Trends in Analytical Chemistry 124 (2020) 115796

Contents lists available at ScienceDirect

Trends in Analytical Chemistry

journal homepage: www.elsevier.com/locate/trac

Deep learning and artificial intelligence methods for Raman and surface-enhanced Raman scattering

Félix Lussier [a, b, 1], Vincent Thibault [c, 1], Benjamin Charron [c, 1], Gregory Q. Wallace [c, 1], Jean-Francois Masson [c, *]

[a] Department of Cellular Biophysics, Max Planck Institute for Medical Research, Jahnstraße 29, 69120, Heidelberg, Germany
[b] Department of Biophysical Chemistry, University of Heidelberg, Im Neuenheimer Feld 253, 69120, Heidelberg, Germany
[c] Département de Chimie, Centre Québécois des Matériaux Fonctionnels (CQMF), and Regroupement Québécois des Matériaux de Pointe (RQMP), Université de Montréal, C.P. 6128 Succ. Centre-Ville, Montréal, QC, H3C 3J7, Canada

ARTICLE INFO

Article history:
Available online 7 January 2020

Keywords:
Deep learning
Machine learning
Artificial intelligence
Artificial neural network
Raman
Surface enhanced Raman scattering
SERS
Sensors

ABSTRACT

Machine learning is shaping up our lives in many ways. In analytical sciences, machine learning provides an unprecedented opportunity to extract information from complex or big datasets in chromatography, mass spectrometry, NMR, and spectroscopy, among others. This is especially the case in Raman and surface-enhanced Raman scattering (SERS) techniques where vibrational spectra of complex chemical mixtures are acquired as large datasets for the analysis or imaging of chemical systems. The classical linear methods of processing the information no longer suffice and thus machine learning methods for extracting the chemical information from Raman and SERS experiments have been implemented recently. In this review, we will provide a brief overview of the most common machine learning techniques employed in Raman, a guideline for new users to implement machine learning in their data analysis process, and an overview of modern applications of machine learning in Raman and SERS.

1. Introduction

Raman spectroscopy is a vibrational spectroscopy which has been extensively applied in the past decades in analytical sciences. Due to the nature of the Raman scattering, a vibrational fingerprint intrinsic to the molecule is acquired, thus enabling the identification of the analyte. As with most spectroscopic techniques, Raman requires advanced data processing to extract meaningful information from spectra. To analyze the optically rich and complex signal, linear regression and multiple multivariate data analysis algorithms have been developed to recognize vibrational fingerprints with high accuracy, sensitivity, and selectivity. Such data analysis typically employed a form of artificial intelligence (AI) to automate data analysis. AI, as far as data science entails, corresponds to the usage of a computing system which applies a mathematical model meticulously developed by the user (*i.e.* hard-coded). For instance, a linear regression is a form of simple AI where strict mathematical rules are applied and followed in order to extract information such as the slope and intercept of a curve. In some cases, an AI system requires the capability to extract its own relevant features, patterns, or knowledge from the data (*i.e.* the machine learns from the data set). This is usually the case when applying AI to vibrational spectroscopy to analyze vibrational fingerprints. This capability of an AI system is referred as machine learning. Thus, it was identified in the 1990s that machine learning methods, could efficiently discriminate between spectra of different molecules [1] and be more effective than linear regression methods for data analysis in Raman spectroscopy [2]. Hence, machine learning algorithms were then applied in fields such as food analysis [3], bacteria identification [4], diagnostic application [5], and material analysis [6].

The interest in machine learning methods has resurged lately, due to the rise of surface-enhanced Raman scattering (SERS) [7], increased computing power, and the availability of open source machine learning libraries (*i.e.* TensorFlow®, Keras and PyTorch) or many application program interface (API) available in commercial multivariate data analysis software. Several characteristics of SERS makes it ideally suited for the use of machine learning data processing. The use of plasmonic materials solved the low sensitivity issue of Raman, and SERS is thus capable of single molecule detection [8,9]. However, the Raman spectra of single molecules show strong variations in intensity and spectral profiles due to the

* Corresponding author.
E-mail address: jf.masson@umontreal.ca (J.-F. Masson).
[1] These Authors contributed equally to the manuscript.

https://doi.org/10.1016/j.trac.2019.115796
0165-9936/

effects of orientation of the molecules with respect to the SERS surface. This leads to the unsuitability of linear based methods (*i.e.* partial least square) to capture the various possible Raman spectra of single molecules. On the contrary, machine learning based methods are well adapted to capture complex relationships within large sets of spectra from single molecule experiments. Machine learning methods can be trained to recognize features in Raman (or SERS) spectra, and assign them to the proper label, which corresponds to the identity of the analyte. Consequently, a plethora of analytical sensors employed a Raman based detection method with machine learning for molecular analysis and chemical sensing, and many of these studies will be discussed later in this review.

1.1. Introduction to machine learning

To identify features and employ them to perform classification or regression, an algorithm will typically employ various forms of machine learning. By definition, machine learning corresponds to a system capable of acquiring knowledge by extracting features from raw data, and then using this newly gained knowledge to tackle problems of the real world by making decisions [10]. For example, simple machine learning models, such as logistic regression, can recommend a caesarean delivery over natural delivery [11], or separate a legitimate e-mail from a malicious one [10]. However, the limited performance of various machine learning models depends mostly on the quality of the features that were presented to the model to perform its classification operation, consequently requiring extensive features extraction and selection. Such limitation was recently overcome using deep learning, a representation-learning based method which autonomously extracts relevant features, and then uses this information to perform classification or regression tasks. This type of model is considered a subset of machine learning, also respecting its definition. To achieve high performance, deep learning models employ multiple levels (or layers) of representation, of increasing abstraction, which ultimately allow models to learn very complex functions and relationships from its inputs. Deep learning models have thus been applied in image [12–14] and speech [15,16] recognition, analysis of data from a particle accelerator [17], predicting the activity of new therapeutics [18], and predicting the effects of potential mutations on non-coding DNA of the genes' expression and diseases state [19,20] to only name a few examples. Thus, it is unsurprising that deep learning has recently been used in many diverse fields, enabling the extraction and analysis of highly valuable information and has attracted other fields of sciences.

1.2. Supervised learning

The most common form of machine/deep learning is known as supervised learning, which implies the training of a model by presenting multiple known examples of the different objects, classes of interest or output (*i.e.* target molecules, disease states, among others) measured in similar conditions. The model then learns by extracting knowledge from the data and performs a generalization of the classification problem. More specifically to the topic of this review, new Raman spectra are sequentially presented to the model, which predicts the corresponding output.

Many supervised learning algorithms have been employed to analyze Raman spectra, which can be separated into distinct sub-methods: (1) methods based on discriminant analysis such as linear discriminant analysis (LDA), partial least-squares followed by a discriminant analysis (PLS-DA), or discriminant function analysis (DFA); (2) artificial neural network (ANN) based models such as the multilayer perceptron (MLP), convolutional neural network (CNN), or support vector machine (SVM); (3) models based on regression analysis like multiple linear regression (MLR), principal components regression (PCR), or partial least squares (PLS); (4) models based on regression trees like random forests (RF), and classification and regression trees (CART); and finally (5) evolutionary based algorithms such as genetic algorithm (GA), genetic programming and computing (respectively GP and GC), evolutionary algorithm (EA), and evolutionary programming (EP) [21]. A review of the most important models is provided in the next section, while the following section will provide the reader with useful information on how to apply these models in their research.

2. Overview of supervised learning methods

2.1. Discriminant analysis-based methods

Linear discriminant analysis (LDA) and discriminant function analysis (DFA), also referred to as canonical variates analysis (CVA), are supervised learning methods largely employed in the field of spectroscopy. These methods are based on the extraction of different hyperplanes, or linear functions, that effectively enable discrimination between two or more classes in multivariate space. These discriminant algorithms search for an optimal hyperplane which maximizes the variance between the different classes, while minimizing the variance within the same class. However, LDA and DFA cannot be directly employed if the data exhibit collinearity or if the number of variables (*i.e.* tracked wavenumbers) is larger than the number of observations [22]. To circumvent this limitation, principal component analysis (PCA) is usually employed to reduce data collinearity and the number of variables [23]. These characteristics make for algorithms that give results quickly and efficiently for spectra that are very similar within a class. However, some accuracy might be lost for unstable or small datasets.

2.2. Artificial neural network

From the different supervised learning methods, models employing ANNs, such as the MLP, were significantly improved since the advent and major development of deep learning [24,25]. In its most fundamental representation, an artificial neuron, or node, corresponds to a basic processing unit, which attempts to mimic the biological functionality of a biological neuron. To achieve this task, artificial neurons compute the sum of a variable number of weighted inputs, and then apply a non-linear function in order to generate an output. By grouping and connecting these neurons together, neural networks are created (see Fig. 1). In a typical neural network, the information is processed and passed in a single direction, from the input layer towards the output layer, where the final output is employed for predictions. Such ANN are referred as a feedforward neural network. In order to perform predictions on realistic and relevant data, the ANN is trained on known examples and the weights, corresponding to the learning parameters, are adjusted in order to minimize the error between the predicted and the true output. To be considered deep, a neural network must be composed of at least three connected neural layers. In a deep neural network, the level of abstraction of the generated output increases as the number of layer increases. This level of abstraction enables the learning of very complex relationships between data, thus conferring high analytical performance (such as high sensitivity and specificity) to deep neural networks for classifying data. However, all types of neural networks have the regrettable tendency to overfit the dataset. More specifically, the network can simply learn the output and fail to generalize the conclusions necessary for prediction. Thus, the network requires (1) a meticulously developed neural architecture; (2) extensive hyperparameter selection and most of all; (3) a rigorous validation of the

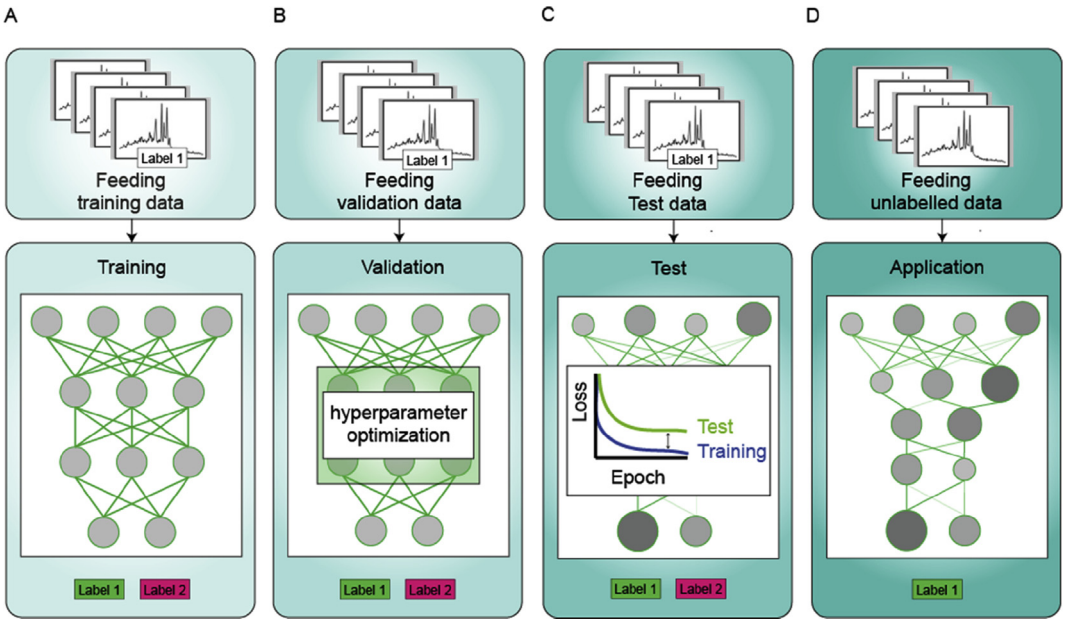

Fig. 1. General scheme presenting the generation pipeline of an artificial neural network applied to vibrational spectroscopy. In a first steps, three independently acquired Raman/SERS spectra dataset (respectively training, validation, and test) are preprocessed identically. Each spectrum composing the various datasets are then labelled to its corresponding identity, defined as label. The training dataset is then presented to the untrained artificial neural network (A). An artificial neural network is composed of artificial neurons, or nodes, arranged and layered to create a complex network capable of extracting information from its input. Herein, the nodes are schematized as grey spheres and correspond to the basic computational units of the network. Weights, the learning parameter of a neural network, are represented as the green lines connecting the various nodes of the network. During the training phase, the predicted output is compared to the true label of the input data, and then used to evaluate the error of the prediction. During the training phase, several metrics may be employed, such as the mean squared error, the recall, or even the differentiable F1 score function where the monitored metric needs to be judicially chosen. The learning algorithm then attempts to minimize, or maximize, the selected metric to improve the neural network performance. As a first try, users might attempt to minimize the error of the network in order to improve its accuracy by using the ADAM optimizer, a commonly employed learning algorithm. Jointly with the training phase, the validation dataset is presented to the network to evaluate the choice of hyperparameters (e.g. number of nodes, number of layers, types of layer such as convolutional or densely connected, learning rate, regulation coefficient, optimizer, etc.) (B). The step of validation may be performed by using algorithms such as grid/random search or Bayesian optimization in order to find the optimal set of hyperparameters that maximize the performance of the network. Typically, the hidden layers will imply the most important changes during the process (B). The trained model composed of an optimal number or nodes, layers and weight values, represented as darker green lines and/or grey spheres is employed to predict the output on the test dataset. (C). Herein, the network has been optimized for the labels presented during training and validation. If the performance of the network is sufficient (i.e. the distance between the loss (error) on the training set and the test set is at its minimal), the artificial neural network can be deployed for applications (D).

trained model on independent datasets. This typical checklist has been extensively reviewed elsewhere [26]. Hence, ANNs offer high performance regarding complex datasets. The major drawback is that the user must be careful when using them as it is easy to treat these models as “black boxes” and misinterpret poor or incorrect results as acceptable. A good knowledge of the theory behind the used algorithm and the origin of the dataset is necessary to truly understand where the issues may come from. Furthermore, these deep learning methods require more time and/or computing power. These algorithms can also, if the user is not careful, be trained towards a local minimum rather than an absolute solution, a common problem of stochastic models and non-convex optimization.

Beside the conventional feedforward architecture, various neural networks, deep or not, presenting different architectures have been applied to Raman and SERS data analysis. CNNs, which employ shared weight filters and pooling layers within their architecture, have shown a higher specificity and sensitivity compared to conventional MLPs [27]. CNNs were recently applied to SERS optophysiology measurements to analyze metabolite gradients near various cell lines, including HeLa and HUVEC [28], thus enabling the acquisition of up to nine molecular gradients simultaneously. Mixture analysis by Raman spectroscopy was also performed by employing a CNN, which showed improved sensitivity and specificity compared to other machine learning methods such as k-nearest-neighbor (kNN), RF, and logistic regression (LR) but also compared to other neural network architectures like fully connected neural network and backpropagation based artificial neural network [27]. CNN models were recently applied to identify tongue squamous cell carcinoma based on the distinct vibrational signature between healthy and cancerous cells, thus presenting the promising potential of machine learning based technique for intraoperative evaluation of the margin of resection site. A CNN model was also employed for pattern recognition of raw SERS spectra [29]. While this list of examples is not exhaustive, it demonstrates the various use of CNN in Raman spectroscopy. Prior knowledge gained by a CNN could serve to facilitate its adaptation to new uses. By generating open-access spectral library, CNNs could be perceived as a powerful candidate for component identification by employing transfer learning, corresponding to the process of training a model on new data based on a previous training. Hence, the initial values of the learning parameters are already close to the optimal value and thus training is accelerated.

2.3. Support vector machine

Support vector machine (SVM) is an optimization-based discrimination model, which attempts to generate optimal hyperplanes, or decision boundaries, that best separate different classes in a high dimensional space. These hyperplanes are typically constructed by analysing data points which will most likely be misclassified (*i.e.* data points near the candidate hyperplanes). These data points, referred to as support vectors, are then iteratively weighted during the training phase in order to maximize the distance (*i.e.* margin) between the classes separated by the hyperplane [30]. Seeking classes' boundaries in the form of hyperplanes is

adequate in cases where the different classes are linearly separable. To handle more complex segregation, the original input space, where classes are not linearly separable, is transformed into a new higher dimensional feature space where samples are projected by using a feature (*i.e.* kernels) function [30]. This powerful manoeuvre, known as the *kernel trick,* allows SVM to seek optimal separating hyperplanes in higher dimensions where classes are thus linearly separable. An online library containing various support vector machines' kernels, referred as the LIBSVM, is available online for the interested readers [31]. As opposed to ANNs, SVMs are deterministic and their solution is global and unique. They also have the advantage that they do not overfit like ANNs. However, the size of the solution is not defined. The support vectors being selected from the training data, it requires sufficient data to avoid the issue of a number of vectors necessary to separate the data equal to the amount of data. Picking the best kernel for the problem is also trial and error.

SVM was employed to analyze SERS data in various fields of applications such as the quantification of thiophanate-methyl and its associated metabolites carbendazim in red bell pepper (*Capsicum annuum* L.) [32]. It was also utilized to detect labelled nanoparticles through various bones thickness by surface-enhanced spatially offset Raman spectroscopy (SESORS) demonstrating the capability of SESORS and SVM for a minimally invasive *in vivo* sensor [33]. SVM models were also developed to detect DNA damaged after x-ray radiation on nasopharyngeal carcinoma cell line (CNE2) [34], and the presence of age-related macular degeneration and/or diabetic macular edema in aqueous humors for point of care diagnostics [35]. The combination of micro-Raman and SVM enabled the identification of various pathogenic bacteria [36], originating from various cultured conditions [37], and the rapid identification of *staphylococcus* strains [38].

2.4. k-nearest neighbor

One of the simplest forms of machine learning based method corresponds to the *k*-nearest-neighbor (kNN) which can be applied for both regression and classification problems. For classification purposes, the kNN assumes that similar objects will exist in close proximity in their feature space. Thus, presented samples are classified based on the *k* number of examples of the different classes located in the closest neighborhood in this feature space. Weights, which typically corresponds to the *Euclidean* distance, are used in order to evaluate which classes are the closest to the unknown example. However, alternative methods such as the Manhattan, Chebyshev and Hamming distance may be used for evaluating distance between data within the feature space. Although easy to implement, the memory usage increases significantly as the number of class dimension increases. The combination of Raman spectroscopy with kNN has been implemented in studies involving detecting different types of cancers, such as breast [39] and colon [40,41].

2.5. Random forest

The RF classification algorithm belongs to a supervised learning method referred to as decision trees. In the case of RF, multiple uncorrelated decision trees are generated to form the RF classifier. By using a process referred to as bootstrapping, also called bagging, the same number of input data with replacement is presented to each decision tree to train the RF model. Each tree then selects features that maximize separation of the dataset in two classes. Each and every tree differ in its selected feature as they all have a randomly assigned dedicated training dataset. Subsequently, new examples are presented to the RF model in which each tree will sort the example based on its own selected features. Representative architectures are shown in Fig. 2A. The output of each decision tree is then pooled, leading to the final prediction of the model. The high analytical performance, such as high sensitivity and specificity of the RF classifiers, resides in the usage of multiple uncorrelated trees to sort the presented examples, where each uncorrelated tree circumvents the individual prediction error of the other trees. RF offers several advantages, which were presented elsewhere [42]. Specifically, RFs are robust to outliers, noisy data [22] and relatively tolerant to overfitting, in addition to the ability of handling data without the prerequisite of preprocessing. The principal disadvantages of RFs relate to their complexity. As additional trees are added for finer classification, the architecture of the forest becomes more elaborate. When this happens, greater computational requirements are needed, both for training and prediction. Additionally, RFs are especially not well suited for extrapolation. In a recent example, RF was used to analyze SERS signals of accumulated lipids such as sphingomyelin and cholesterol in lysosomes resulting from exposition to various tricyclic antidepressants [43]. In another study, RF was utilized in a SERS-based immunoassay [44]. The specific architecture for that classification tree is shown in Fig. 2B.

3. Getting started with machine learning in Raman spectroscopy

Now that the different models have been presented, this section of the review aims at defining some common language in machine learning as well as providing some advices to ease the entry barrier to apply machine learning for Raman data analysis. Going in-depth on every machine learning algorithm is beyond the scope of this review and therefore, most of the further discussion will focus on supervised learning methods as these are more common in contemporary Raman spectroscopy. More in depth resources are available elsewhere [10,45,46].

3.1. Selecting the right tool

As discussed in section 2, the various machine learning algorithms all have pros and cons. The first step is to choose the appropriate method for a dataset. This can be a tedious task as there is no absolute method, unless the algorithm has mathematical constraints with respect to the input. Experimenting with the different methods is appropriate for selecting the best method for a specific problem [47,48]. A great starting point is to distinguish between supervised and unsupervised learning. Unsupervised learning, in the case of characterization, usually refers to clustering methods that do not require any labels. Spectra are separated based on a space where every pixel is considered a dimension, effectively making those spectra points in that space. Those points can then be grouped in the space through various mathematical functions. K-means clustering optimizes centroid centers to group the spectra while distribution-based and density-based algorithms optimize the shape of the clustering space. However, in order to identify what is inside a cluster, some knowledge about the data is needed. For example, if the user knows each type of spectra to be assigned, a known standard for each of them can be clustered with the data to determine what each cluster represents [49,50]. This is a good and efficient method when the data have well defined differences, even if the said differences cannot be seen in other ways. Otherwise, overlapping of the results might occur for spectra that are not as well defined. If correctly labelled data is available, supervised learning is likely a better choice as the program will know to look for differences in spectra with different labels rather than blindly

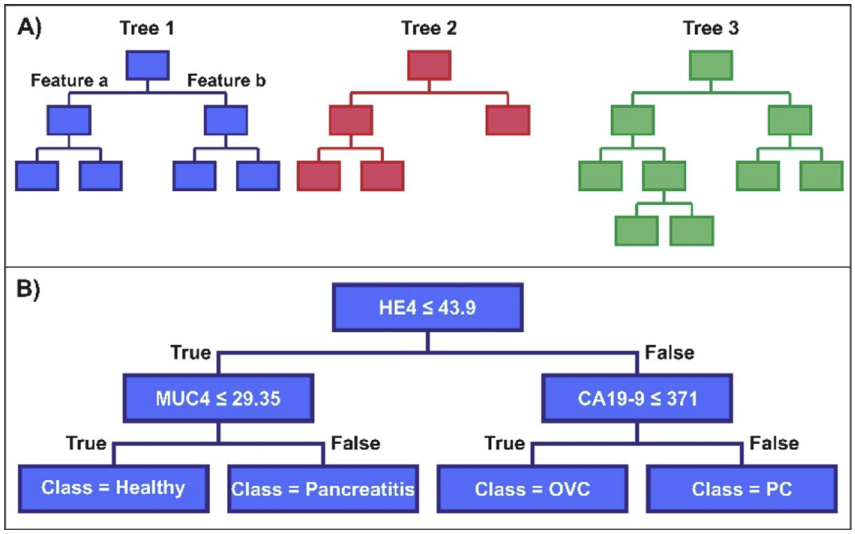

Fig. 2. (A) Random forest (RF) classifiers originate from the generation of multiple uncorrelated decision trees. Each tree separates the presented data by the means of randomly selected *features*, whom maximize the separation of the data into two subsets of data with similar size leading to the generation of various branches. Each uncorrelated and independent tree will have its own set of features, leading to the generation of trees exhibiting different architectures. (B) Classification tree for a SERS-based immunoassay that used in the intensity of a Raman reporter with each biomarker to determine if the sample was healthy or in a disease state [44]. HE4: Human epididymis protein 4. MUC4: Mucin 4. CA19-9: Carbohydrate 19−9. OVC: Ovarian cancer. PC: Pancreatic cancer. Reprinted and adapted from Ref. [44]. Published by The Royal Society of Chemistry.

searching all the data for similarities and potentially finding incoherent differences instead [4].

Another important aspect is how complex the spectra are. Machine learning algorithms study features of the input spectra. If the spectra of a same class can change over time, like in single molecule SERS, or if features from various classes overlap, a more in-depth algorithm (CNN, among others) would be preferable in order to not only extract common features, but also different possibilities. On the contrary, for spectra that are very consistent, like Raman spectra of bulk material, a more basic algorithm (MLP, PLS, among others) can be as efficient without requiring much effort for an easier and faster classification.

Additionally, feature reduction algorithms such as PCA can be implemented before the main algorithm, whether it is a supervised or unsupervised method. The choice to use it or not depends on the information contained in the spectra. Again, for complex spectra that can vary such as in the case of single molecule SERS experiments, important features could be lost in feature reduction. For very similar spectra however, feature reduction can eliminate features that do not help separation and could make the training and processing much faster. For choosing a specific architecture, the best measure would probably be to get inspired by the general shape of one from the winners of a software contest, such as Kaggle (https://www.kaggle.com/competitions) or ImageNet Large Scale Visual Recognition Challenge (http://image-net.org/challenges/LSVRC/) and tweak it to towards a specific application. There are many 2D classification models and anomaly detection models that can be converted to 1D to accept Raman spectra with optimization and following those type of contests will allow users to see hands-on usage of the latest developments in machine learning.

3.2. Gathering data

The quality of a dataset is critical, referring to the proverbial "Garbage in, garbage out" in data science. It points out the fact that models based on false or flawed data will yield false or flawed models as well. This is one of the greatest risks of machine learning. As mentioned in section 2, flawed models usually take the form of a model that overfits its training dataset but could also simply be a model that tries to find a correlation when there is none. This is avoidable through the acquisition of proper training data, careful selection of proper hyperparameters (top-level or user-defined parameters), and a robust validation method. The latter two will be discussed in section 3.3.

Datasets should basically be acquired in conditions as similar as possible to the dataset that will require classification or regression through the machine learning algorithm. Depending on the algorithm, spectral range and center could be perceived and influence the algorithm. A change in background could also have an impact and change the class to which a specific spectrum will be assigned. Therefore, the camera's temperature or integration time, the sample's form (solid or in solution) and the condition of analysis (continuous flow or static measurement, SERS or normal Raman, among others) should also be as similar as possible. If data were acquired on different instruments or at different times, it is possible to adjust for some hardware changes with preprocessing of the spectra. Preprocessing of all spectra, whether it is the training set or the data to identify, should be exactly the same in order to obtain relevant results. Common preprocessing steps include smoothing, baseline correction, cosmic ray removal, cropping and normalization [51]. It is however important to remember that preprocessing also means removing some features from the raw data. Nearly indistinguishable features might be hidden close to the baseline and although unseen by the naked eye, may be picked up by a machine learning algorithm. Although some machine learning algorithms benefit from a proper preprocessing, others like MLPs, are usually not affected by it [29]. Here again, trial and error will show what is best for each case, but it is important to never assume preprocessing will yield better results as it is often not the case [51].

3.3. Training and evaluating the model

Supervised learning will always require a training dataset. A test set is another dataset that is generally mentioned in the literature. Finally, some reports will mention or use a validation dataset. The difference between these three can be confusing to a machine learning beginner as validation and test sets are sometimes wrongly used interchangeably to describe the same data. The

training set is used by the algorithm to extract meaningful information towards classifying the spectra. The validation set is used when optimizing the hyperparameters. The user will train models with different hyperparameter values, measure their accuracy on the validation set (*i.e.* how correct is the identification of the new data), and select the best value selection to move forward based on these values. Although the algorithm never learns from the validation data, a bias is added by the user selecting the best values for that set. A different user trying to reproduce that accuracy on another dataset with the same model might not be able to achieve it as the values may not generalize the features properly and simply fit the validation data better. This leads to the test set which is similar to the validation set but is absolutely independent from the model's training and selection. It should only be used at the end to report a purely unbiased accuracy of the model. Data scientists should always use a proper test set to back their claims. Although it would be best to have three fully independent datasets that are acquired in the same manner, but during different experiments, they are usually extracted from the same large dataset as this is more convenient. A model that would achieve great separation of the training/validation data but identify test spectra poorly is most likely to have experienced overfitting of the training set. To analytical scientists, the training set can be seen as the prepared standards of a calibration curve and the test set is equivalent to a quality control sample bought and tested by an external supplier. The relative proportion of one to another is not an exact science. In some algorithms, a good training will require a lot of training spectra while others may only require a few, depending on the noise level of the data and the number and complexity of the features it needs to model. In a similar fashion, a model that has many internal parameters, such as an ANN's weights and biases, will be more difficult to test and will require bigger validation and test datasets while a model with few to none will need smaller sets. Large datasets are normally preferred, as it is always possible to subdivide it into smaller datasets if needed. Generally, the validation and test sets contain a number of spectra corresponding to between 10% and 40% of the training set [45].

Model type and hyperparameter tuning is another important task of data analysis process. Unfortunately, only trial and error will give a definitive answer as to which conditions are best for a specific case. A few tricks can however ease up the process. It is a good idea before optimizing an algorithm to research what parameters and layers can be implemented as some optional inputs can result in immense improvements of the final algorithm. As an example, ANNs can use a certain amount of dropout. This type of layer will randomly deactivate neurons during training to ensure the final model does not rely only on just a few nodes and overfits the data because it's relying on so few features [52]. In order to find the best values for each parameter, interesting approaches can be a grid search, a random search, or a Bayesian optimization of various values on a single dataset, but sometimes at the expense of being time consuming.

Often, the test set's results will be classified in a confusion matrix, a table representation of the actual labels to the assigned labels (Fig. 3A). This allows the extraction of the probability of true positive (TP), false positive (FP), true negative (TN) and false negative (FN), giving great insight in the capabilities of the current model. These values can be used to calculate various evaluation metrics, most notably the accuracy. It is the most commonly used metric but is often misleading when used alone and when data is not evenly distributed in all classes [53]. For example, a binary model simply attributing all data to a single value would be right 99.9% of the time if 99.9% of the spectra belonged to this class. It would however solely rely on statistics to assign a spectrum to a single class and not on the spectrum's features. These first level metrics can then be combined further into the F1 score or a receiver operating characteristic curve (ROC curve). The F1 score is a metric which considers both how often the model is right when it should be right and how often the model is right when it says it is right (Fig. 3A). The ROC curve shows the general overview of the model and gives an area under the curve (AUC) which represent the trade-off between the recall and specificity. The ROC curve is excellent at conveying information about binary classifiers but fails to do so in multiclass models. Confusion matrices are more suitable in that last case, since they help visualize how good a model is for each class. An example of a confusion matrix for staphylococci strain identification is included in Fig. 3B [38]. Overall, the confusion matrix is a good way to convey the general quality of the model, but the ROC curve and other metrics can be more relevant in some situations.

There are many means of evaluating the performances of a model, oftentimes performed after training, as mentioned earlier, but also throughout its learning phase. It is possible to simply use the result of the loss function to determine if the model is converging and its quality relative to another model. The model can also use a specific and absolute quality metric to determine which values of the internal parameters will be kept without resorting to the loss value, which is relative. Whichever metric is chosen, it will always be easier to converge to its optimal value if the loss function has a good correlation with it. For example, if one wants to optimize a model based on the F1 score, there is an optimization function based on this quality metric that can be used as a loss function inside the learning algorithm. This function guides the program's training towards a model with the best possible F1 score [54]. The AUC of the ROC curve is also a good option for balanced dataset. Using cancer identification as an example, falsely identifying a scan as non-cancerous and letting a tumor grow is costlier than falsely identifying a scan as cancerous and having more tests needed. Hence, a metric like the F1 score, which penalizes false negatives more heavily, would be preferable. A similar idea can be implemented in the loss function. An example would be a function where the loss is proportional to the classes' size. An error on a class containing less data would be penalized more strongly then larger classes, but a proper assignment on this same class would also be rewarded more greatly then the larger one. It is recommended that users consult the various communities of users online in the optimization of the models for their applications.

4. Applications

In this section, we will highlight existing and emerging applications involving the combination of deep learning and artificial intelligence methods with Raman and/or SERS. We have chosen to focus on four of the most common categories: food and beverage, forensics, bacteria and viruses, and medical diagnostics. This is by no means a list of all possible applications, it is simply designed to demonstrate the broad scope of applications.

4.1. Food and beverage

One means of cost-saving in the food industry is to use ingredients that are less expensive. In 2013, some meat products labelled as containing beef in parts of Europe were found to contain or made entirely of horse meat. The ability to therefore identify the origin of a food product, such as meat, can potentially avoid such events. For poultry, it was previously demonstrated that the leg and breast meat of turkeys and chickens can be distinguished from one another [49]. This idea can be expanded to other common foodstuffs, such as margarine in butter [55], and luxurious food items, notably caviar [56]. Fraud in caviar production often occurs by mixing roe from different species of sturgeon, or by mixing roe obtained from low-cost farmed sturgeon with that of wild sturgeon. The latter is particularly advantageous as although the diets of the fish are quite

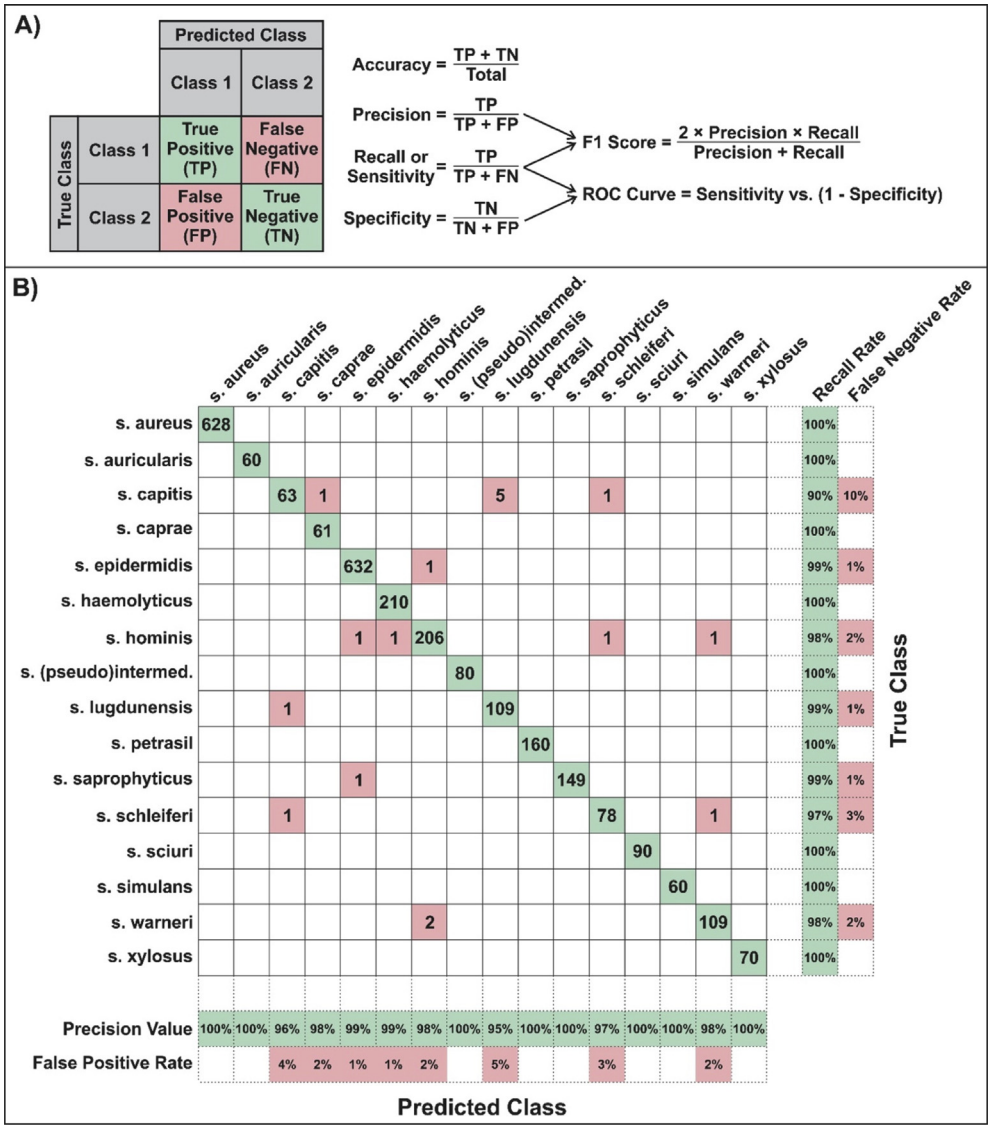

Fig. 3. (A) Representation of a confusion matrix (left), the first level metrics that can be extracted from it (middle), and the more efficient second level metrics formulas (right). (B) Confusion matrix for the result of 5-fold cross-validation of bacterial strain identification [38]. The numbers in the cells correspond to the number of spectra that were correctly classified (diagonal, green) or misclassified (off-diagonal, red). Reprinted and adapted from Ref. [38]. Published by Spring Nature Limited.

unique, by appearance, the caviar does not appear different. Beyond the cost of a product, the origin may also influence its nutritional value. It has been shown that milk obtained from humans, cows, buffalo, and goats can be distinguished from each other by Raman spectroscopy (Fig. 4A) [57]. The variations in the Raman spectra that allowed for this differentiation were attributed to the wide variation in macronutrients present in the milk of different animal species. Included in the inset of Fig. 4A is the out-of-bag classification error analysis that the authors used to determine the number of trees that were needed to make the model computationally efficient while maintaining low statistical variance with the model (n = 100).

Detection of molecules that have been added to food, can be improved upon using vibrational spectroscopy. Synthetic dyes are often intentionally added to foodstuffs to enhance the colour. In some areas, their use is regulated, and for some dyes, even banned. As dye molecules are often used to evaluate the performance of SERS structures, detection of dyes in foodstuffs by SERS is a natural progression. Dyes that have been studied by SERS coupled with multivariate analysis include: the known carcinogen Sudan-1 [58]; carmine [59]; food blue, tartrazine, sunset yellow, and acid red [60]. In the latter study, the authors collected SERS spectra for each of the dyes at concentrations ranging from 10^{-3} to 10^{-8} M. SERS spectra for tartrazine and sunset yellow are shown in Fig. 4B. PCA was then used to create fitting curves for quantitative detection of the dyes (Fig. 4B). The use of pesticides and fungicides in agriculture leads to the exterior being contaminated with these molecules. Ensuring that these molecules are removed is important to the health of the consumer. As these are surface contaminants, different approaches have been used to determine if pesticides or fungicides are present and if possible, to quantify the amount. Methods include extraction from a dissolved sample [61], swabbing the surface and releasing the contents of the swab [62], and, adding and aggregating silver nanoparticles directly on the surface [32]. In that study, the quantity and distribution of both the fungicide (thiophanate-methyl) and its metabolite (carbendazim) could be evaluated.

4.2. Forensics

In addition to characterizing the properties of pharmaceutical drugs [63–65], the detection of illicit drugs can benefit from the

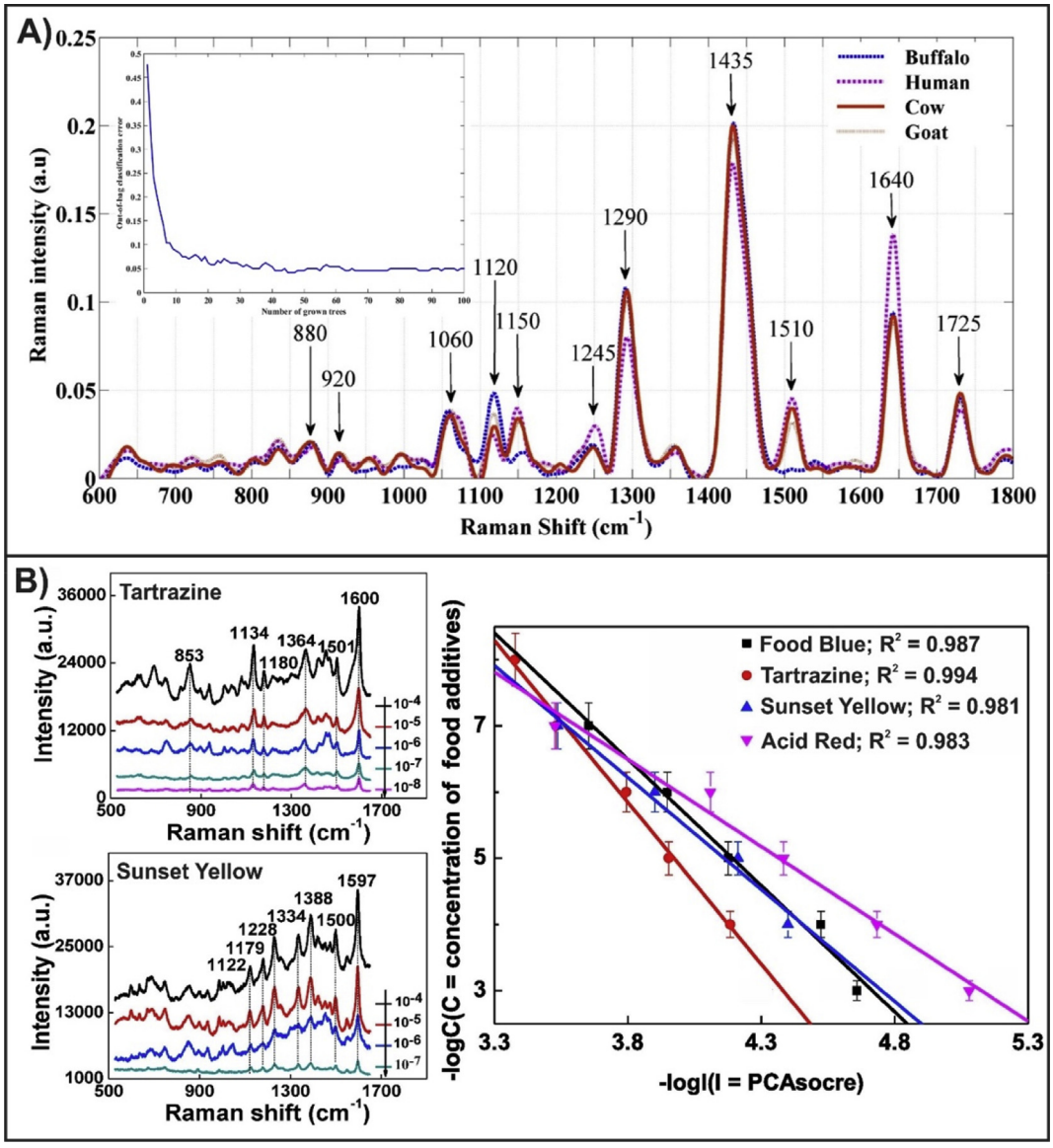

Fig. 4. (A) Mean vector normalized Raman spectra of milk samples from four animal species [57]. Included as an inset is the trend of out-of-bag classification error against the generated number of trees. (B) SERS spectra of different concentrations (mol/L) of tartrazine (top left) and sunset yellow (bottom left) adsorbed onto flower-shaped silver nanoparticles [60]. PCA score fitting curve for four food colourants (right). Equations were: food blue ($y = -3.8157x + 20.994$); tartrazine ($y = -4.8897x + 24.424$); sunset yellow ($y = -3.382x + 19.07$); acid red ($y = -2.6361x + 16.51$). Reprinted and adapted with permission of Ref. [57,60]. Copyrights 2018 Elsevier.

combination of Raman, and more commonly SERS, with multivariate analysis. For example, it was shown that oxycodone, heroin, tetrahydrocannabinol, and cocaine could be readily distinguished by a model combining PCA and SVM (Fig. 5A) [66]. The authors further demonstrated quantification of cocaine drug samples, and that their proposed method could detect cocaine in saliva at a concentration lower than the cut-off for federal workplace testing in the United States (Fig. 5A). Here, we continue to highlight examples involving detection in complex biofluids. These works include the detection of methamphetamine and heroin in saliva [67], tetrahydrocannabinol in saliva [68], and, morphine and different fentanyls in urine [69], all of which involved the use of PCA. The latter study also used a PLS-DA algorithm. For the detection of methamphetamine in saliva [67], the developed model with two principal components could accurately test against a positive control (methamphetamine-positive saliva at 0.5 μM) and a negative control (heroin-positive saliva at 50 μM). The detection of tetrahydrocannabinol used two models each having three principal components [68]. They first demonstrated that SERS spectra of tetrahydrocannabinol (1 pM) in methanol, plasma, and saliva could be differentiated from one another. More interestingly, the second model was then used to predict the metabolism of tetrahydrocannabinol in saliva (10 μM). A 5-fold cross-validation was used to test the metabolism prediction model. The calibration and test datasets achieved root mean square errors of 0.9318 and 1.2833 respectively (R^2 values of 0.9820 and 0.8305). A series of models each with two principal components were used for the detection of fentanyls in morphine, the main metabolite of heroin, for SERS spectra obtained using dry and wet preparations [69]. In urine samples, fentanyl could be determined down to 50 ng/mL from 10,000-fold higher concentration of morphine. The sensitivity of five fentanyls (fentanyl, carfentanil, 4-fluorobutyryl fentanyl, norfentanyl, and remifentanil) ranged from 50 to 2000 ng/mL.

Beyond illicit drugs, other areas of forensic sciences can benefit from Raman spectroscopy [70]. Crime scene analysis is one, though here we will highlight some of the different types of information

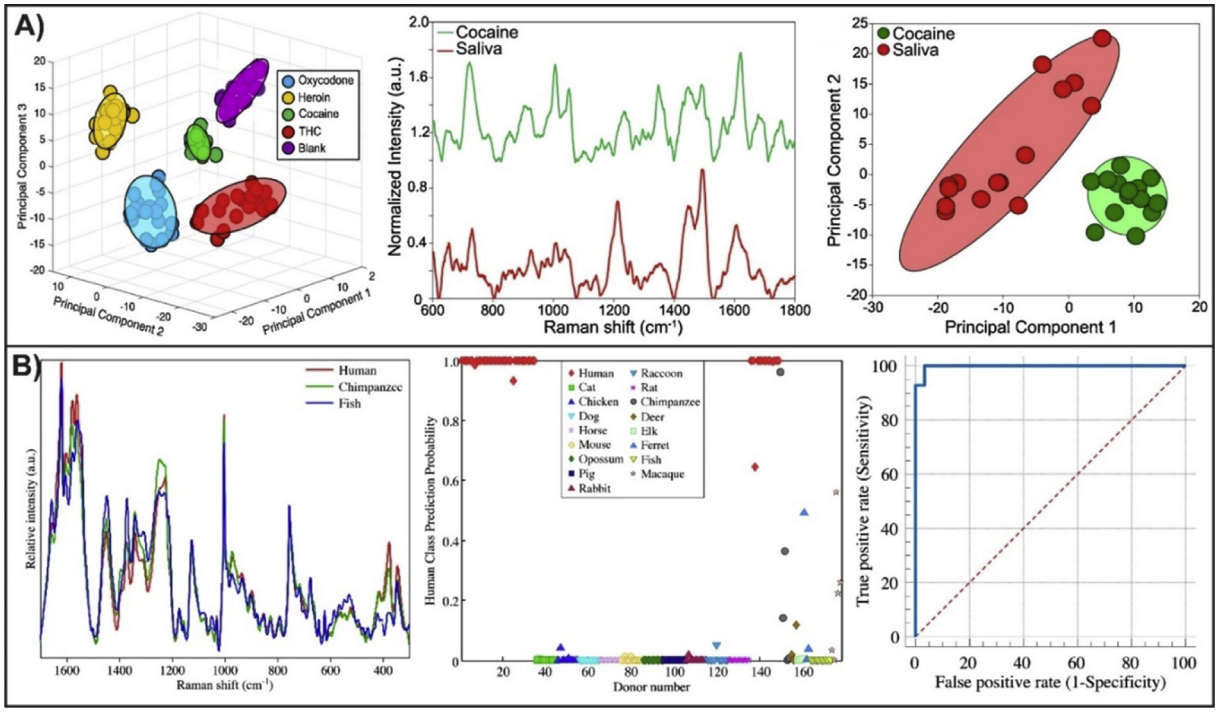

Fig. 5. (A) Plot of first three principal components used to cluster the spectra from different illicit drug analytes for identification (left) [66]. SERS spectra for processed cocaine-containing saliva, as well as processed saliva (middle). Plot of first 2 PCs for SERS spectra of cocaine-spiked and unspiked saliva (right). All clusters consist of 20 spectra, and ellipses indicate 95% confidence intervals. (B) Raman spectra of human, chimpanzee (most genetically similar to humans), and fish (least genetically similar to humans) donors overlaid for enhanced visual comparison (left) [71]. Scores plots generated from the binary PLSDA model constructed to differentiate human blood from animal blood showing both the prediction probability values for the internal calibration spectra (spectra 1–135) used to build the model, the original human donors used for external validation (spectra 136–139), the new human donors used for external validation (spectra 140–149), and the six new animal species used as external validation (spectra 149–178) (middle). A "score" of 1.0 correlates to 100% probability that a spectrum was predicted as human blood. The ROC curve showing the true positive rate values (y-axis) versus the false positive rate values (x-axis) where an area under the curve (AUC) of 0.99 was achieved (right). The solid blue line is the actual ROC curve generated from the human class prediction values of the external validation spectra. The red diagonal dotted line represents the 0.5, or random chance, line. Reprinted and adapted with permission of Ref. [66,71]. Copyrights 2018 Elsevier.

that can be obtained by analyzing a bloodstain. Understanding the source of the blood (animal vs. human) is an important and sometimes overlooked piece of information [71]. In that work, blood samples from 16 species of animals were compared to that of human blood. A comparison of Raman spectra of human blood with that chimpanzee and fish is shown in Fig. 5B. The developed PLS-DA model correctly predicted all external human blood donor samples as human, and 28 of 29 animal blood donor samples as nonhuman. The ROC curve in Fig. 5B had an AUC of 0.99, indicating that for a randomly chosen sample, the PLS-DA model only had a 1% chance of incorrectly predicting a nonhuman blood sample as being human. Other characteristics of bloodstains of human origin that have been explored include gender [72] and selected age groups [73]. Under optimized chemometric conditions, 59 of 60 samples were correctly classified based on gender [72], and the lowest cross-validated sensitivity and specificity values were 0.96 and 0.97 respectively for determining age grouping [73]. In addition to blood, an earlier work also found that gender classification by Raman and chemometrics could be obtained using fingernail clippings [74]. Here, the classification accuracy for males and females by leave-one-out cross-validation was 87.5% and 93.3% respectively. Although all these results are highly encouraging, especially those of the bloodstain analysis, there are three important aspects that need to be developed further: (i) bloodstains on different types of surfaces and materials, (ii) the use of hand-held Raman spectrometers for field-deployability, and (iii) combined chemometric parameters. The latter is especially important as it would demonstrate how combined variables (such as gender and age) influence the sensitivity and specificity.

Chemical forensics can be used to determine the origin of compounds found at a crime scene. One such example is the determination of the synthetic route used to prepare mustard gas, a widely recognized chemical warfare agent [75]. In that study, 8 syntheses were compared, with the subsequent Raman measurements performed with the samples inside glass vials after the samples had been aged for six months. The Raman spectra showed variability within the replicate measurements obtained from the same synthetic route. Further hindered by the small number of samples, only 78% of the spectra (18 of 23) were correctly classified within the test set. This was marginally improved by coupling Raman and IR measurements (83%), but nevertheless demonstrates that such analysis is feasible.

As opposed to putting a physical barrier (glass vial) between the sample and the instrument (and operator), one can consider using distance. By developing a stand-off hyperspectral Raman imaging system that incorporated the use of a telescope, it was possible to safely acquire Raman spectra of explosives at a distance of 15 m [76]. With this instrument, a sample holder with a side length of 80 mm can be scanned from 800 to 3101 cm^{-1} in approximately 50 min. This study explored five pure explosives (ammonium nitrate, sodium chlorate, pentarythritol tetranitrate, 1,3,5-trinitro-1,3,5-triazinane, and trinitrotoluene), three plasticized samples each containing one of the pure explosives, and the commercially available explosive Rowodyn. Using an initial random forest that considered the full spectral range, the out-of-bag average errors ranged from 1.44% (sodium chlorate) to 7.82% (pentarythritol tetranitrate). By creating a new RF that only considered a single band that would be assigned to each explosive, the measurement time

decreased by a factor of 10, and the out-of-bag average errors improved (0.37% for sodium chlorate, 5.2% for pentarythritol tetranitrate).

4.3. Bacteria and viruses

In Raman microspectroscopy, the diameter of the laser beam is typically in the range of 1 μm. As this is comparable to the size of many bacteria, Raman microspectroscopy is a valuable technique for the analysis of individual bacteria. As a result, spectral databases of bacteria can be built. However, it is important that such databases consist of a large number of spectra per type of bacterium as a considerable degree of variability can occur based on the growth stage and conditions [37]. For example, when comparing 11 bacterial species that can potentially cause urinary tract infections, an average of 268 spectra were acquired for each species during construction of an SVM [77]. 10-fold cross-validation showed an accuracy of 92% for the developed model. The model was subsequently able to correctly predict the dominant species present in 10 infected urine samples. For each sample, the number of spectra assigned to the dominant species ranged from 66.7% (40 of 60 spectra) to 97.8% (45 of 46 spectra). As the infection may also be caused by the presence of more than one species of bacteria, evaluating combinations of bacteria is also important [78]. A sample containing two defined species was prepared along with a separate sample containing three species. With the binary mixture, 19 of 20 spectra were labelled to the species present, with 29 of 30 spectra labelled in the same manner for the trinary mixture. Analysis of bacteria can be rapidly expanded by examining strains of species [38]. This was done for 277 strains from 16 species of *staphylococci*. Here, the idea is not to be able to identify the specific strain, but to improve the database for each species by incorporating as few as 6 strains (*S. auricularis*, *S. caprae*, *S. simulans*) to as many as 63 (*S. epidermidis*). It is important to note that for this study, a minimum of 10 spectra from 3 bacterial colonies were used for each strain. The resulting model, by 5-fold cross-validation had

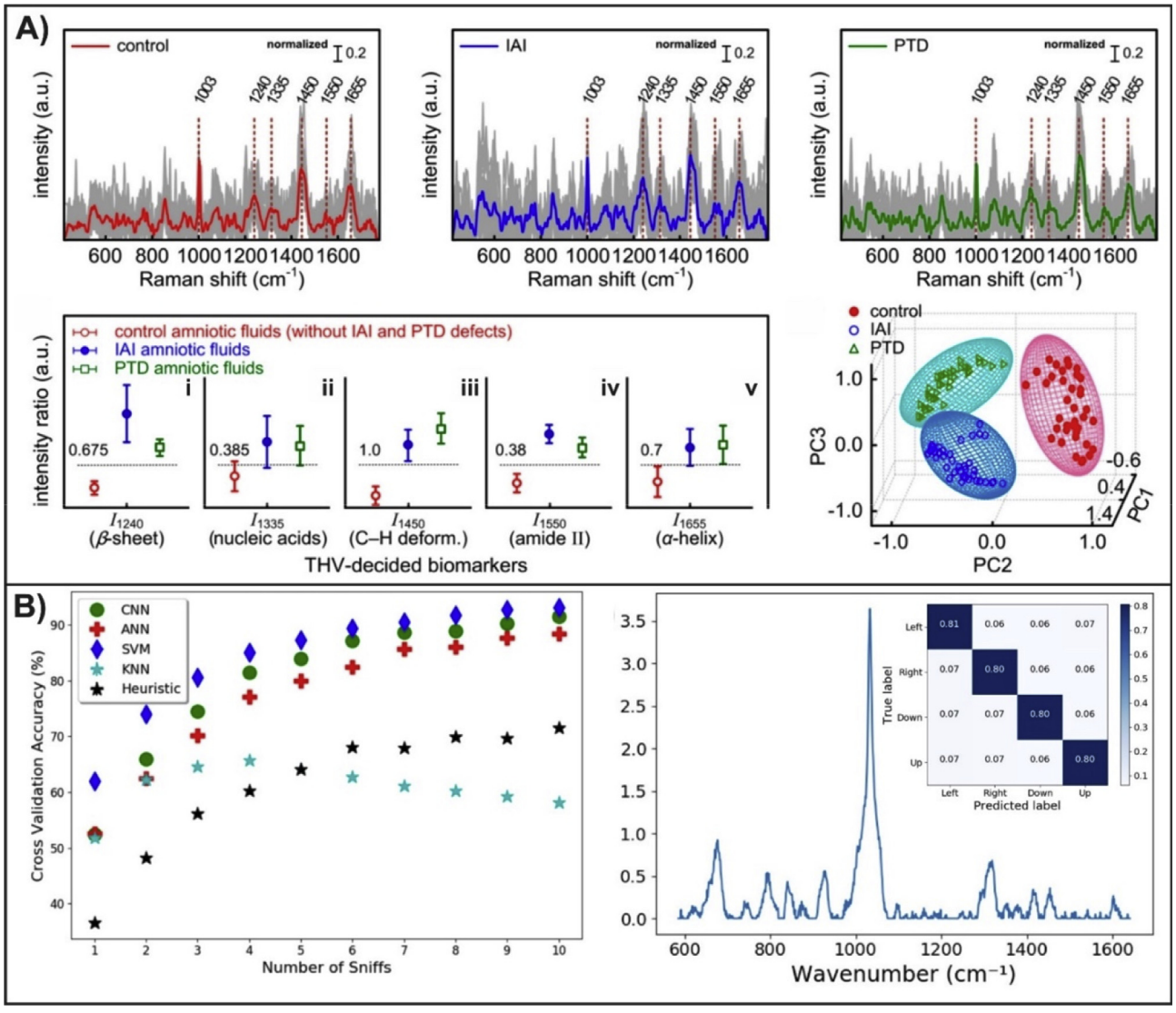

Fig. 6. (A) Representative Raman spectra of the control, intra-amniotic infection (IAI), and preterm delivery (PTD) amniotic fluids on the SERS chip (top) [84]. Five threshold biomarkers with dominant Raman intensities were proposed to discriminate between these fluids (bottom left). The amniotic fluids with IAI and PTD showed an increase in their Raman intensities compared to the control. A three-dimensional PCA score plot of PC1 to PC3 (n = 40) to classify the types of prenatal diseases (bottom right). (B) Cross validation accuracy of the models used (left) [86]. Negative matrix factorization (NMF) component acquired from bacterial VOC training dataset used in bacterial odor source localization. Included as an inset is the normalized confusion matrix produced by 1 sniff SVM model applied to bacterial VOC test dataset (right). Reprinted and adapted with permission of Ref. [84,86]. Copyrights 2018 and 2019 American Chemical Society.

a success rate of 99% for individual spectra. The confusion matrix from this work is shown in Fig. 3B. Viral infections such as influenza can be explored, with recent interest focussing on identifying new emerging influenza viruses [79]. In its simplest form, blood samples acquired from patients that are positive or negative for a bacterial or viral infection can be used to generate predictive models for future patients. Examples include Dengue fever [80,81], Hepatitis C [82], and Tuberculosis [83]. Likewise, intra-amniotic infections caused by different bacteria can also benefit from the combination of SERS and machine learning [84]. After acquiring SERS spectra of three amniotic fluid groups (Fig. 6A), the authors proposed the use of threshold values based on normalized Raman intensities of five vibrational modes. However, it was determined that such an approach would unlikely be useful. Instead, a combination of PCA and SVM was used. The resulting three-dimensional PCA score plot (Fig. 6A) shows that the developed algorithm can identify characteristics of the different amniotic fluids.

Another means of detecting and/or differentiating bacteria is to look for small molecules that are secreted by the bacteria. One means of ascertaining the pathogenicity of toxigenic *Clostridium difficile* in stool samples is to determine if toxin A (TcdA) or toxin B (TcdB) is present [85]. A total of 36 stool samples and 360 Raman spectra were used to prepare the models. It was found that when comparing spiked and unspiked stool samples, different machine learning methods were better at different toxin concentrations. Specifically, stochastic gradient boosting machine worked best at the highest concentration (1000 pg/mL), and either SVM linear kernel or, most predominantly, PCA-LDA worked best for the other concentrations (0.1–100 pg/mL). Using the diffusion gradients of a vapour containing volatile organic molecules, the origin of the molecules, such as the position of bacteria or biofilm, can be determined with an odor compass (3 × 3 array of sensing areas containing gold nanospheres) [86]. A variety of different algorithms were first evaluated (Fig. 6B), and it was found that the SVM model was the best. In the biofilm experiments, a biofilm of *E. coli* was placed adjacent to one side of the SERS sensor. After six hours of exposure, the sensor was characterized *ex situ*. The negative matrix factorization (NMF) component (spectrum in Fig. 6B) was found to match well with that of previous reports of *E. coli*, with an intense band at 1023 cm^{-1} attributed to volatile metabolites. A categorical accuracy of 82.95% was achieved using the SVM model for a test dataset (normalized confusion matrix in Fig. 6B), considerably better than the 25% associated with chance.

4.4. Medical diagnostics

As mentioned in the previous section, proof-of-concept studies have been performed where serum samples from patients of known infections were used to make predictive models for future diagnoses. This same type of approach can be applied to medical disorders. Here, we highlight a few other proof-of-concept studies that have focussed on identifying or classifying aspects of medical disorders. Screening of diabetes mellitus typically requires that the patient provides a blood sample, which is an invasive approach. In this regard, screening using a non-invasive approach that yields rapid results would be ideal. One idea was to use Raman spectra acquired at different sites on the body (earlobe, inner arm, thumbnail, and median cubital vein) instead of blood samples [87]. Using spectra acquired from eleven patients with type 2 diabetes, and eleven patients without diabetes, artificial neural networks were developed. Although the authors report a highest classification accuracy of 96% for spectra acquired at the inner arm, a considerably greater number of patients is vital to determine how effective this approach would be in a clinical setting. In the case of kidney stones, it is not about determining if the patient has them, but instead, to determine what the composition is, and thus, find its origin [88]. 135 kidney stone samples of various chemical compositions: L-cysteine (34), purines (34), phosphates (32), oxalates (35) were obtained from patients. 5-fold cross-validation for kNN and SVM models yielded classification accuracies ranging from 88.1 to 96.3%. SERS has also shown potential as a point-of-care test for eye disorders, specifically cataracts, and oxidative stress-induced age-related macular degeneration and diabetic macular edema [35]. In this proof-of-concept study, aqueous humour samples from 15 patients (5 from each disorder), were collected. With an optimized algorithm incorporating PCA-SVM, the authors reported sensitivities, specificities, and accuracies of 100% for the eye disorders evaluated.

With little doubt, the most commonly studied medical condition by Raman/SERS and machine learning is cancer. In 2014, a review on the use of vibrational spectroscopy and machine learning for the study of various cancer types was published [89]. In the

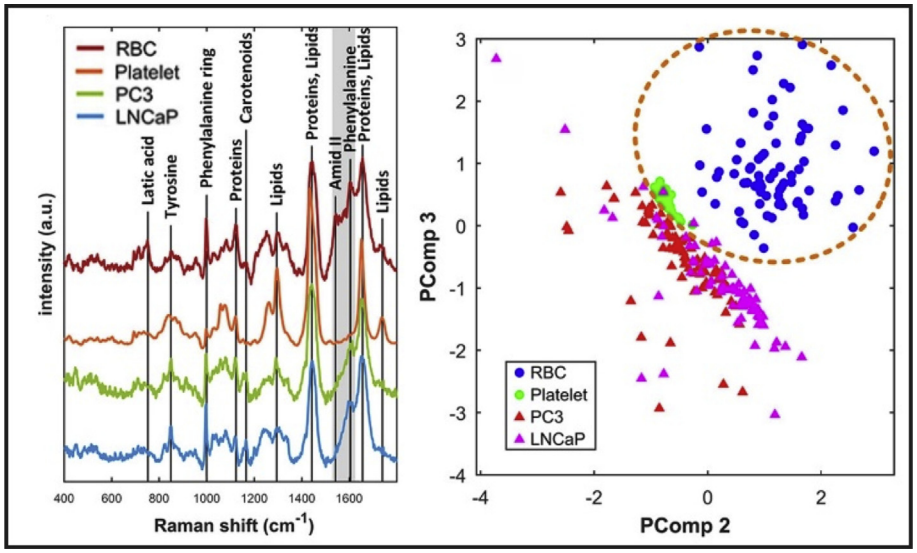

Fig. 7. Raman spectra of four EV subtypes (left) [98]. PCA score plots for the Raman spectra obtained from the subtypes (right). Reprinted and adapted with permission of Ref. [98]. Copyright 2018 American Chemical Society.

intervening years, other cancers not discussed in the earlier review including liver [90], nasopharyngeal [91], osteosarcoma [92], thyroid [93], and tongue squamous cell carcinoma [94] have all been explored by Raman/SERS and multivariate analysis. Although not a new concept, high-resolution Raman imaging has become an important technique to better understand and detect non-cancerous and cancerous cells, especially when coupled with machine learning [92,93].

An important diagnostic tool for determining if cancer is present and its origin is to identify the presence of cancer markers, typically within serum. In the case of small molecules, diagnostic approaches can focus on, but are not limited to; metabolites (*i.e.* tryptophan, phenylalanine, proline, valine, adenine, thymine) in serum [95], modified nucleosides in serum [96], and volatile organic compounds found in breath [97]. The use of proteins as biomarkers has great potential for the diagnosis of cancer. Although some proteins are used as diagnostic tools for specific cancers, it is now recognized that a single protein can also be overexpressed by other types of cancers or the result of another disease. For example, human epididymis protein 4 is an approved marker for ovarian cancer, but it has been shown that patients with prostate cancer, endometrial cancer, and lung cancer can have elevated levels of this protein. Therefore, immunoassays, including SERS-based ones [44], that consider a variety of protein markers is a preferable approach to diagnosing the cause of overexpression. In that study, five protein markers were examined, with sera samples from healthy patients and people suffering from prostate cancer, ovarian cancer, and pancreatitis. The assay used a sandwich approach, with detection based on the presence of a Raman reporter molecule. From there, two algorithms were used (classification tree, kNN) that considered the signal intensity of the Raman reporter molecule for a single peak or the entire spectrum respectively. Even though a sensitivity

Table 1
Chemometric approaches used in research presented in Section 4.

Category	Sample(s) and/or Analyte(s)	Raman and/or SERS	Method(s) Used	Ref.
Food and Beverage	Muscle meat	Raman	PCA, DFA	[49]
	Butter adulteration	Raman	PCA, PCR, PLS, ANNs	[55]
	Caviar quality	Raman	PCA, LDA, ANN, PLSR	[56]
	Animal milk	Raman	PCA, RF	[57]
	Sudan-1	SERS	PCA, PLSR, ANNs, SVR	[58]
	Carmine dye	SERS	PCA	[59]
	Food colourants	Raman and SERS	PCA	[60]
	Pirimiphos-Methyl	SERS	PLSR, SVMR, RF, PCA	[61]
	Fenthion	SERS	RF	[62]
	Thiophanate-methyl and carbendazim	SERS	LS, SVM	[32]
Forensics	Pharmaceutical tablets	Raman	SVM	[63]
	Cocrystal formulations	Raman	PLS, ANNs	[64]
	Dissolution of pharmaceutical tablets	Raman	PLSR, ANN, PCA	[65]
	Illicit drugs	SERS	PCA, SVM	[66]
	Heroin and methamphetamine	SERS	PCA	[67]
	Tetrahydrocannabinol	SERS	PCA, PLSR	[68]
	Fentanyls	SERS	PCA, PLS, DA	[69]
	Human and animal blood	Raman	PLS, DA	[71]
	Human blood	Raman	GA, SVM, DA, ANN, kNN, PCA	[72]
	Human blood	Raman	SVM, DA, PLS	[73]
	Human fingernail clippings	Raman	PCA, SVM	[74]
	Chemical warfare agent	Raman	RF	[75]
	Explosives	Raman	RDF	[76]
Bacteria and Viruses	*Staphylococci* species and growth conditions	Raman	HCA, SVM	[37]
	Urinary tract infections	Raman	SVM	[77]
	Urinary tract infections	Raman	PCA, PLS, DA, SVM	[78]
	Staphylococci strains	Raman	PCA, LDA, 1NN, SVM	[38]
	New influenza viruses	SERS	PCA	[79]
	Dengue fever infections	Raman	PCA, SVM	[80]
	Dengue fever infections	Raman	PCA, RF	[81]
	Hepatitis C infections	Raman	PCA, kNN, RF, LDA, SVM	[82]
	Tuberculosis infections	SERS	PCA, SVM, DTA, RF	[83]
	Prenatal diseases	SERS	PCA, SVM	[84]
	C. difficile toxins	Raman	PCA, LDA, SVM, RF, GBM	[85]
	Odor compass for biofilms	SERS	NMF, MCR, ALS, CNN, ANN, kNN, SVM	[86]
Medical Diagnostics	Diabetes mellitus screening	Raman	ANN, PCA, SVM	[87]
	Kidney stone classification	Raman	PCA, kNN, SVM	[88]
	Eye disorders	SERS	PCA, SVM	[35]
	Liver diseases	SERS	PLS, DA, SVM, ANN	[90]
	Nasopharyngeal and liver cancer	SERS	PLS, PCA, SVM, RBF	[91]
	Osteosarcoma	Raman	KCA, PCA	[92]
	Thyroid cancer	Raman	AHC	[93]
	Tongue squamous cell carcinoma	Raman	CNN, PCA, LDA, SVM	[94]
	Hepatocellular carcinoma metabolic profile	SERS	PLS, DA	[95]
	Modified nucleosides	SERS	PLS, DA	[96]
	Breath analysis of gastric cancer patients	SERS	PCA	[97]
	Various cancer biomarkers	SERS	kNN, RF	[44]
	Prostate cancer extracellular vesicles	Raman	PCA	[98]
	Exosome-like vesicle identification	SERS	PLS, DA, MCR, ALS	[99]
	Breast cancer exosomes	SERS	PCA	[100]
	Exosome differentiation	SERS	PCA	[101]
	Lung cancer exosomes	SERS	PCA	[102]

of 86%, a specificity of 93%, and an accuracy of 91% was achieved for the kNN algorithm, with an extremely limited dataset of 5 samples per classification, more samples are required before the developed algorithms could be used as a diagnostic tool.

Other markers in sera include larger biomaterials, such as exosomes. Exosomes are nanoscale membrane vesicles that are released by a variety of cell and can be found throughout various biofluids. What is important to cancer diagnostics is that the chemical and biochemical components of the exosomes are related to their cell-type of origin. This offers the potential for differentiation not only based on cancerous vs. noncancerous origins, but also on the type of cancer. Raman spectra of cancerous and non-cancerous-derived exosomes are shown in Fig. 7 [98]. The challenges to using Raman spectroscopy alone to determine origin is that the composition of individual exosomes from the same cell line will vary. As well, in SERS experiments, the molecules that are being enhanced will also vary, especially for dried samples where the internal components must now also be considered. Therefore, it is necessary to create a spectral database for exosomes of different origin, prior to using machine learning to determine if differentiation is possible. Examples of this combined approach include differentiating exosomes derived melanoma [99], prostate cancer [98], breast cancer [100], and non-small-cell lung cancer [101,102] from exosomes originating from various non-cancerous cells (red blood, platelets, and alveolar). In these studies, the number of samples studied ranged from 25 [99] to 75 [98]. As shown in Fig. 7, although the PCA score plot shows that 98% of the spectra can be grouped as being healthy or cancerous, the clusters can be rather large, such as that of the red blood cells [98]. As such, we strongly believe that in instances where it is recognized that a variety of spectra are to be expected for the same classification, it is critical that researchers develop a large enough dataset such that there are sufficient spectra both for testing and training purposes.

4.5. Common chemometric methods in the applications

Focussing only on the original research papers presented in this section, we have prepared Table 1 and Fig. 8. In Table 1, the method(s) for each specific reference are provided. For simplicity, we have separated techniques that had been used sequentially, such as PLS-DA, into their separate methods (PLS and DA). One of the things that we noticed was that many of the references explored multiple methods so as to better convey the advantages of their developed algorithm or approach. This is why the frequency of methods shown in Fig. 8, is far greater than the number of references. It is also important to note that for simplicity, in this figure, only those methods that were used by two or more references are included. As expected, we noticed that within the same application, different methods could be used, with some techniques being used throughout, such as PCA with exosome differentiation. Furthermore, none of the commonly used methods were exclusively applied to either Raman or SERS experiments. The combination of Table 1 with Fig. 8 provides greater insight for a beginner to machine learning as combined, they better show not only often how specific methods are used, but also, which methods are often used together or tested against each other. With these considerations in mind, it is important to recognize and remember that although the various machine learning methods have advantages and disadvantages (as discussed in Section 2) and can be utilized across multiple applications, only by evaluating multiple methods is it possible to determine the ideal technique for any given experiment and/or dataset.

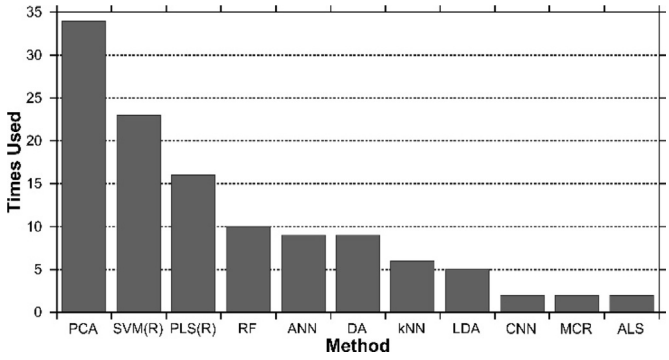

Fig. 8. Most commonly used chemometric methods in the 52 original research papers referenced in Section 4.

5. Outlook and concluding remarks

Deep learning and artificial intelligence methods have had a significant impact on the development of contemporary Raman and SERS sensors, and are poised to have a greater impact in the near future of the field. The development of open source algorithms in addition to commercial software are lowering the entry barrier of these data processing methods in the analysis of Raman and SERS spectra. The improvement of computing power from modern personal computer also facilitates the use of these methods that can be computationally intensive. Hence, wider use of these methods is on the horizon. To provide simpler entry in the field of deep learning and artificial intelligence in Raman or other spectroscopies, this review provides an overview of the most common methods employed in such data processing, advices for selecting the proper tools and validating them, and a series of examples highlighting their use in real-world applications. The next generation of deep learning and artificial intelligence tools will be using more advanced algorithms, or neuronal architecture to further improve the analytical performance of spectra classification. Getting inspiration by the field of computer vision will surely accelerate the development of more robust deep learning model, which could then be applied to detect simple features such as vibrational spectra. For instance, the recent application of CNN based model as in image recognition problems was used to identify pattern in SERS and Raman spectra. CNN models significantly improved selectivity and specificity over conventional ANN for classification of SERS and Raman vibrational spectra, in addition to benefit from advances in Raman spectrophotometers and microscopes to collect better datasets. Lastly, it will facilitate the adoption of methods relying on single molecule Raman scattering, as these datasets are currently challenging to analyze due to the large variability in the spectra of single molecules. As these methods are applicable to other types of spectroscopies, as already demonstrated, the field of deep learning and artificial intelligence data processing in spectroscopy is bound to grow in the near future.

Acknowledgements

This work was funded by the Natural Science and Engineering Research Council (NSERC) of Canada (Grant no. RGPIN/03864-2016) and by the Fonds de Recherche du Québec, Canada (Grant no. 2019-AUDC-262722). The authors would like to dedicate this paper to Prof. Frank Vogt from the University of Tennessee, who suddenly passed away in late 2018, with whom the authors had many discussions on the use of deep learning and artificial intelligence in vibrational spectroscopy. F.L. acknowledges the Alexander von Humboldt Foundation for support.

References

[1] Y. Liu, B.R. Upadhyaya, M. Naghedolfeizi, Chemometric data analysis using artificial neural networks, Appl. Spectrosc. 47 (1993) 12–23.

[2] M. Gniadecka, H.C. Wulf, N.N. Mortensen, O.F. Nielsen, D.H. Christensen, Diagnosis of basal cell carcinoma by Raman spectroscopy, J. Raman Spectrosc. 28 (1997) 125–129.
[3] N.A. Marigheto, E.K. Kemsley, M. Defernez, R.H. Wilson, A comparison of mid-infrared and Raman spectroscopies for the authentication of edible oils, J. Am. Oil Chem. Soc. 75 (1998) 987–992.
[4] R. Goodacre, E.M. Timmins, R. Burton, N. Kaderbhai, A.M. Woodward, D.B. Kell, P.J. Rooney, Rapid identification of urinary tract infection bacteria using hyperspectral whole-organism fingerprinting and artificial neural networks, Microbiology 144 (1998) 1157–1170.
[5] S. Sigurdsson, P.A. Philipsen, L.K. Hansen, J. Larsen, M. Gniadecka, H.C. Wulf, Detection of skin cancer by classification of Raman spectra, IEEE Trans. Biomed. Eng. 51 (2004) 1784–1793.
[6] I.R. Lewis, N.W. Daniel, N.C. Chaffin, P.R. Griffiths, Raman spectrometry and neural networks for the classification of wood types - 1, Spectrochim. Acta A 50 (1994) 1943–1958.
[7] P.L. Stiles, J.A. Dieringer, N.C. Shah, R.P.V. Duyne, Surface-enhanced Raman spectroscopy, Annu. Rev. Anal. Chem. 1 (2008) 601–626.
[8] K. Kneipp, Y. Wang, H. Kneipp, L.T. Perelman, I. Itzkan, R.R. Dasari, M.S. Feld, Single molecule detection using surface-enhanced Raman scattering (SERS), Phys. Rev. Lett. 78 (1997) 1667–1670.
[9] S. Nie, S.R. Emory, Probing single molecules and single nanoparticles by surface-enhanced Raman scattering, Science 275 (1997) 1102–1106.
[10] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press, 2016.
[11] S. Moryosef, A. Samueloff, B. Modan, D. Navot, J.G. Schenker, Ranking the risk-factors for cesarean - logistic-regression analysis of a nationwide study, Obstet. Gynecol. 75 (1990) 944–947.
[12] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, in: Proceedings of the 25th International Conference on Neural Information Processing Systems 1, 2012, pp. 1097–1105.
[13] C. Farabet, C. Couprie, L. Najman, Y. LeCun, Learning hierarchical features for scene labeling, IEEE Trans. Pattern Anal. Mach. Intell. 35 (2013) 1915–1929.
[14] C. Szegedy, L. Wei, J. Yangqing, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, 2015, pp. 1–9.
[15] T.N. Sainath, A. Mohamed, B. Kingsbury, B. Ramabhadran, Deep convolutional neural networks for LVCSR, 2013, pp. 8614–8618.
[16] G. Hinton, L. Deng, D. Yu, G.E. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T.N. Sainath, B. Kingsbury, Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups, IEEE Signal Process. Mag. 29 (2012) 82–97.
[17] T. Ciodaro, D. Deva, J.M. de Seixas, D. Damazio, Online particle detection with neural networks based on topological calorimetry information, J. Phys. Conf. Ser. 368 (2012), 012030.
[18] J. Ma, R.P. Sheridan, A. Liaw, G.E. Dahl, V. Svetnik, Deep neural nets as a method for quantitative structure–activity relationships, J. Chem. Inf. Model. 55 (2015) 263–274.
[19] H.Y. Xiong, B. Alipanahi, L.J. Lee, H. Bretschneider, D. Merico, R.K.C. Yuen, Y. Hua, S. Gueroussov, H.S. Najafabadi, T.R. Hughes, Q. Morris, Y. Barash, A.R. Krainer, N. Jojic, S.W. Scherer, B.J. Blencowe, B.J. Frey, The human splicing code reveals new insights into the genetic determinants of disease, Science 347 (2015) 1254806.
[20] M.K.K. Leung, H.Y. Xiong, L.J. Lee, B.J. Frey, Deep learning of the tissue-regulated splicing code, Bioinformatics 30 (2014) i121–i129.
[21] D.I. Ellis, R. Goodacre, Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy, Analyst 131 (2006) 875–885.
[22] P.S. Gromski, H. Muhamadali, D.I. Ellis, Y. Xu, E. Correa, M.L. Turner, R. Goodacre, A tutorial review: metabolomics and partial least squares-discriminant analysis – a marriage of convenience or a shotgun wedding, Anal. Chim. Acta 879 (2015) 10–23.
[23] N. Kumar, A. Bansal, G.S. Sarma, R.K. Rawal, Chemometrics tools used in analytical chemistry: an overview, Talanta 123 (2014) 186–199.
[24] M.I. Jordan, T.M. Mitchell, Machine learning: trends, perspectives, and prospects, Science 349 (2015) 255–260.
[25] Z. Xiaobo, Z. Jiewen, M.J.W. Povey, M. Holmes, M. Hanpin, Variables selection methods in near-infrared spectroscopy, Anal. Chim. Acta 667 (2010) 14–32.
[26] J. Yang, J. Xu, X. Zhang, C. Wu, T. Lin, Y. Ying, Deep learning for vibrational spectral analysis: recent progress and a practical guide, Anal. Chim. Acta 1081 (2019) 6–17.
[27] X. Fan, W. Ming, H. Zeng, Z. Zhang, H. Lu, Deep learning-based component identification for the Raman spectra of mixtures, Analyst 144 (2019) 1789–1798.
[28] F. Lussier, D. Missirlis, J.P. Spatz, J.-F. Masson, Machine-learning-driven surface-enhanced Raman scattering optophysiology reveals multiplexed metabolite gradients near cells, ACS Nano 13 (2019) 1403–1411.
[29] J. Liu, M. Osadchy, L. Ashton, M. Foster, C.J. Solomon, S.J. Gibson, Deep convolutional neural networks for Raman spectrum recognition: a unified solution, Analyst 142 (2017) 4067–4074.
[30] Y. Xu, S. Zomer, R.G. Brereton, Support vector machines: a recent method for classification in chemometrics, Crit. Rev. Anal. Chem. 36 (2006) 177–188.
[31] C.-C. Chang, C.-J. Lin, LIBSVM: a library for support vector machines, ACM Trans. Intell. Syst. Technol. 2 (2011) 1–27.
[32] J.-L. Li, D.-W. Sun, H. Pu, D.S. Jayas, Determination of trace thiophanate-methyl and its metabolite carbendazim with teratogenic risk in red bell pepper (Capsicum annuum L.) by surface-enhanced Raman imaging technique, Food Chem. 218 (2017) 543–552.
[33] B. Sharma, K. Ma, M.R. Glucksberg, R.P. Van Duyne, Seeing through bone with surface-enhanced spatially offset Raman spectroscopy, J. Am. Chem. Soc. 135 (2013) 17290–17293.
[34] L. Ou, Y. Chen, Y. Su, Y. Huang, R. Chen, J. Lei, Application of silver nanoparticle-based SERS spectroscopy for DNA analysis in radiated nasopharyngeal carcinoma cells, J. Raman Spectrosc. 44 (2013) 680–685.
[35] W. Kim, S.H. Lee, S.H. Kim, J.-C. Lee, S.W. Moon, J.S. Yu, S. Choi, Highly reproducible Au-decorated ZnO nanorod array on a graphite sensor for classification of human aqueous humors, ACS Appl. Mater. Interfaces 9 (2017) 5891–5899.
[36] C.-S. Ho, N. Jean, C.A. Hogan, L. Blackmon, S.S. Jeffrey, M. Holodniy, N. Banaei, A.A.E. Saleh, S. Ermon, J. Dionne, Rapid identification of pathogenic bacteria using Raman spectroscopy and deep learning, Nat. Commun. 10 (2019) 4927.
[37] M. Harz, P. Rösch, K.D. Peschke, O. Ronneberger, H. Burkhardt, J. Popp, Micro-Raman spectroscopic identification of bacterial cells of the genus Staphylococcus and dependence on their cultivation conditions, Analyst 130 (2005) 1543–1550.
[38] K. Rebrošová, M. Šiler, O. Samek, F. Růžička, S. Bernatová, V. Holá, J. Ježek, P. Zemánek, J. Sokolová, P. Petráš, Rapid identification of staphylococci by Raman spectroscopy, Sci. Rep. 7 (2017) 14846.
[39] Q. Li, W. Li, J. Zhang, Z. Xu, An improved k-nearest neighbour method to diagnose breast cancer, Analyst 143 (2018) 2807–2811.
[40] X. Li, T. Yang, S. Li, D. Wang, Y. Song, S. Zhang, Raman spectroscopy combined with principal component analysis and k nearest neighbour analysis for non-invasive detection of colon cancer, Laser Phys. 26 (2016), 035702.
[41] P. Reokrungruang, I. Chatnuntawech, T. Dharakul, S. Bamrungsap, A simple paper-based surface enhanced Raman scattering (SERS) platform and magnetic separation for cancer screening, Sens. Actuators B Chem. 285 (2019) 462–469.
[42] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32.
[43] V. Živanović, S. Seifert, D. Drescher, P. Schrade, S. Werner, P. Guttmann, G.P. Szekeres, S. Bachmann, G. Schneider, C. Arenz, J. Kneipp, Optical nanosensing of lipid accumulation due to enzyme inhibition in live cells, ACS Nano 13 (2019) 9363–9375.
[44] N. Banaei, J. Moshfegh, A. Mohseni-Kabir, J.M. Houghton, Y. Sun, B. Kim, Machine learning algorithms enhance the specificity of cancer biomarker detection using SERS-based immunoassays in microfluidic chips, RSC Adv. 9 (2019) 1859–1868.
[45] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, second ed., Springer, New York, 2009.
[46] G. Zaccone, Getting Started with TensorFlow, Packt Publishing, 2016.
[47] V. Estivill-Castro, Why so many clustering algorithms: a position paper, SIGKDD Explor. Newsl. 4 (2002) 65–75.
[48] D.H. Wolpert, W.G. Macready, No free lunch theorems for optimization, IEEE Trans. Evol. Comput. 1 (1997) 67–82.
[49] D.I. Ellis, D. Broadhurst, S.J. Clarke, R. Goodacre, Rapid identification of closely related muscle foods by vibrational spectroscopy and machine learning, Analyst 130 (2005) 1648–1654.
[50] D. Naumann, Infrared and NIR Raman Spectroscopy in Medical Microbiology, SPIE, 1998.
[51] J. Engel, J. Gerretzen, E. Szymańska, J.J. Jansen, G. Downey, L. Blanchet, L.M.C. Buydens, Breaking with trends in pre-processing? Trends Anal. Chem. 50 (2013) 96–106.
[52] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, A simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (2014) 1929–1958.
[53] V. Ganganwar, An overview of classification algorithms for imbalanced datasets, Int. J. Emerg. Technol. Adv. Eng. 2 (2012) 42–47.
[54] H. Huang, H. Xu, X. Wang, W. Silamu, Maximum F1-score discriminative training criterion for automatic mispronunciation detection, IEEE/ACM Trans. Audio Speech Lang. Process. 23 (2015) 787–797.
[55] R.S. Uysal, I.H. Boyaci, H.E. Genis, U. Tamer, Determination of butter adulteration with margarine using Raman spectroscopy, Food Chem. 141 (2013) 4397–4403.
[56] H. Mohamadi Monavar, N.K. Afseth, J. Lozano, R. Alimardani, M. Omid, J.P. Wold, Determining quality of caviar from Caspian Sea based on Raman spectroscopy and using artificial neural networks, Talanta 111 (2013) 98–104.
[57] A. Amjad, R. Ullah, S. Khan, M. Bilal, A. Khan, Raman spectroscopy based analysis of milk using random forest classification, Vib. Spectrosc. 99 (2018) 124–129.
[58] W. Cheung, I.T. Shadi, Y. Xu, R. Goodacre, Quantitative analysis of the banned food dye Sudan-1 using surface enhanced Raman scattering with multivariate chemometrics, J. Phys. Chem. C 114 (2010) 7285–7290.
[59] Y.-x. Wu, P. Liang, Q.-m. Dong, Y. Bai, Z. Yu, J. Huang, Y. Zhong, Y.-C. Dai, D. Ni, H.-b. Shu, C.U. Pittman, Design of a silver nanoparticle for sensitive surface enhanced Raman spectroscopy detection of carmine dye, Food Chem. 237 (2017) 974–980.
[60] Y.-j. Ai, P. Liang, Y.-x. Wu, Q.-m. Dong, J.-b. Li, Y. Bai, B.-J. Xu, Z. Yu, D. Ni, Rapid qualitative and quantitative determination of food colorants by both Raman spectra and surface-enhanced Raman scattering (SERS), Food Chem. 241 (2018) 427–433.

[61] S. Weng, S. Yu, R. Dong, J. Zhao, D. Liang, Detection of pirimiphos-methyl in wheat using surface-enhanced Raman spectroscopy and chemometric methods, Molecules 24 (2019) 1691.
[62] S. Weng, M. Qiu, R. Dong, F. Wang, L. Huang, D. Zhang, J. Zhao, Fast detection of fenthion on fruit and vegetable peel using dynamic surface-enhanced Raman spectroscopy and random forests with variable selection, Spectrochim. Acta A 200 (2018) 20–25.
[63] Y. Roggo, K. Degardin, P. Margot, Identification of pharmaceutical tablets by Raman spectroscopy and chemometrics, Talanta 81 (2010) 988–995.
[64] P. Barmpalexis, A. Karagianni, I. Nikolakakis, K. Kachrimanis, Artificial neural networks (ANNs) and partial least squares (PLS) regression in the quantitative analysis of cocrystal formulations by Raman and ATR-FTIR spectroscopy, J. Pharm. Biomed. Anal. 158 (2018) 214–224.
[65] B. Nagy, D. Petra, D.L. Galata, B. Démuth, E. Borbás, G. Marosi, Z.K. Nagy, A. Farkas, Application of artificial neural networks for Process Analytical Technology-based dissolution testing, Int. J. Pharm. 567 (2019) 118464.
[66] H. Dies, J. Raveendran, C. Escobedo, A. Docoslis, Rapid identification and quantification of illicit drugs on nanodendritic surface-enhanced Raman scattering substrates, Sens. Actuators B Chem. 257 (2018) 382–388.
[67] R. Salemmilani, B.D. Piorek, R.Y. Mirsafavi, A.W. Fountain, M. Moskovits, C.D. Meinhart, Dielectrophoretic nanoparticle aggregation for on-demand surface enhanced Raman spectroscopy analysis, Anal. Chem. 90 (2018) 7930–7936.
[68] K. Sivashanmugan, K. Squire, A. Tan, Y. Zhao, J.A. Kraai, G.L. Rorrer, A.X. Wang, Trace detection of tetrahydrocannabinol in body fluid via surface-enhanced Raman scattering and principal component analysis, ACS Sens. 4 (2019) 1109–1117.
[69] K. Wang, B. Xu, J. Wu, Y. Zhu, L. Guo, J. Xie, Elucidating fentanyls differentiation from morphines in chemical and biological samples with surface-enhanced Raman spectroscopy, Electrophoresis 40 (2019) 2193–2203.
[70] S.R. Khandasammy, M.A. Fikiet, E. Mistek, Y. Ahmed, L. Halámková, J. Bueno, I.K. Lednev, Bloodstains, paintings, and drugs: Raman spectroscopy applications in forensic science, Forensic Chem. 8 (2018) 111–133.
[71] K.C. Doty, I.K. Lednev, Differentiation of human blood from animal blood using Raman spectroscopy: a survey of forensically relevant species, Forensic Sci. Int. 282 (2018) 204–210.
[72] A. Sikirzhytskaya, V. Sikirzhytski, I.K. Lednev, Determining gender by Raman spectroscopy of a bloodstain, Anal. Chem. 89 (2017) 1486–1492.
[73] K.C. Doty, I.K. Lednev, Differentiating donor age groups based on Raman spectroscopy of bloodstains for forensic purposes, ACS Cent. Sci. 4 (2018) 862–867.
[74] E. Widjaja, G.H. Lim, A. An, A novel method for human gender classification using Raman spectroscopy of fingernail clippings, Analyst 133 (2008) 493–498.
[75] D. Wiktelius, L. Ahlinder, A. Larsson, K. Höjer Holmgren, R. Norlin, P.O. Andersson, On the use of spectra from portable Raman and ATR-IR instruments in synthesis route attribution of a chemical warfare agent by multivariate modeling, Talanta 186 (2018) 622–627.
[76] C. Gasser, M. Göschl, J. Ofner, B. Lendl, Stand-off hyperspectral Raman imaging and random decision forest classification: a potent duo for the fast, remote identification of explosives, Anal. Chem. 91 (2019) 7712–7718.
[77] S. Kloß, B. Kampe, S. Sachse, P. Rösch, E. Straube, W. Pfister, M. Kiehntopf, J. Popp, Culture independent Raman spectroscopic identification of urinary tract infection pathogens: a proof of principle study, Anal. Chem. 85 (2013) 9610–9616.
[78] M. Yogesha, K. Chawla, A. Bankapur, M. Acharya, J.S. D'Souza, S. Chidangil, A micro-Raman and chemometric study of urinary tract infection-causing bacterial pathogens in mixed cultures, Anal. Bioanal. Chem. 411 (2019) 3165–3177.
[79] J.-y. Lim, J.-s. Nam, H. Shin, J. Park, H.-i. Song, M. Kang, K.-i. Lim, Y. Choi, Identification of newly emerging influenza viruses by detecting the virally infected cells based on surface enhanced Raman spectroscopy and principal component analysis, Anal. Chem. 91 (2019) 5677–5684.
[80] S. Khan, R. Ullah, A. Khan, N. Wahab, M. Bilal, M. Ahmed, Analysis of dengue infection based on Raman spectroscopy and support vector machine (SVM), Biomed. Opt. Express 7 (2016) 2249–2256.
[81] S. Khan, R. Ullah, A. Khan, A. Sohail, N. Wahab, M. Bilal, M. Ahmed, Random forest-based evaluation of Raman spectroscopy for dengue fever analysis, Appl. Spectrosc. 71 (2017) 2111–2117.
[82] A. Sohail, S. Khan, R. Ullah, S.A. Qureshi, M. Bilal, A. Khan, Analysis of hepatitis C infection using Raman spectroscopy and proximity based classification in the transformed domain, Biomed. Opt. Express 9 (2018) 2041–2055.
[83] R. Botta, P. Chindaudom, P. Eiamchai, M. Horprathum, S. Limwichean, C. Chananonnawathorn, V. Patthanasettakul, B. Kaewseekhao, K. Faksri, N. Nuntawong, Tuberculosis determination using SERS and chemometric methods, Tuberculosis 108 (2018) 195–200.
[84] W. Kim, S.H. Lee, J.H. Kim, Y.J. Ahn, Y.-H. Kim, J.S. Yu, S. Choi, Paper-based surface-enhanced Raman spectroscopy for diagnosing prenatal diseases in women, ACS Nano 12 (2018) 7100–7108.
[85] S.K. Koya, S. Yurgelevic, M. Brusatori, C. Huang, L.N. Diebel, G.W. Auner, Rapid detection of clostridium difficile toxins in stool by Raman spectroscopy, J. Surg. Res. 244 (2019) 111–116.
[86] W.J. Thrift, A. Cabuslay, A.B. Laird, S. Ranjbar, A.I. Hochbaum, R. Ragan, Surface-enhanced Raman scattering-based odor compass: locating multiple chemical sources and pathogens, ACS Sens. 4 (2019) 2311–2319.
[87] E. Guevara, J.C. Torres-Galván, M.G. Ramírez-Elías, C. Luevano-Contreras, F.J. González, Use of Raman spectroscopy to screen diabetes mellitus with machine learning tools, Biomed. Opt. Express 9 (2018) 4998–5010.
[88] X. Cui, Z. Zhao, G. Zhang, S. Chen, Y. Zhao, J. Lu, Analysis and classification of kidney stones based on Raman spectroscopy, Biomed. Opt. Express 9 (2018) 4175–4183.
[89] M. Sattlecker, N. Stone, C. Bessant, Current trends in machine-learning methods applied to spectroscopic cancer diagnosis, Trends Anal. Chem. 59 (2014) 17–25.
[90] X. Li, T. Yang, S. Li, L. Jin, D. Wang, D. Guan, J. Ding, Noninvasive liver diseases detection based on serum surface enhanced Raman spectroscopy and statistical analysis, Opt. Express 23 (2015) 18361–18372.
[91] Y. Yu, Y. Lin, C. Xu, K. Lin, Q. Ye, X. Wang, S. Xie, R. Chen, J. Lin, Label-free detection of nasopharyngeal and liver cancer using surface-enhanced Raman spectroscopy and partial lease squares combined with support vector machine, Biomed. Opt. Express 9 (2018) 6053–6066.
[92] J. Li, J. Qin, X. Zhang, R. Wang, Z. Liang, Q. He, Z. Wang, K. Wang, S. Wang, Label-free Raman imaging of live osteosarcoma cells with multivariate analysis, Appl. Microbiol. Biotechnol. 103 (2019) 6759–6769.
[93] J.N. Taylor, K. Mochizuki, K. Hashimoto, Y. Kumamoto, Y. Harada, K. Fujita, T. Komatsuzaki, High-resolution Raman microscopic detection of follicular thyroid cancer cells with unsupervised machine learning, J. Phys. Chem. B 123 (2019) 4358–4372.
[94] H. Yan, M. Yu, J. Xia, L. Zhu, T. Zhang, Z. Zhu, Tongue squamous cell carcinoma discrimination with Raman spectroscopy and convolutional neural networks, Vib. Spectrosc. 103 (2019) 102938.
[95] R. Xiao, X. Zhang, Z. Rong, B. Xiu, X. Yang, C. Wang, W. Hao, Q. Zhang, Z. Liu, C. Duan, K. Zhao, X. Guo, Y. Fan, Y. Zhao, H. Johnson, Y. Huang, X. Feng, X. Xu, H. Zhang, S. Wang, Non-invasive detection of hepatocellular carcinoma serum metabolic profile through surface-enhanced Raman spectroscopy, Nanomedicine 12 (2016) 2475–2484.
[96] S. Feng, Z. Zheng, Y. Xu, J. Lin, G. Chen, C. Weng, D. Lin, S. Qiu, M. Cheng, Z. Huang, L. Wang, R. Chen, S. Xie, H. Zeng, A noninvasive cancer detection strategy based on gold nanoparticle surface-enhanced Raman spectroscopy of urinary modified nucleosides isolated by affinity chromatography, Biosens. Bioelectron. 91 (2017) 616–622.
[97] Y. Chen, Y. Zhang, F. Pan, J. Liu, K. Wang, C. Zhang, S. Cheng, L. Lu, W. Zhang, Z. Zhang, X. Zhi, Q. Zhang, G. Alfranca, J.M. de la Fuente, D. Chen, D. Cui, Breath analysis based on surface-enhanced Raman scattering sensors distinguishes early and advanced gastric cancer patients from healthy persons, ACS Nano 10 (2016) 8169–8179.
[98] W. Lee, A. Nanou, L. Rikkert, F.A.W. Coumans, C. Otto, L.W.M.M. Terstappen, H.L. Offerhaus, Label-free prostate cancer detection by characterization of extracellular vesicles using Raman spectroscopy, Anal. Chem. 90 (2018) 11290–11296.
[99] S. Stremersch, M. Marro, B.-E. Pinchasik, P. Baatsen, A. Hendrix, S.C. De Smedt, P. Loza-Alvarez, A.G. Skirtach, K. Raemdonck, K. Braeckmans, Identification of individual exosome-like vesicles by surface enhanced Raman spectroscopy, Small 12 (2016) 3292–3301.
[100] N. Ferreira, A. Marques, H. Águas, H. Bandarenka, R. Martins, C. Bodo, B. Costa-Silva, E. Fortunato, Label-free nanosensing platform for breast cancer exosome profiling, ACS Sens. 4 (2019) 2073–2083.
[101] Z. Yan, S. Dutta, Z. Liu, X. Yu, N. Mesgarzadeh, F. Ji, G. Bitan, Y.-H. Xie, A label-free platform for identification of exosomes from different sources, ACS Sens. 4 (2019) 488–497.
[102] J. Park, M. Hwang, B. Choi, H. Jeong, J.-h. Jung, H.K. Kim, S. Hong, J.-h. Park, Y. Choi, Exosome classification by pattern analysis of surface-enhanced Raman spectroscopy data for lung cancer diagnosis, Anal. Chem. 89 (2017) 6695–6701.