure and produce desired transmission spectra considering the polarization of the incident and transmitted light (Figure 6C).\(^1^0^2\) The generator provides completely arbitrary patterns for a structure that fulfills desired spectra beyond constraints with few parameters; e.g., width, height, period, thickness, and gap. A GAN has been used indirectly in the design of photonic crystal fiber-based SPR sensors by augmenting training data that were fed into MLP.\(^1^0^3\)

A conditional GAN\(^1^0^4\) in conjunction with a genetic algorithm has been introduced to the design of a dielectric spacer for a gap plasmon-based aluminum structure. The proposed network alleviated the requirement of generating numerous training data from numerical simulations by performing an inverse design and optimization.\(^1^0^5\) A conditional deep convolutional GAN (cDCGAN) can be utilized by integrating a CNN and conditional GAN to achieve a stable Nash equilibrium solution.\(^1^0^6\) The design of the silver structure was demonstrated by employing the cDCGAN, providing a high degree of design freedom with arbitrary patterns of a structure.\(^1^0^7\) A VAE has been employed to design patterns with desired reflectance spectra without a pre-trained forward model, where the latent variables allowed diverse solutions to the inverse design, achieving one-to-many mapping consistent with physical intuition.\(^1^0^8\) A deep learning model that can learn physical principles is highly desired for more general nanostructure design problems. Such a model can overcome weakness under structural constraints and decrease the computational load for simulation of optical responses of plasmonic nanostructures.

**CHALLENGES AND OPPORTUNITIES**

**Importance of data quality**

ML generates output and solutions based on training data, and the data can significantly affect performance. As a result, gathering high-quality data is vital for efficient application of ML. The training data would be experimental by source, which can be obtained by incorporation and optimization of hardware/software for efficient and consistent data acquisition. In some cases, however, obtaining a large amount of data can be challenging, in which case simulation can be performed and used for training an ML model. Collecting training data by simulation requires accurate physical and mathematical modeling of experimental conditions and environments. Noise components, such as Gaussian and/or speckle noise, that can be inevitably present in the experiment should be considered to generate more realistic data.
and, hence, improve the performance of the ML model. Noise can be modeled by considering sources; for instance, thermal and shot noise. One may use transfer learning and one-/few-shot learning techniques, which were introduced to utilize other domain data and label for application of the target domain data. No matter how one prepares training data for ML, it is imperative to have insights into the data domain and collaborate with domain experts in the specific research area to lessen the burden of obtaining training data of high quality and, thus, good performance.

Physicality of ML
Another challenge in ML lies in the difficulty of ensuring physicality and whether the algorithm learned physical principles that govern the characteristics of light and matter. Designing nanostructures is largely conducted by neural network-based methods because the task is a highly complex and ill-posed problem. A neural network is a black box in nature in the sense that it is difficult to obtain insights into the way the model works and to achieve good performance. Despite numerous efforts in the field of computer science to overcome the issue, it remains insufficient to reveal specific black box properties. The results obtained from the learned model may not reflect the physicality. This prevents the model from being useful in general applications and often makes it difficult to predict the data beyond training. Because of these issues, it is still difficult to assure the physicality of ML.

Application specificity of ML algorithms
We also emphasize that ML algorithms have different roles, scopes, and advantages. Thus, it is important to select a suitable algorithm for analysis of data from plasmonic biosensing and imaging systems. Although there is no unique and straightforward solution for this issue, several aspects of ML are worth considering; for instance, format, linearity of data, and memory usage. Specifics in this direction to select ML algorithms are left to guides that are well explained in other studies, and it will be helpful to test various ML algorithms a priori. Although traditional data analysis techniques have been largely limited to data with a small number of parameters, ML can simultaneously consider many features that affect sensing and imaging data by effectively examining the entire data and extracting features using an advanced and optimized algorithm. ML is more suitable for analyzing multimodal and high-throughput sensors, and high-dimensional data can be effectively reduced to low-dimensional data, yielding better reliability and accuracy. ML can be less effective given an insufficient amount of data or when performing relatively simple tasks. In this case, an approach that compromises ML and conventional statistical analysis methods may be desirable. Sensor-related research, including plasmonic sensors, has shown consistently to enhance SNRs, and diverse ML-based methods have been developed for applications working with many other types of signals and images. These methods can be adapted with proper modification for applications of plasmonics in biology and biomedical engineering by providing enhanced sensitivity and improved resolution. To date, most ML studies for plasmonic biosensing and imaging have relied on data-driven approaches without considering plasmonic physical models. However, this can be developed to incorporate theoretical aspects of plasmonics by introducing physics-informed ML, which can enforce physics laws in ML models. This approach has already been applied to diverse research areas, yielding physically consistent results with improved generalization.

Code sharing and standardization
ML and deep learning grow as powerful tools by sharing code and high-quality data while collaborating on projects with enhanced productivity through various platforms and communities. One may begin writing analysis codes for data from sensor
and imaging systems by adopting and modifying ML models that are shared by many researchers instead of starting from scratch. On the other hand, in terms of data, it may be relatively challenging to import shared data and apply one’s own analysis of plasmonic sensor and imaging data because experimental conditions, including equipment and noise level, are different. It is necessary to consider a way for proper adoption of data; for instance, utilization of transfer learning. ML is no panacea for all applications; in this regard, it is important that users understand beforehand whether and how an application can benefit from ML.

**Conclusion**

In this perspective, we reviewed various approaches and applications of ML for plasmonic sensing and imaging in biology. Diverse ML-based methods, such as classification/regression, image recognition, and dimension reduction, were utilized for data processing of SERS, SPRM images and SEIRA, and exploration of the relationship between plasmonic structure and optical response. ML-based approaches provide huge benefits and, at the same time, a cumbersome burden of deciphering complicated patterns by extraction of important features from complex and multidimensional data from plasmonic biosensor and imaging systems. Sophisticated algorithms, such as tandem neural networks and generative models, have been utilized in the design of plasmonic structures ranging from nanoparticles to arbitrary shapes. We anticipate that the use of ML for plasmonic biosensing and imaging will continue to grow because of the unique advantages of efficient extraction of meaningful information based on a data-driven approach and extensibility of incorporating physical models, and in-depth collection of high-quality training data and selection of optimum ML algorithms will remain essential.

**ACKNOWLEDGMENTS**

This work was supported by the National Research Foundation (NRF) (NRF-2022R1A4A2000748) and funded by the Korean Government and the Korea Medical Device Development Fund (RS-2020-KD000088).

**AUTHOR CONTRIBUTIONS**

G.M. wrote the manuscript primarily. G.M., J.L., and H.L. performed and analyzed the literature. G.M. and H.L. prepared figures. H.Y., K.K., and S.I. helped analyze the literature and write the manuscript. D.K. supervised the overall writing of the manuscript.

**DECLARATION OF INTERESTS**

The authors declare no competing interests.

**REFERENCES**

1. Kravets, V.G., Schedin, F., Jalil, R., Britnell, L., Gorbachev, R.V., Ansell, D., Thackray, B., Novoselov, K.S., Geim, A.K., Kabashin, A.V., and Grigorenko, A.N. (2013). Singular phase nano-optics in plasmonic metamaterials for label-free single-molecule detection. Nat. Mater. 12, 304–309.
2. Taylor, A.B., and Zijlstra, P. (2017). Single-molecule plasmon sensing: current status and future prospects. ACS Sens. 2, 1103–1122.
3. Clavero, C. (2014). Plasmon-induced hot-electron generation at nanoparticle/metal-oxide interfaces for photovoltaic and photocatalytic devices. Nat. Photonics 8, 95–103.
4. Kim, Y.J., Moon, G., Hong, H., Yun, J., Kim, S.I., Chae, Y., Kim, D., and Ryu, W. (2020). Plasmon-stimulated bioprotovoltic cells based on thylakoid–AuNR conjugates. J. Mater. Chem. A 8, 24192–24203.
5. Kwon, M.K., Kim, J.Y., Kim, B.H., Park, I.K., Cho, C.Y., Byeon, C.C., and Park, S.J. (2008). Surface-plasmon-enhanced light-emitting diodes. Adv. Mater. 20, 1253–1257.
6. Homola, J., Yee, S.S., and Gauglitz, G. (1999). Surface plasmon resonance sensors. Sens. Actuator. B Chem. 54, 3–15.
7. Xue, T., Liang, W., Li, Y., Sun, Y., Xiang, Y., Zhang, Y., Dai, Z., Duo, Y., Wu, L., Qi, K., et al. (2019). Ultrasensitive detection of miRNA with an antimonene-based surface plasmon resonance sensor. Nat. Commun. 10, 28–29.
8. Lee, H., Rhee, W.J., Moon, G., Im, S., Son, T., Shin, J.-S., and Kim, D. (2021). Plasmon-enhanced fluorescence correlation spectroscopy for super-localized detection of nanoscale subcellular dynamics. Biosens. Bioelectron. 184, 113219.
9. Rich, R.L., and Myszka, D.G. (2000). Advances in surface plasmon resonance biosensor analysis. Curr. Opin. Biotechnol. 11, 54–61.
10. Masson, J.-F. (2020). Portable and field-deployed surface plasmon resonance and plasmonic sensors. Analyst 145, 3776–3800.
11. Byun, K.M., Yoon, S.J., Kim, D., and Kim, S.J. (2007). Experimental study of sensitivity enhancement in surface plasmon resonance biosensors by use of periodic metallic nanowires. Opt. Lett. 32, 1902–1904.

12. Schlucker, S. (2014). Surface-Enhanced Raman spectroscopy: concepts and chemical applications. Angew. Chem. Int. Ed. Engl. 53, 4756–4792.

13. Langer, J., Jimenez de Aberasturi, D., Aizpurua, J., Alvarez-Puebla, R.A., Augüé, B., Baumberg, J.J., Bazan, G.C., Bell, S.E.J., Boisen, A., Brolo, A.G., et al. (2020). Present and future of surface-enhanced Raman scattering. ACS Nano 14, 28–117.

14. Wang, Y., and Irudayaraj, J. (2013). Surface-enhanced Raman spectroscopy at single-molecule scale and its implications in biology. Philos. Trans. R. Soc. Lond. B Biol. Sci. 368, 20120026.

15. Ataka, K., and Heberle, J. (2007). Biochemical applications of surface-enhanced infrared absorption spectroscopy. Anal. Bioanal. Chem. 388, 47–54.

16. Stiles, P.L., Dearing, J.A., Shah, N.C., and Van Duyne, R.P. (2008). Surface-enhanced Raman spectroscopy. Annu. Rev. Anal. Chem. 1, 601–626.

17. Lakowicz, J.R. (2005). Radiative decay engineering 5: metal-enhanced fluorescence and plasmon emission. Anal. Biochem. 337, 171–194.

18. Kinkhabwala, A., Yu, Z., Fan, S., Avlasevich, Y., Mullen, K., and Moerner, W.E. (2009). Large single-molecule surface plasmon enhancements produced by a bowtie nanoantenna. Nat. Photonics 3, 654–657.

19. Kim, K., Oh, Y., Ma, K., Sim, E., and Kim, D. (2009). Plasmon-enhanced total-internal-reflection fluorescence by momentum-mismatched surface nanostructures. Opt. Lett. 34, 3905–3907.

20. Meinerz, N., Barnes, W.L., and Hooper, I.R. (2014). Plasmonic meta-atoms and metamaterials. Nat. Photonics 8, 889–898.

21. Cetin, A.E., Etezadi, D., Galarreta, B.C., Busson, M.P., Eksoulhu, Y., and Altug, H. (2015). Plasmonic nanoarray holes on a robust hybrid substrate for highly sensitive label-free biosensing. ACS Photonics 2, 1167–1174.

22. Son, T., Moon, G., Lee, H., and Kim, D. (2018). Metallic 3D random nanocomposite islands for near-field spatial light switching. Adv. Opt. Mater. 6, 1701219.

23. Forestiere, C., Pascaule, A.J., Capretti, A., Miano, G., Tamburino, A., Lee, S.Y., Reinhard, B.M., and Dal Negro, L. (2012). Genetically engineered plasmonic nanoarrays. Nano Lett. 12, 2037–2044.

24. Son, T., Lee, D., Lee, C., Moon, G., Ha, G.E., Lee, H., Kwak, H., Cheong, E., and Kim, D. (2019). Superlocalized three-dimensional live imaging of mitochondrial dynamics in neurons using plasmonic nanohole arrays. ACS Nano 13, 3063–3074.

25. Kim, K., Choi, J.W., Ma, K., Lee, R., Yoo, K.H., Yun, C.O., and Kim, H. (2017). Nanoisland-based random activation of fluorescence for visualizing endocytotic internalization of adenosine. Small 6, 1293–1299.

26. Lee, W., Kinosita, Y., Oh, Y., Mikami, N., Yang, H., Miyata, M., Nishizaka, T., and Kim, D. (2015). Three-dimensional superlocalization imaging of glicy Mycoplasma mobile by extraordinary light transmission through arrayed nanoholes. ACS Nano 9, 10896–10906.

27. Yoo, H., Lee, H., Rhee, W.W., Moon, G., Lee, C., Lee, S.A., Shin, J.-S., and Kim, D. (2021). Lid sensing nanocomposite islands for nanospeckle illumination microscopy in wide-field super-resolution imaging. Adv. Opt. Mater. 9, 2100211.

28. Yang, Y., Shen, G., Wang, H., Li, H., Zhang, T., Tao, N., Ding, X., and Yu, H. (2018). Interferometric plasmonic imaging and detection of single exosomes. Proc. Natl. Acad. Sci. USA 115, 10275–10280.

29. Son, T., Lee, C., Moon, G., Lee, D., Cheong, E., and Kim, D. (2019). Enhanced surface plasmon microscopy based on multi-channel spatial light switching for label-free neuronal imaging. Biosens. Bioelectron. 146, 111738.

30. Son, T., Seo, J., Choi, I.-H., and Kim, D. (2018). Label-free quantification of cell-to-substrate separation by surface plasmon resonance microscopy. Opt. Commun. 422, 64–68.

31. Lee, H., Kang, K., Mochizuki, K., Lee, C., Toh, K.-A., Lee, S.A., Fujita, K., and Kim, D. (2020). Surface plasmon localization-based super-resolved Raman microscopy. Nano Lett. 20, 8951–8958.

32. Lee, H., Yoo, H., Moon, G., Toh, K.-A., Mochizuki, K., Fujita, K., and Kim, D. (2021). Super-resolved Raman microscopy using random structured light illumination: concept and feasibility. J. Phys. Chem. Phys. 154, 144202.

33. Yu, H., Peng, Y., Yang, Y., and Li, Z.-Y. (2019). Plasmonic-enhanced light–matter interactions and applications. NPJ Comput. Mater. 5, 1–14.

34. Kasari, S., Hagemann, M., Fenhagen, J., and Scherman, O.A. (2014). Quantitative multiplexing with nano-self- assemblies in SERS. Sci. Rep. 4, 6785–6786.

35. Wang, S., Dong, H., Shen, W., Yang, Y., Li, Z., Liu, Y., Wang, C., Gu, B., and Zhang, L. (2021). Rapid SERS identification of methicillin-susceptible and methicillin-resistant Staphylococcus aureus via aptamer recognition and deep learning. RSC Adv. 11, 34425–34431.

36. Cheng, N., Chen, D., Lou, B., Fu, J., and Wang, H. (2021). A biosensing method for the direct serological detection of liver diseases by integrating a SERS-based sensor and a CNN classifier. Biosens. Bioelectron. 186, 113246.

37. Rahman, A., Kang, S., Wang, W., Huang, Q., Kim, I., and Vikesland, P.J. (2022). Lectin-modified bacterial cellulose nanocrystals decorated with Au nanoparticles for selective detection of bacteria using surface-enhanced Raman scattering coupled with machine learning. Int. J. Nurs. Sci. 9, 259–267.

38. Yang, Y., Zhai, C., Zeng, Q., Khan, A.L., and Yu, H. (2020). Multifunctional detection of extracellular vesicles with surface plasmon resonance microscopy. Anal. Chem. 92, 4884–4890.

39. Brown, C., Goncharov, A., Ballard, Z.S., Fordham, M., Clemens, A., Qiu, Y., Rivenson, Y., and Ozcan, A. (2021). Neural network-based on-chip spectroscopy using a scalable plasmonic encoder. ACS Nano 15, 6305–6315.

40. Moon, G., Choi, J.R., Lee, C., Oh, Y., Kim, K.H., and Kim, D. (2020). Machine learning-based design of meta-plasmonic biosensors with negative index metamaterials. Biosens. Bioelectron. 164, 112335.

41. Moon, G., Son, T., Lee, H., and Kim, D. (2019). Deep learning approach for enhanced
detection of surface plasmon scattering. Anal. Chem. 91, 9338–9545.

53. Guselnikova, O., Trelin, A., Skvortsova, A., Ulbrich, P., Postnikov, P., Pershin, A., Sykora, D., Svorcik, V., and Lutukov, O. (2019). Label-free surface-enhanced Raman spectroscopy with artificial neural network technique for recognition photoinduced DNA damage. Biosens. Bioelectron. 145, 111178.

54. Ezra, M., Trelin, A., Guselnikova, O., Dvoranskova, B., Stnadova, K., Perminova, A., Ulbrich, P., Mares, D., Jerabek, V., Elashnikov, R., et al. (2020). Precise cancer detection via the combination of functionalized SERS surfaces and convolutional neural network with independent inputs. Sens. Actuator. B 308, 111760.

55. Lenssen, J.E., Shpacovitch, V., and Weichert, M. (2015). Convolutional neural networks for Raman spectroscopy-based classification in biomedical applications. Opt. Commun. 308, 127660.

56. Kazemzadeh, M., Hisey, C.L., Zargar-Shoshtari, K., Xu, W., and Broderick, N.G.R. (2022). Label-free classification of bacterial extracellular vesicles by combining nanoplasmonic sensors with machine learning. IEEE Sens. J. 22, 1128–1137.

57. Lin, C., Liang, S., Li, Y., Peng, Y., Huang, Z., Li, Z., Yang, Y., and Luo, X. (2022). Localized plasmonic sensor for direct identifying lung and colon cancer from the blood. Biosens. Bioelectron. 211, 114372.

58. Liu, Y., and Zheng, Y.F. (2005). One-against-all multi-class SVM classification using reliability measures. Proc. Int. Jt. Conf. 849–854.

59. Corticone, E., Pfeifer, D., Hentschel, M., Giessen, H., and Tarin, C. (2021). Machine learning methods of regression for plasmonic nanoantenna glucose sensing. Sensors 22, 7.

60. Chio, W.I.K., Liu, J., Jones, T., Perumal, J., Dinish, U.S., Parkin, I.P., Olivo, M., and Lee, T.-C. (2021). SERS multiplexing of methylamine drug isomers via host–guest size matching and machine learning. J. Mater. Chem. C 9, 12624–12632.

61. Yan, S., Liu, C., Fang, S., Ma, J., Qiu, J., Xu, D., Li, L., Yu, J., Li, D., and Liu, Q. (2020). SERS-based lateral flow assay combined with machine learning for highly sensitive quantitative analysis of Escherichia coli. H7. Anal. Bioanal. Chem. 412, 7881–7890.

62. Thrift, W.J., and Ragan, R. (2019). Quantification of analyte concentration in the single molecule regime using convolutional neural networks. Anal. Chem. 91, 13337–13342.

63. Du, Q., Zhang, Q., and Liu, G. (2021). Multiple solution solving in plasmon sensing by deep learning: determination of a layer refractive index and thickness in one experiment. Opt. Lett. 46, 5667–5670.

64. Song, M.K., Chen, S.X., Hu, P.P., Huang, C.Z., and Zhou, J. (2021). Automated plasmonic resonance scattering imaging analysis via deep learning. Anal. Chem. 93, 2619–2626.

65. Thadson, K., Visitsattapongse, S., and Pechprasarn, S. (2021). Deep learning-based single-shot phase retrieval algorithm for surface plasmon resonance microscopy based refractive index sensing application. Sci. Rep. 11, 16289.

66. Xu, Y., Wang, X., Zhai, C., Wang, J., Zeng, Q., Yang, Y., and Yu, H. (2021). A single-shot autofocus approach for surface plasmon resonance microscopy. Anal. Chem. 93, 2433–2439.

67. Van Der Maaten, L., Postma, E., and Van den Herik, J. (2009). Dimensionality reduction: a unified solution for Raman spectroscopy-based classification in biomedical applications. Opt. Spectrosc. 31, 7881–7890.

68. Forestiere, C., Donelli, M., Walsh, G.F., Zeni, E., Miano, G., and Dal Negro, L. (2010). Particle swarm optimization of broadband nanoplasmonic arrays. Opt. Lett. 35, 133–135.

69. Zeng, Z., Venuthurumilli, P.K., and Xu, X. (2021). Inverse design of plasmonic structures with FDTT. ACS Photonics 8, 1489–1496.

70. Sajedian, I., Kim, J., and Rho, J. (2019). Finding the optical properties of plasmonic structures by image processing using a combination of convolutional neural networks and recurrent neural networks. Nanosyst. Nanotechnol. Nanomed. 5, 27–28.

71. Li, X., Shu, J., Gu, W., and Gao, L. (2019). Deep neural network for plasmonic sensor modeling. Opt. Mater. Express 9, 3857–3862.

72. Yan, R., Wang, T., Jiang, X., Zhong, Q., Huang, X., Wang, L., and Yue, X. (2020). Design of high-performance plasmonic nanosensors by particle swarm optimization algorithm combined with machine learning. Nanotechnology 31, 375202.

73. Wiesz, C.R., and Muskens, O.L. (2020). Deep learning meets nearfield nanophotonics: a generalized accurate predictor for near fields and far fields of arbitrary 3D nanostructures. Nano Lett. 20, 329–338.

74. Toh, K.-A. (2018). Kernel and range approach to analytic network learning. Int. J. Networked Distrib. 7, 20–28.

75. Kim, J., Lee, H., Im, S., Lee, S.A., Kim, D., and Toh, K.-A. (2021). Machine learning-based leaky momentum prediction of plasmonic random nanosubstrate. Opt Express 29, 30625–30636.

76. Liu, D., Tan, Y., Khoram, E., and Yu, Z. (2018). Training deep neural networks for the inverse design of nanophotonic structures. ACS Photonics 5, 1365–1369.
94. Lin, R., Zhai, Y., Xiong, C., and Li, X. (2020). Inverse design of plasmonic metamaterials by convolutional neural network. Opt. Lett. 45, 1362–1365.

95. Malkiel, I., Mrejen, M., Nagler, A., Arieli, U., Wolf, L., and Suchowski, H. (2018). Plasmonic nanostructure design and characterization via deep learning. Light Sci. Appl. 7, 60–68.

96. He, J., He, C., Zheng, C., Wang, Q., and Ye, J. (2019). Plasmonic nanoparticle simulations and inverse design using machine learning. Nanoscale 11, 17444–17459.

97. Du, Q., Zhang, Q., and Liu, G. (2021). Deep learning: an efficient method for plasmonic design of geometric nanoparticles. Nanotechnology 32, 055607.

98. Ma, W., Cheng, F., and Liu, Y. (2018). Deep-learning-enabled on-demand design of chiral metamaterials. ACS Nano 12, 6326–6334.

99. Ashalley, E., Acheampong, K., Besteiro, L.V., Ademgil, H. (2021). Generative adversarial inverse design of metasurfaces. Nano Lett. https://doi.org/10.1021/acs.nanolett.1c01280.

100. Wu, Q., Li, J., Wang, W., Zhou, J., Wang, D., Fan, B., Zhang, X., Sun, D., Gong, G., Suolang, S., and Li, B. (2021). Deep neural network for designing near-and far-field properties in plasmonic antennas. Virus 13, 1907–1917.

101. Wu, Q., Li, X., Wang, W., Dong, Q., Xiao, Y., Cao, X., Wang, L., and Gao, L. (2021). Comparison of different neural network architectures for plasmonic inverse design. ACS Omega 6, 23076–23082.

102. Liu, Z., Zhu, D., Rodrigues, S.P., Lee, K.-T., and Cai, W. (2018). Generative model for the inverse design of metasurfaces. Nano Lett. 18, 6570–6576.

103. Zelaci, A., Yashi, A., Kalyoncu, C., and Ademgil, H. (2021). Generative adversarial neural networks model of photonic crystal fiber based surface plasmon resonance sensor. J. Lightwave Technol. 39, 1515–1522.

104. Mirza, M., and Osindero, S. (2014). Conditional generative adversarial nets. Preprint at arXiv. https://doi.org/10.48550/arXiv.1411.1784.

105. Mall, A., Patil, A., Sethi, A., and Kumar, A. (2020). A cyclical deep learning based framework for simultaneous inverse and forward design of nanophotonic metasurfaces. Sci. Rep. 10, 1–12.

106. Radford, A., Metz, L., and Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. Preprint at arXiv. https://doi.org/10.48550/arXiv.1511.06434.

107. So, S., and Rho, J. (2019). Designing nanophotonic structures using conditional deep convolutional generative adversarial networks. Nanophotonics 8, 1255–1261.

108. Ma, W., Cheng, F., Xu, Y., Wen, Q., and Liu, Y. (2019). Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy. Adv. Mater. 31, 1901111.

109. Hain, R., Kahler, C.J., and Tropea, C. (2007). Comparison of CCD, CMOS and intensified cameras. Exp. Fluids 42, 403–411.

110. Bonculet, C. (2009). Image Noise Models. The Essential Guide to Image Processing (Elsevier), pp. 143–167.

111. Ober, R.J., Ram, S., and Ward, E.S. (2004). Localization accuracy in single-molecule microscopy. Biophys. J. 86, 1185–1200.

112. Raschka, S. (2018). Model evaluation, model selection, and algorithm selection in machine learning. Preprint at arXiv. https://doi.org/10.48550/arXiv.1811.12808.

113. Luo, G. (2016). A review of automatic selection methods for machine learning algorithms and hyper-parameter values. Netw. Model. Anal. Health Inform. Bioinforma. 5, 1–16.

114. Zhang, K., Zuo, W., Chen, Y., Meng, D., and Zhang, L. (2017). Beyond a Gaussian denoiser: residual learning of deep cnn for image denoising. IEEE Trans. Image Process. 26, 3142–3155.

115. Tian, C., Fei, L., Zheng, W., Xu, Y., Zuo, W., and Lin, C.-W. (2020). Deep learning on image denoising: an overview. Neural Netw. 131, 251–275.

116. Huang, Y., Zhang, N., and Hao, Q. (2021). Real-time noise reduction based on ground truth free deep learning for optical coherence tomography. Biomed. Opt Express 12, 2027–2040.

117. Karniadakis, G.E., Kevcakis, I.G., Lu, L., Perdikaris, P., Wang, S., and Yang, L. (2021). Physics-informed machine learning. Nat. Rev. Phys. 3, 422–440.

118. Cai, S., Wang, Z., Fuest, F., Jeon, Y.J., Gray, C., and Karniadakis, G.E. (2021). Flow over an espresso cup: inferring 3-D velocity and pressure fields from tomographic background oriented Schlieren via physics-informed neural networks. J. Fluid Mech. 915, A102.

119. Wang, F., Bian, Y., Wang, H., Lyu, M., Pedrini, G., Osten, W., Barbastathis, G., and Situ, G. (2020). Phase imaging with an untrained deep neural network. Light Sci. Appl. 9, 77.

120. Kashinath, K., Mustafa, M., Albert, A., Wu, J.L., Jiang, C., Esmaili-zadeh, S., Aizzadeneshelli, K., Wang, R., Chattopadhyay, A., Singh, A., et al. (2021). Physics-informed machine learning: case studies for weather and climate modelling. Philos. Trans. A Math. Phys. Eng. Sci. 379, 20200093.