



Spectral classification by generative adversarial linear discriminant analysis

Ziyi Cao^a, Shijie Zhang^b, Youlin Liu^a, Casey J. Smith^a, Alex M. Sherman^a, Yechan Hwang^a, Garth J. Simpson^{a,*}

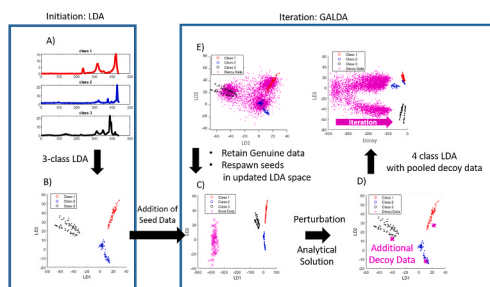
^a Department of Chemistry, Purdue University, 560 Oval Drive, West Lafayette, IN, 47907, USA

^b Takeda Pharmaceuticals International Co, 35 Landsdowne Street, Cambridge, MA, 02139, USA

HIGHLIGHTS

- Generative adversarial linear discriminant analysis (GALDA) was developed for spectral classification.
- A theoretical foundation for implementing GALDA for spectral dimension reduction and classification was derived.
- Simulations with known ground truth spectral classes supported the assessment of GALDA.
- Application of GALDA improved classification accuracy in polymorph discrimination by conventional Raman spectroscopy.
- Pixel-wise application of GALDA produced composition maps in good agreement with manual assignments for model system.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:

Linear discriminant analysis (LDA)
Generative adversarial networks (GAN)
Raman spectroscopy
THz Raman spectroscopy

ABSTRACT

Generative adversarial linear discriminant analysis (GALDA) is formulated as a broadly applicable tool for increasing classification accuracy and reducing overfitting in spectrochemical analysis. Although inspired by the successes of generative adversarial neural networks (GANs) for minimizing overfitting artifacts in artificial neural networks, GALDA was built around an independent linear algebra framework distinct from those in GANs. In contrast to feature extraction and data reduction approaches for minimizing overfitting, GALDA performs data augmentation by identifying and adversarially excluding the regions in spectral space in which genuine data do not reside. Relative to non-adversarial analogs, loading plots for dimension reduction showed significant smoothing and more prominent features aligned with spectral peaks following generative adversarial optimization. Classification accuracy was evaluated for GALDA together with other commonly available supervised and unsupervised methods for dimension reduction in simulated spectra generated using an open-source Raman database (Romanian Database of Raman Spectroscopy, RDRS). Spectral analysis was then performed for microscopy measurements of microspheroids of the blood thinner clopidogrel bisulfate and in THz Raman imaging of common constituents in aspirin tablets. From these collective results, the potential scope of use for GALDA is critically evaluated relative to alternative established spectral dimension reduction and classification methods.

* Corresponding author.

E-mail address: gsimpson@purdue.edu (G.J. Simpson).

<https://doi.org/10.1016/j.aca.2023.341129>

Received 8 November 2022; Received in revised form 20 February 2023; Accepted 21 March 2023

Available online 28 March 2023

0003-2670/© 2023 Published by Elsevier B.V.

1. Introduction

The increasing volume and complexity of information from modern chemical instrumentation places growing importance on methods to aid in data-intensive visualization and analysis [1,2]. Dimension reduction is a key arrow in the analysis quiver, in which raw measurements in a high dimensional space (e.g., spectral space) are reduced to a handful of manageable feature-dimensions [3,4]. Transformation to feature-space provides several key advantages: i) ease of visualization of inherent clustering veiled at high dimension, ii) signal to noise enhancement through suppression of directions in measurement space that contribute predominantly to noise, and iii) improved simplicity for statistical hypothesis testing at low dimension. Dimension reduction can be performed as a stand-alone operation for data visualization or in conjunction with classification, in which the output is a class assignment (e.g., through identification of decision boundaries within a reduced-dimensional space). The number of methods available for dimension reduction is many, ranging from classic linear methods such as principal component analysis (PCA) and linear discriminant analysis (LDA) through complex neural network architectures for navigating highly nonlinear interrelationships in the data [5,6]. However, even in these nonlinear cases, complementary analysis using linear methods can provide useful benchmarking and enable estimates for quantitative analysis using well-developed statistical methods for hypothesis testing [5].

Classification of supervised high dimensional data is intimately connected with dimension reduction operations. As in dimension reduction, a diverse suite of algorithms exists for classification of high-dimensional data. LDA [7] and soft independent modeling by class analogy (SIMCA) [8,9] are two of the earliest classification approaches finding widespread adoption. When used for classification, LDA assigns classes to data based on the Mahalanobis distance to the class mean [9]. For any pair of classes with identical intraclass covariance matrices, the dividing surface of equal probability for class assignment forms a linear decision boundary within the reduced dimensional space. Extension to quadratic discriminant analysis (QDA) relaxes the assumption of identical intraclass covariance matrices, producing decision boundaries with quadratic curvature [7]. In SIMCA, classification is evaluated by first performing unsupervised dimension reduction (e.g., via PCA) within each class independently, followed by assessment of distance (Euclidean or Mahalanobis) from the resulting PCA-based hypersurfaces [8,9]. More recently, artificial neural networks (ANNs) have found widespread use for classification of high-dimensional data [10], notably exemplified in image [10–12] and speech analysis [12–14]. ANNs are particularly advantageous in navigating nonlinear relationships between high-dimensional data (e.g., translation, rotation, rescaling) that notoriously frustrate linear methods. The number of methods available for classification has ballooned rapidly in recent years with the expanded use of ANNs. While laudable, these collective successes in classification do not always cleanly translate to analogous successes in dimension reduction for data visualization. For example, no universally excepted approaches are available for using QDA [15,16] or SIMCA for dimension reduction. Similarly, self-organizing maps (SOMs) produced by ANNs can be interpreted as feature extraction representations [17]. However, the nonlinear relationships inherent in ANNs can complicate intuitively relating interconnectivity mapping in SOMs back to the initial input data.

The two simplest and most universally adopted linear methods for dimension reduction are arguably PCA and LDA, representing archetypical unsupervised and supervised dimension reduction approaches, respectively [18,20]. In PCA, the data are pooled, and dimensions are selected that maximize the variance within the projected data [19,21]. As an unsupervised learning approach, PCA is often used to identify intrinsic patterns and groupings within high-dimensional data sets. When class information is available, LDA identifies the dimensions that maximize the resolution between the projected data through

maximization of the Fisher linear discriminant function J (defined explicitly in the Supporting Information). Based on these metrics, one might initially expect LDA to be always preferred over PCA when classification information is available (i.e., with supervised data), as PCA neglects the additional classification information in the analysis. However, in practice, LDA is often ill-posed and prone to computational instabilities [20], most notably when the number of training spectra is small relative to the dimensionality of the measurements (e.g., spectra) [21,22].

The reason for the failure of LDA in these limits is tied to a matrix inversion step in its evaluation [23]. The Fisher linear discriminant function J maximized in LDA is given by the ratio of the between-class variance divided by the within-class variance, evaluated following projection of the input spectra onto a given test vector $\mathbf{w}$ in spectral-space or matrix of test vectors $\mathbf{W}$ [24]. The linear algebra manipulations to determine the optimal projections $\mathbf{w}^*$ to maximize J involve identification of the eigenvectors of the product of two matrices: one given by the inverse of a matrix describing the within-class variance $\mathbf{S}_W$ and the other describing the between-class variance $\mathbf{S}_b$. Because of the matrix inversion operation, the eigenvectors and eigenvalues of $\mathbf{S}_W^{-1} \mathbf{S}_b$ may not be uniquely solvable or may exhibit large uncertainties. Such numerical instabilities are most likely to arise when the dimensionality of the measurements p greatly exceeds the number of replicates n in each class (i.e., $p > n$) [25], which is often the case in both spectral and image analyses. One common consequence of this instability is high overfitting in poorly-posed LDA evaluations, corresponding to a gap in resolution between training and testing data sets.

A number of regularization strategies have been proposed for addressing computational instabilities associated with the matrix inversion step with $p > n$, and the corresponding propensities for overfitting that arise as a consequence. Arguably, the simplest regularization approach is the “shrunk centroids” method proposed by Guo, Hastie, and Tibshirani [26], in which scaled addition of an identity matrix to $\mathbf{S}_W$ results in a mathematically stable matrix inversion operation [27]. Alternatively, Mehay, Cai, and Harrington demonstrated the use of singular value decomposition (SVD) to estimate the pseudo-inverse (e.g., the Moore-Penrose pseudo-inverse) for regularizing the evaluation of $\mathbf{S}_W^{-1} \mathbf{S}_b$ [28]. Friedman also proposed a regularized discriminant analysis integrating linear and quadratic discriminant analysis, combined with a weighted identity matrix sharing similarities with the shrunk centroids method [27]. While generally very successful in enabling full-dimension LDA of spectral data, regularization can be quite sensitive to both the method and the degree of regularization [29].

Beyond regularization, a host of alternative approaches are available for feature extraction to reduce the spectral dimensionality prior to performing LDA as a means of overcoming the limitations of direct LDA at full dimension. Arguably, the simplest method is to first perform PCA as an initial dimension reduction step, then perform LDA within the lower-dimensional PCA-space. Partial least-squared discriminant analysis (PLS-DA) is also widely for spectral classification [30], as well as dimension reduction [31]. Fourier transformation and discrete wavelet transforms have also been implemented for reducing dimensionality in spectral space prior to LDA [32]. While generally quite successful, the information lost by dimension reduction prior to LDA cannot be recovered by subsequent operations, such that the method adopted for feature extraction can significantly impact the quality of the final outcomes.

Complementing feature extraction approaches, a suite of methods has also been developed for variable selection, in which only the most informative dimensions within the raw input data (e.g., spectral peaks) are retained for initial dimension reduction. Variable selection can be as simple as manually windowing within the high dimensional inputs to reject featureless sections in subsequent analyses. Improved generalizability and reproducibility can be achieved by formalizing criteria for retention or rejection during variable selection [33–37]. In order to select the optimal subset of variables, forward selection initially adds significant variables (e.g., based on F-test assessments relative to

variance within the total set of variables) [37,38]. Backward elimination subsequently refines by removing variables with least contributions (e.g., based on F-test assessments from within the forward-selected subset) [38]. Among forward, backward, and stepwise (forward and backward) variable selection, stepwise selection arguably yields the most stable result [38]. As one might expect, variable selection generally performs well when signal information is localized to small numbers of discrete spectral domains (e.g., sharp spectral peaks), and suffers when information is broadly distributed across spectral space. In such cases, coordinate transformation (e.g., via PCA, Fourier transformation, or wavelet transformation) can substantially improve the benefits of variable selection. In this limit, variable selection and feature extraction become effectively synonymous.

Transfer learning strategies, broadly defined, have also been implemented with considerable success for dimension reduction and spectral analysis in the $p > n$ regime. Calibration transfer methods have a long and successful history of leveraging large volumes of well-characterized measurements to inform on spectral analyses when few case-specific spectra are available [39–41]. In work using artificial neural networks for spectral analysis, Li et al. [42] developed a deep transfer learning based near infrared spectroscopy multi-manufacturer drug identification method. This method achieved higher classification accuracy and scalability in multi-variety and multi-manufacturer NIR compared with current popular methods, such as support vector machines (SVM), back propagation (BP), the use of an autoencoder (AE) and extreme learning machines (ELM). Zhang et al. [43] proposed an approach using a transfer-learning model pretrained on a standard Raman spectral database for the identification of Raman spectra of organic compounds that were not included in the database and with limited data. Zhu et al. [44] first demonstrated usefulness and effectiveness of GANs for classification of hyperspectral images (HSIs), using training samples to fine-tune a discriminative CNN for image classification. However, the application of the paper was limited to remote sensing with a focus on classification rather than dimension reduction, with no chemical spectral analysis. Later, Yu et al. [45] demonstrated classification of pathogens by Raman spectroscopy combined with generative adversarial networks to analyze the most salient identification regions in the real spectrum. Much of this collective body of work centers on classification, with fewer options for dimension reduction to aid in visualization of high dimensional spectral data. Furthermore, the inherent stochastic nature of neural network training and decision-making complicates reliance on artificial neural network architectures for decision-making requiring compliance with federal guidance, such as in the pharmaceutical industry.

The numerical instability in full-dimension LDA with small training size has some qualitative similarities to over-fitting effects arising in machine learning tools, such as artificial neural networks [46]. Neural networks also generally possess a substantially greater number of adjustable parameters relative to the number of inputs used in their training [47,48]. As one example, the convolutional neural network ImageNet contained 6×10^7 network parameters for classification of image data with only $\sim 10^3$ images per class [48]. In the limit of a small number of training spectra, repeated optimization of an ANN during training can result in increasing reliance on noise in driving classification/regression in addition to the signal. As a consequence, over-fitting results in an increasing disparity between the accuracy of ANNs when evaluated with training versus testing datasets [7]. In the case of ANNs, several strategies to address the consequences of overfitting in data-limited settings have found widespread adoption. The most common are arguably transfer learning (TL) [49] and the use of generative adversarial networks (GANs) [50]. By analogy with calibration transfer in chemometrics introduced in the preceding paragraph, TL in neural network applications leverages a pre-trained network to serve as a foundation for the extension to new systems [51]. In applications involving neural networks, TL requires access to a pre-trained network, which may not be available in many instances. Physics-informed simulations for data augmentation approaches introduced in the discussion

of spectral data are also widely used in neural networks [52], in which neural networks are trained taking into consideration partial differential equations that represent physics-informed constraints. Another common approach is the use of generative adversarial networks (GANs), in which a second competing neural network or algorithm is introduced to improve the statistical reliability of an ANN. In brief, the generator is most often designed to convert a random initial seed to an input registered as being genuine and of a particular target class. A re-optimized ANN is produced to reject the generated “decoy” data as false, which is then targeted again by an updated GAN. Iteration between attack and defense improves the broader utility of the ANN in data-limited applications [53]. Between TL networks and GANs, GANs have the broader utility of the two approaches, as they can operate wholly independently for a given new problem without prior knowledge of related systems or pre-training. Consistent with this expectation, Pavlou et al. [62] recently demonstrated the use of a GAN for spectral augmentation to minimize overfitting in PCA + LDA spectral analysis of bone tissue samples by Raman spectroscopy. In that work, a GAN was used to generate synthetic spectra that were statistically indistinguishable from genuine data within the reduced-dimensional space. Compiled analysis of the augmented data set resulted in improved accuracy in subsequent PCA + LDA spectral analyses.

In light of the successes of generative adversarial approaches to minimize over-fitting artifacts in ANNs [54,55] and for data augmentation, we hypothesize that analogous benefits may be realized to address overfitting in LDA by integrating adversarial updates directly into the LDA operation itself. We developed an analog of the nonlinear processes intrinsic in GANs but built around linear transformations inherent in LDA. Specifically, we developed a linear mathematical framework for optimally perturbing a random input seed to generate decoy spectra for data augmentation and used LDA under computationally stable conditions of $p < n$ to optimally separate genuine and generated training data, and iteratively optimized the processes to compete against each other by analogy with established GAN architectures [53]. In contrast to calibration transfer and other transfer learning methods for “nudging” models by rebuilding, the generative adversarial approach leaves both the model and the training data fixed, and instead adversarially augments real data with additional generated spectra, kept separate from the real data in the analysis but constrained to co-locate with real data in the reduced-dimensional space. In this sense, GALDA is a data augmentation approach for dimension reduction by LDA, in which overfitting is suppressed by locating and rejecting the regions in the transformed space in which the real data are *not* located, in contrast with most common chemometric approaches designed broadly around data reduction prior to LDA. Because the raw data are transformed by simple matrix multiplication in GALDA rather than through nonlinear transformations inherent in artificial neural networks, the outcomes are intuitively interpretable and amenable to standard approaches for uncertainty propagation.

In this work, a mathematical framework for GALDA is proposed. Implementation of GALDA is compared with other common linear methods for dimension reduction using simulated spectra generated from an archived spectral database. Following these proof-of-concept studies using simulated data sets with known ground-truth outcomes, the utility of GALDA is assessed for two classes of Raman imaging data: i) dimension reduction of a set of spectra from Raman microscopy of clopidogrel bisulfate microspheroids, and ii) dimension reduction and classification of THz Raman measurements of common constituents of aspirin tablets. The strengths and limitations of GALDA are then critically evaluated for dimension reduction and classification of spectrochemical data.

2. Materials and methods

2.1. Overall workflow

Fig. 1 provides an overview of the workflow for GALDA, in which LDA is used for dimension reduction. Classified initial input data in Fig. 1A are projected into a reduced dimensional space defined by LDA, indicated in Fig. 1B. Next, a generative adversarial iteration (Fig. 1C through 1E) is performed. The panel in Fig. 1C shows the projection of randomly generated inputs serving as initial “decoy” data projected into LDA-space. In Fig. 1D, the decoy data are then modified by the linear addition of a perturbation in spectral space to induce classification as one of the input classes within LDA-space. In Fig. 1E, LDA is performed again, but now with an additional class for the decoy data. The generation of decoy data, perturbations, and LDA is performed iteratively until convergence is achieved (Fig. 1C through 1E). Convergence is realized when additional adversarial iterations provide negligible change in testing and training resolution (illustrated by plateaus in Supporting Information Figures SI. 2 and SI. 6). Details on implementation of LDA are summarized in the Supporting Information.

2.2. Launching an attack to generate decoy spectra

By analogy with GANs, initial random inputs were used to seed the generation of decoy data. For the simulated data, the random seeds were produced from a uniform distribution with an upper bound of 6×10^{-3} arbitrary amplitude units (a.u.) and a lower bound of 0 a.u. These bounds were chosen to reflect the upper bound and lower bound of the RDRS data following normalization. Random seeds were then projected onto the reduced dimensional space using the eigenvectors from LDA, producing broad distributions of projections.

An adversarial attack was performed by identifying the perturbation

δ in the spectral domain that optimally altered the classification of an initial spectrum x_i (e.g., random seed) to a predetermined target class in LDA-space, subject to constraints imposed by a cost function. A perturbed spectrum, x' , was given by $x' = x_i + \delta$. The general strategy in the optimization of the attack perturbation δ is illustrated Fig. 2, which is intended to serve as a graphical depiction of an attack shown in the reduced-dimensional space (e.g., as determined by LDA). Each wavelength channel in the original spectral space results in a “nudge” to collectively contribute to the position of the spectrum in the reduced dimensional space. While the primary spectral features (indicated by the thin black arrows) combine to dictate the general position within the reduced dimensional space, randomness within the noise (indicated by the short red arrows in Fig. 2). Fig. 2A and B produces a spread about that mean position.

Additional “nudges” by perturbation to each wavelength channel of the original spectrum can relocate the position of the initial spectrum in the reduced dimensional space to one significantly closer to the target, as illustrated in Fig. 2C. The vector of deviations d from the initial sample spectrum, x_s , to the “target”, x_t , in the reduced dimensional space is given by the following expression.

$$d = W \cdot [(x_s + \delta) - x_t] \quad \text{Eq. 1}$$

The matrix W is comprised of the set of eigenvectors that project the high-dimensional data to a lower-dimensional space (such as PCA or LDA). In the absence of other considerations, the optimal perturbation, $\hat{\delta}$, will be one that maximizes the probability that the perturbed spectrum will be classified as the target, subject to the cost function. In the absence of a cost function, the reduction in dimension associated with W is an underdetermined problem; an infinite number of selections for δ will generally produce comparable values for d .

The selection of one among the innumerable possible perturbations was performed by addition of the cost function in Eq. (2) to minimize the

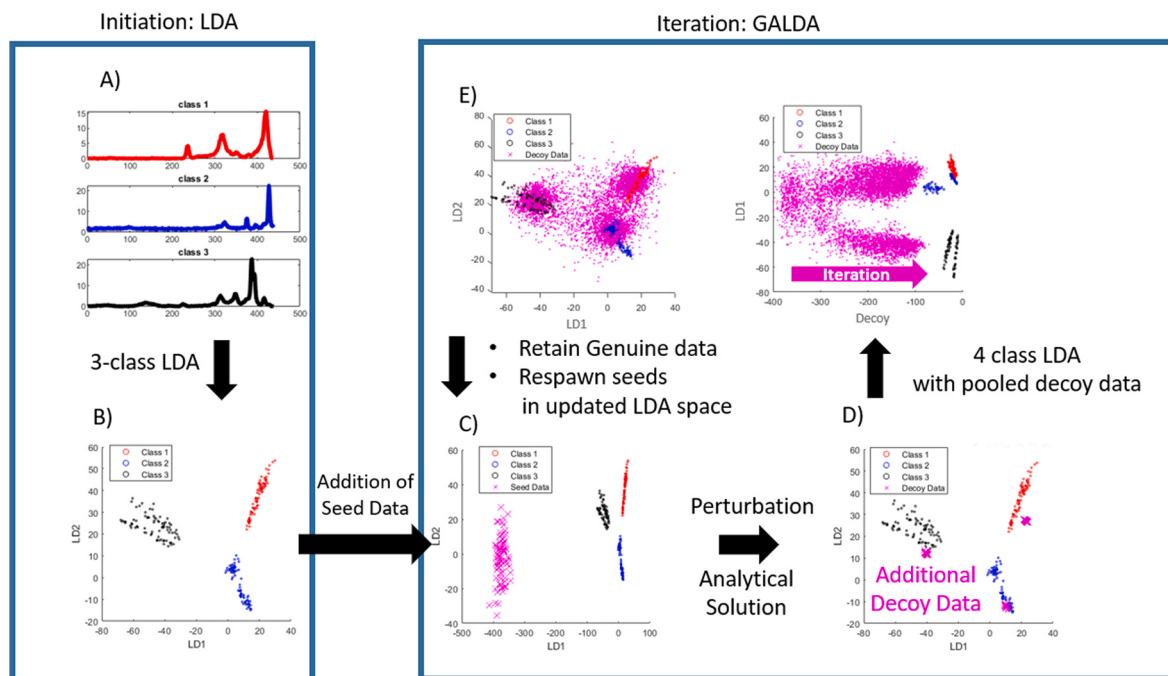


Fig. 1. Workflow of the generative adversarial linear analysis algorithm. (A, B) Initial data that fall into distinct classes are projected onto a reduced dimensional space. (C) Random inputs serve as initial decoy data projected into this LDA-space, depicted in magenta here in 2D. (D) The decoy data classes (red, blue, black) are then “attacked” to induce classification as one of the original input classes in LDA-space while minimizing the magnitude of the perturbation depicted in magenta here in spectral space. (E) LDA is then performed again, but now with an additional class and corresponding dimension for isolating the decoy data. Early generations of decoy data are well separated from the genuine data with separation narrowing in the subsequent interactions. The processes C-E are iterated until convergence is achieved.

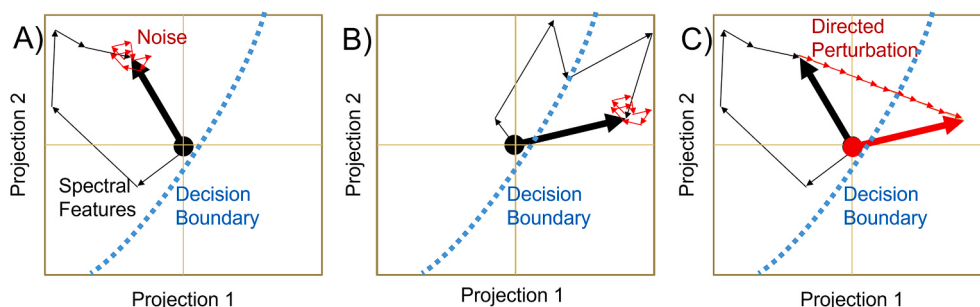


Fig. 2. Conceptual illustration of a spectroscopic adversarial attack. Major spectral peaks drive the position of spectra in lower-dimensional projections, demonstrated by the set of thin, long black arrows for Classes A and B, respectively. In C) the addition of patterned perturbations in the vector δ , illustrated by the set of directional red arrows, optimally relocates the position from the source Class A to Class B in this reduced dimensional space.

squared magnitude of the perturbation. The total cost function for the reduced dimensional analysis was given by the sum of the two terms, which collectively minimized the sum of squared deviations to the target in the reduced dimensional space while simultaneously minimizing the overall squared magnitude of the perturbation, δ , in spectral space, both evaluated as the squared L_2 norms.

$$\hat{\delta} = \underset{\delta}{\operatorname{argmin}} [\|W \cdot (x_s + \delta - x_t)\|_2^2 + \beta \|\delta\|_2^2] \quad \text{Eq. 2}$$

The scalar parameter β in Eq. (2) allows for empirical adjustment of the relative weight given to proximity to the target relative to perturbation of the major spectral features. The first term in the cost-function is designed to “fool the classifier” by minimizing the distance to the target in the reduced dimensional space, while the second term is targeted to “fool the human” by minimizing the perturbation magnitude in spectral space. Unless explicitly stated otherwise, a value of $\beta = 1$ was used throughout for simplicity.

The optimal perturbation was determined analytically by rewriting the cost function F in terms of the perturbation.

$$F(\delta) = \|W \cdot (x_s + \delta - x_t)\|_2^2 + \beta \|\delta\|_2^2 \quad \text{Eq. 3}$$

To solve for the optimal perturbation, the minimum of $F(\delta)$ was found by setting the gradient to zero ($\nabla F = 0$). Before performing the gradient, it is convenient to explicitly evaluate the squared magnitudes analytically, which can be simplified by defining $C = W^T W$ and $v = x_s - x_t$.

$$F(\delta) = \|W \cdot (v + \delta)\|_2^2 + \beta \|\delta\|_2^2 = v^T C v + 2v^T C \delta + \delta^T (C + \beta I) \delta \quad \text{Eq. 4}$$

The following vector gradient operations [57] for a general function $f(x)$ can be used to simplify Eq. (4).

$$f(x) = A x \Rightarrow \nabla f(x) = A \quad \text{Eq. 5}$$

$$f(x) = x^T A x \Rightarrow \nabla f(x) = (A^T + A)x \quad \text{Eq. 6}$$

Using the relations in Eq. (5) and Eq. (6) noting that $v^T C = C^T v$ and $C^T = C$, ∇F can be written in the following form.

$$\nabla F(\delta) = 2(C + \beta I)\delta + 2Cv \quad \text{Eq. 7}$$

Setting $\nabla F = 0$ and solving for δ yields the following expression for the optimized perturbation.

$$\hat{\delta} = -(C + \beta I)^{-1} C(x_s - x_t) \quad \text{Eq. 8}$$

In the limit of $\beta = 0$ and with no rescaling, the perturbation δ modifies the source spectrum (e.g., random seed) to match exactly the target in spectral space.

The script to perform gradient calculations is available at <http://github.itap.purdue.edu/Simpson-Laboratory-for-Nonlinear-Optics/GALDA-public> along with the other scripts for performing

adversarial and GALDA iterations.

2.3. Raman simulations

Simulations were performed to assess the performance of the GALDA algorithm for a dataset with known ground-truth results. Six Raman spectra selected from the Romanian Database of Raman Spectroscopy [56] (magnetite, galena, molybdenite, goethite, stibnite, pyrolusite) recorded with 1024 wavelength channels served as ground truth source spectra. Prior to addition of noise, noise-free source spectra were normalized to produce squared magnitudes of one. Following normalization, 100 simulated spectra for each of the six were generated by addition of noise with a mean of 0 for different assumed distributions. To generate bimodally distributed classes, the spectra were grouped pairwise into three classes, with data distributed about two separate modes in each class. This pairwise grouping (analogous to a single beach with two kinds of pebbles) was designed as a simple model for classes with non-normal probability distributions. Particularly in biological assemblies, spectral classification can routinely and reliably be performed even when each class is comprised of a complex multi-component mixture. A bimodal distribution represents a simple, minimal model for accounting for diversity in composition within a given spectral class. Unless stated otherwise, five-fold cross-validation was performed on all reported simulated and experimental data, resulting 80%–20% split for training and testing.

2.4. Raman measurements

Clopidogrel bisulfate Form I and Form II were produced in-house at Dr. Reddy's Laboratories and were used as received. Both the Form I and Form II particles were spherical with similar particle size distributions (diameter: $\sim 25 \mu\text{m}$). A set of 252 Raman spectra was collected from the clopidogrel samples and separated into three classes (84 spectra per class) – Form I, Form II, and background (glass slide). The ground truth identity of these samples was known by visual inspection. Raman spectra were acquired using a custom Raman microscope, built in-house, and described in detail previously [57]. In brief, a continuous wave diode laser (Toptica, RPS785/24) was collimated by a fused silica lens and directed through an X–Y scan head composed of two galvanometer scanning mirrors. Two additional fused silica lenses formed a 4f configuration to deliver a collimated beam to the back of a 10x objective (Nikon). The Raman signal from the sample was collected through the same objective and descanned back through the same beam path into the Raman probe. A notch filter was built in the Raman probe to reject the laser signal. Raman spectra were acquired using an Acton SP-300i spectrometer with a 100×1340 CCD array and controlled by a computer running WinSpec32 software. The laser power measured at the sample was ~ 30 mW. The exposure time was 0.5 s per spectrum. To

achieve a higher signal to noise ratio for high-quality training data for classification, 30 consecutive frames were averaged for each spectrum acquired over a spot size of $\sim 2\text{--}3\ \mu\text{m}$ diameter within the field of view. After acquiring the Raman images, 84 spectra per class were randomly sampled from the images from multiple particles locating at different positions within the field of view to ensure a representative distribution within the samples selected. A Savitzky-Golay filter [58] was applied to smooth the spectra, and a rolling ball filter was used to remove the fluorescence background. Finally, the spectra were normalized to their integrated intensities, i.e., the areas under the curves. The optimal perturbation $\hat{\delta}$ evaluated using Eq. (6) was added to each random seed to produce generative adversarial decoy spectra.

2.5. THz Raman measurements

Acetylsalicylic acid, sucrose, and magnesium stearate were purchased from Sigma, VWR Chemicals, and Alfa Aesar, respectively, and mixed with mortar and pestle. THz Raman spectra were collected using a DXR2xi Raman imaging microscope (Thermo Scientific) with a custom modified incident beam path for compatibility with a ONDAX THz-Raman System and a ONDAX Cleanline Laser System (Coherent, 785 nm CW laser). The laser output was directed through a 10x objective and a 50 μm slit, and then focused onto a motorized sample stage (Prior). A custom grating was used for THz Raman spectroscopy for a Raman shift range of $-500\ \text{cm}^{-1}$ to $2750\ \text{cm}^{-1}$ with a spectral resolution of $2\ \text{cm}^{-1}$. Exposure time was 0.1 s, and signal averaging of 5 scans was performed to generate each THz Raman spectrum. THz Raman spectroscopic imaging was achieved by point measurements and sample scanning. Image pixel resolution was 3 μm .

2.6. Methods for comparison

Implementations of GALDA were benchmarked against multiple established methods for dimension reduction. Benchmarking algorithms included PCA, PCA-LDA (i.e., two-stage dimension reduction consisting first of PCA followed by LDA in PCA-space), PLS-DA (partial least squares discriminant analysis), and RLDA (regularized LDA, using “shrunk centroids” regularization [26,59]). These methods were selected based on their simplicity and ubiquity and are not meant to represent an exhaustive set of comparators for dimension reduction. PCA-LDA and RLDA were evaluated both using default parameters (naïve implementations) and following hyperparameter optimization (i.e., the number of PCA-dimensions in PCA-LDA and the regularization parameter in RLDA). The naïve implementations of PCA-LDA are similar to results obtained using a default metric (e.g., 70%) for the retained variance [60]. For the optimized PCA-LDA analyses of the simulated data, the exact probability density function (PDF) describing the data distribution was known *a priori*. As such the theoretical maximum resolution and minimum in overfitting could be evaluated exactly as well, corresponding to implementation of PCA with six retained dimensions followed by LDA in the six-dimensional PCA-space. Optimization of PCA-LDA for the measured spectral datasets was performed using a “scree” test⁶³.

All results are reported for projection into spaces defined by the intrinsic $c-1$ rank of LDA (e.g., 2-dimensional representations for three-class datasets). Results from regularization of LDA using the Moore-Penrose pseudo-inverse (PLDA) as described by Harrington and co-workers [28] to evaluate the matrix inversion operation were also performed using the built-in “pinv()” function in MATLAB using the default tolerance for singular value decomposition (SVD). With the exception of PCA as a stand-alone method, all other comparators were supervised methods that retain the label information when performing dimension reduction, analogous to GALDA. Additional details on the hyperparameter optimizations are summarized in the Supplemental Information.

GALDA was performed with two different implementations. In the first, generative adversarial attacks were performed in spectral-space, followed by dimension reduction by PCA-LDA (labeled as “GA-PCALDA”). GA-PCALDA implementations were initially computationally well-posed, with $p \leq n$. In addition, GALDA was implemented for the initially ill-posed problem of direct LDA analysis in spectral-space, for which $p > n$ prior to generative attack for the spectral data sets used. Values for the testing resolution for GALDA in this initially ill-posed case are only reported in the regime $p < n_{\text{tot}}$, once the combined number of initial and generated spectra (n_{tot}) exceeded the number of spectral wavelength channels.

3. Results and discussion

3.1. Spectral adversarial attack

Consistent with Eq. (2), an adversarial attack is defined herein as an additive perturbation to an initial seed or spectrum in spectral-space with an optimally small overall magnitude intentionally designed to induce misclassification in the reduced dimensional space (e.g., LDA-space). An example of such can be found in Fig. 3A, in which an initial source spectrum from Class 3 (an equal mixture of molybdenite and goethite) of a simulated dataset derived from a mineral database is perturbed to register as a spectrum from Class 1 (an equal mixture of magnetite and galena). Fig. 3B contains the applied perturbed spectrum, which is significantly smaller in amplitude comparing to the initial source spectrum Fig. 3C shows the classification in the reduced-dimensional LDA-space, in which the initial seed spectrum falls into the cluster of data points belonging to Class 3, consistent with the genuine class for the spectrum. Using Eq. (8), the optimal perturbation direction in spectral space was calculated, with the green trace in with the corresponding indicating the trajectory in LDA-space produced upon the addition of the perturbation to the initial seed. No constraints other than proximity to target in LDA-space and minimization of magnitude in spectral-space were imposed on the cost function given in Eq. (3).

Interestingly, the perturbations shown in Fig. 3B that optimally produced changes in classification did not affect the major spectral features in the source spectrum, as it is visually difficult to detect the perturbation in spectral space that squarely relocated the spectrum to the target classes in the reduced dimensional space. Rather, the optimal perturbations to induce misclassification as seen in Fig. 3B were characterized by high-frequency content likely dominated by noise. This result is initially somewhat counter-intuitive, as one might expect perturbations to recover the prominent low-frequency spectral features clearly apparent in the mean target spectrum. In the absence of noise, precisely such an outcome would be expected. However, the presence of significant noise provides a route for misclassification through many small perturbations distributed throughout the entire spectral range, illustrated graphically in Fig. 2. These results are analogous to adversarial examples observed in the image recognition field, in which “inputs formed by applying small but intentionally worst-case perturbations to examples from the dataset, such that the perturbed inputs result in the model outputting an incorrect answer with high confidence.” [61].

3.2. Generative adversarial linear discriminant analysis

Fig. 1 provides an overview of the GALDA iterative “attack and defend” process for the Raman spectroscopy of simulated data in the reduced-dimensional space. Initially, a 3-class LDA of the training data (100 spectra from each class), resulted in a clear separation between the three bimodal classes, indicated by the red, blue, and black projected data points, respectively. Next, uniformly distributed random seed data were generated and projected onto this initializing 2D LDA-space, resulting in a broad distribution indicated by the magenta data points in Fig. 1C. Perturbation through adversarial attack in spectral space transformed the seed data to decoy data, which subsequently projected

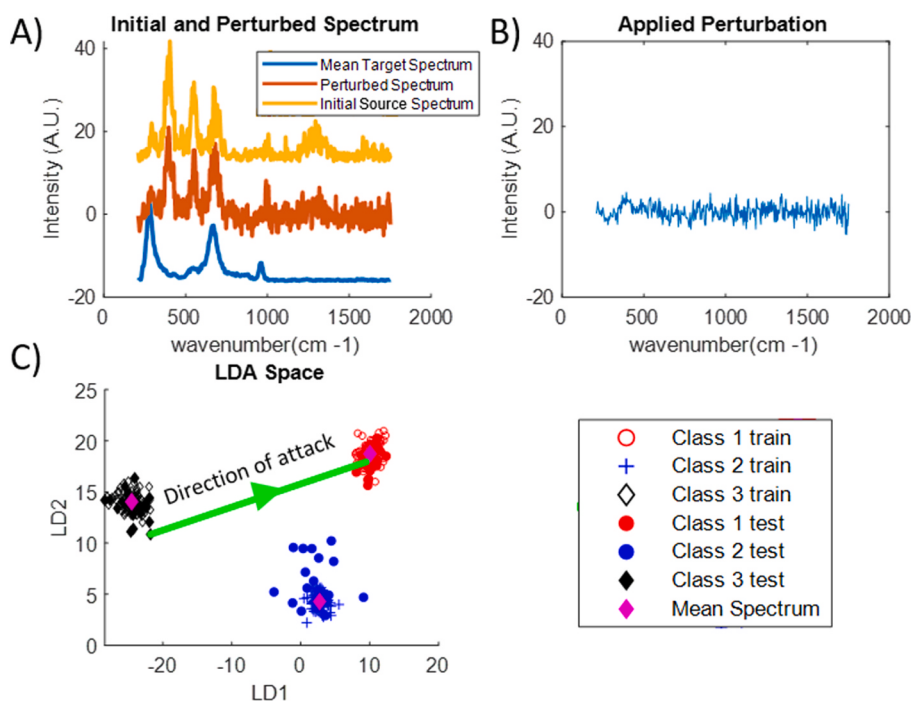


Fig. 3. Example of a spectroscopic adversarial attack, in which a Class 3 spectrum is perturbed such that it is misclassified as Class 1, performed with bimodal simulated data. A) Includes the initial source spectrum, the Class 1 mean target spectrum, and the optimally perturbed post-“attack” spectrum. Spectra are vertically offset to aid visualization. B) Indicates the optimal perturbation δ . In C), the direction of the perturbation and the final position are shown in the reduced dimensional LDA-space.

close to targets randomly assigned from the training set, as demonstrated in Fig. 1D. Note that “target” is defined here as the centroid of a given target class selected randomly for each attack irrespective of proximity to the initial random seed. Subsequently performing a 4-class LDA produced separation between the generated and genuine training data by assigning the generated “decoy” data as a fourth class. Upon projection in 3D space, the LDA coordinate with the greatest eigenvalue (LD1) distinctly separated the training and decoy data. The remaining two coordinates (LD2 and LD3) served as the updated reduced dimensional LDA-space for visualization of the training and testing data.

3.3. Classification accuracy

The classification accuracy of generative adversarial linear analyses was evaluated using simulated datasets with known ground truths for different distributions of testing and training spectra. The optimal reduced-dimensional projections for both PCA and LDA are implicitly formulated on the assumption of normally distributed noise about each class mean. In practice, distributions in testing data can deviate substantially from this assumption through non-normally distributed noise and through offsets in the mean by calibration shifts or addition of unaccounted impurities. In order to approach the inherent deviations away from normal PDFs expected in realistic datasets, while simultaneously supporting facile replication of results by others, several non-standard PDFs were assumed to evaluate the flexibility of PCA, LDA, and GALDA to different PDFs describing the data distribution. The results summarized in Fig. 4 were performed to evaluate the resilience of classification within reduced dimensional spaces to such deviations from ideal behavior. Spectral features from an additional reference spectrum (goethite) were added to the testing set but not included in the training set to simulate an unidentified impurity. In addition, classification accuracy for addition of Poisson and uniformly distributed noise was considered. Decision boundaries within the reduced dimensional space were determined based on shortest Euclidean distance to the class mode.

Relative to classical linear methods, generative adversarial linear analyses exhibited improvements in classification accuracy upon addition of impurity spectral contributions. Naïve PCA-LDA exhibited substantial reductions in classification accuracy upon introduction of

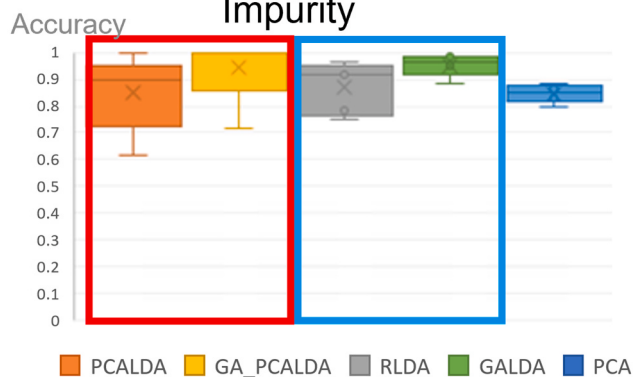
impurity contributions (85%), while the corresponding GA-PCALDA method retained 94% classification accuracy for the simulated Raman spectra. Similar advantages were observed when comparing RLDA and GALDA (both of which are evaluated directly by LDA without invoking PCA for initial dimension reduction). Coordinates for dimension reduction by RLDA produced substantial misclassification rates (87% mean accuracy) when spectral features not included in the training were added to the testing data to simulate the presence of an unknown impurity. In contrast, loading plots generated by GALDA retained 95% classification accuracy for the same testing and training spectra. The improvement of the classification accuracy is tentatively attributed to further suppression of noise in the loading plots due to adversarial data augmentation.

Analysis of the bimodally distributed data suggest similar advantages of generative adversarial algorithms relative to their common analogous linear methods in the presence of non-normally distributed noise, the results of which are also summarized in Fig. 4. Bimodal distributions within each class were adopted to simulate a nontrivial intraclass distribution. For simulated classes with substantial non-normally distributed noise, the classification accuracy was notably low for RLDA (49% for uniformly distributed noise and 44% for Poisson-distributed noise). However, these accuracies improved substantially upon application of GALDA (to 84% and 91%, respectively). Analogous improvements were generally observed for PCA-LDA. Following addition of uniform and Poisson noise to the testing and training data, initial accuracies for naïve PCA-LDA of 92% and 91% recovered to accuracies of 98% and 99%, respectively, by GA-PCALDA. The improvements in classification accuracy for GALDA and GA-PCALDA in the presence of non-normally distributed data is tentatively attributed to the absence of obvious implicit assumptions regarding the functional form of the data distribution within the adversarial stage of the generative adversarial algorithms.

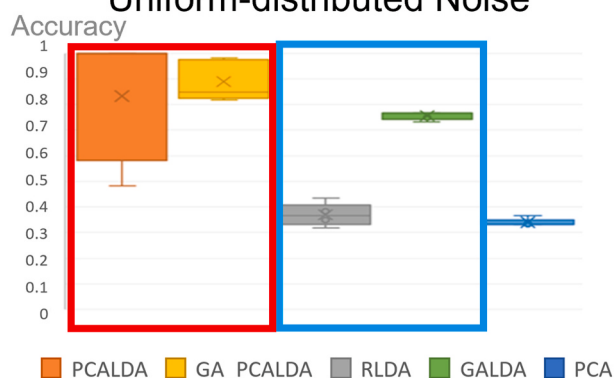
3.4. Analysis of clopidogrel bisulfate

Following the preceding proof of concept assessment based on simulations with known distributions, GALDA was used to analyze Raman microspectroscopy measurements of clopidogrel bisulfate microspheroidal formulations. Microparticles of clopidogrel bisulfate in two

A) Simulated Monomodal Data Impurity



B) Simulated Bimodal Data Uniform-distributed Noise



C) Simulated Bimodal Data Poisson Noise

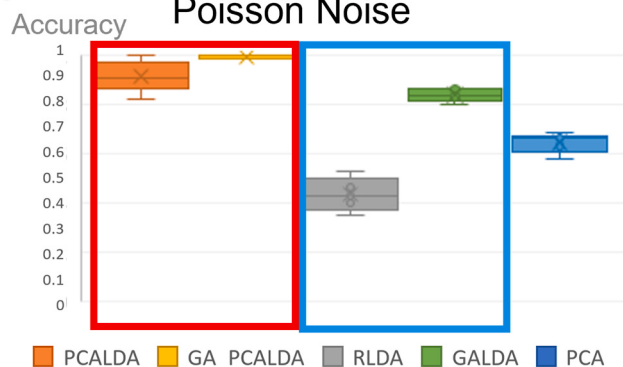


Fig. 4. Classification result of simulated monomodal Raman spectra from the mineral database following five-fold cross-validation. The red rectangle groups PCALDA with GA-PCALDA and the blue rectangle groups LDA with GALDA. In A), an impurity of geothite spectrum of was added to noise free data scaling between 0% and 50% of the total intensity. In B) and C), classification accuracy is compared for simulations of bimodal Raman spectra produced with addition of uniformly distributed B) and Poisson distributed C) noise. Poisson noise was generated by proportionally setting the rescaled ground truth source spectrum equal to the mean of a Poisson distribution at each wavelength. Uniform noise was added with a variance of the input data variance. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

different crystal forms (Form I and Form II) were physically mixed, dispersed on a microscope slide, and analyzed by Raman microscopy to determine the dominant crystal form within each individual microparticle. Spectra were assigned to three classes (Form I, Form II, and

background) based on manual visual inspection of the spectra, as detailed previously[57]. Forms I and II spectra contained several distinctive spectral features in the fingerprint region from 800 cm^{-1} – 1200 cm^{-1} , while normalized background spectra were characterized by a rolling feature centered at $\sim 1375\text{ cm}^{-1}$ tentatively assigned to a spurious fluorescence interference.

Mirroring trends observed in the simulated data sets, application of GA linear analyses improved the classification accuracy of clopidogrel bisulphate spectral assignments. A summary of class accuracy, together with loading plots, confusion matrices, and representative projections of testing and training data in the corresponding reduced dimensional spaces are shown in Fig. 5. Relative to PCALDA for dimension reduction followed by a maximum likelihood estimate (MLE) distance linear classifier, the use of GA-PCALDA improved the accuracy of class assignment from 79% to 97%. A similar comparison between RLDA and GALDA realized improvements from 80% to 94%, respectively.

This improvement in class accuracy was reflected by a corresponding reduction in noise within the loading plots and a reduction in overfitting when comparing the testing and training resolution (summarized in the Supporting Information). Comparisons of the loading plots produced by GALDA and RLDA in Fig. 5B and C, respectively, yields smoother curves with more clearly discernible spectral features upon iterative optimization by GALDA. For spectra with peaks spanning multiple wavelength channels, content at much higher frequencies is generally attributable to noise, consistent with observed improvements in the accuracy by GALDA. As with the simulated datasets, the improvement arising upon application of GA methods is attributed to the rejection of high dimensional contributions from locations where the genuine data do *not* reside through adversarial data augmentation.

3.5. Analysis of THz Raman data

Low frequency THz Raman spectra were acquired using a Raman imaging microscope to map powdered blends of constituents typically found in commercially available aspirin tablets, including aspirin, sucrose, and magnesium stearate. In contrast to conventional Raman spectroscopy of molecular vibrations, THz Raman spectroscopy is typically dominated by weaker intermolecular interactions and librational motions. As such, THz Raman is often highly sensitive to differences in solid-state form, producing distinctive spectral features from molecular crystals.

Prior to Raman imaging measurements of heterogeneous samples, training and cross-validation was performed using spectra obtained from thin films of pure components. Measurements of accuracy following dimension reduction with MLE distance decision boundaries are shown in Fig. 6. As in the preceding simulated and clopidogrel bisulfate experimental datasets, the GA versions of common linear analysis algorithms resulted in notable improvements in classification accuracy under otherwise similar conditions. Implementation of GA-PCALDA produced similar mean accuracy of discrimination between aspirin, sucrose, and magnesium stearate of 99% by naïve PCALDA and GA-PCALDA. However, the mean accuracy of RLDA improved from 86% to 97% upon implementation of GALDA. Interesting, the optimal distribution of spectra projected into the reduced-dimensional space for GA-PCALDA was notably T-shaped compared to projections by PCALDA alone, with the magnesium stearate spectra distributing broadly across the first linear discriminant axis.

Following training and cross-validation, classifiers designed by GA-PCALDA were used to perform chemical speciation in THz Raman microscopy measurements of an aspirin, sucrose, and magnesium-stearate powdered blend, the results of which are shown in Fig. 7. Magnesium stearate particles were characteristically only a few μm in diameter, tentatively assigned as dark particulates dispersed throughout the bright field images. From bright-field images, initial boundaries were identified for different crystal domains corresponding to sucrose and aspirin. Tentative borders were identified for the larger crystalline particulates

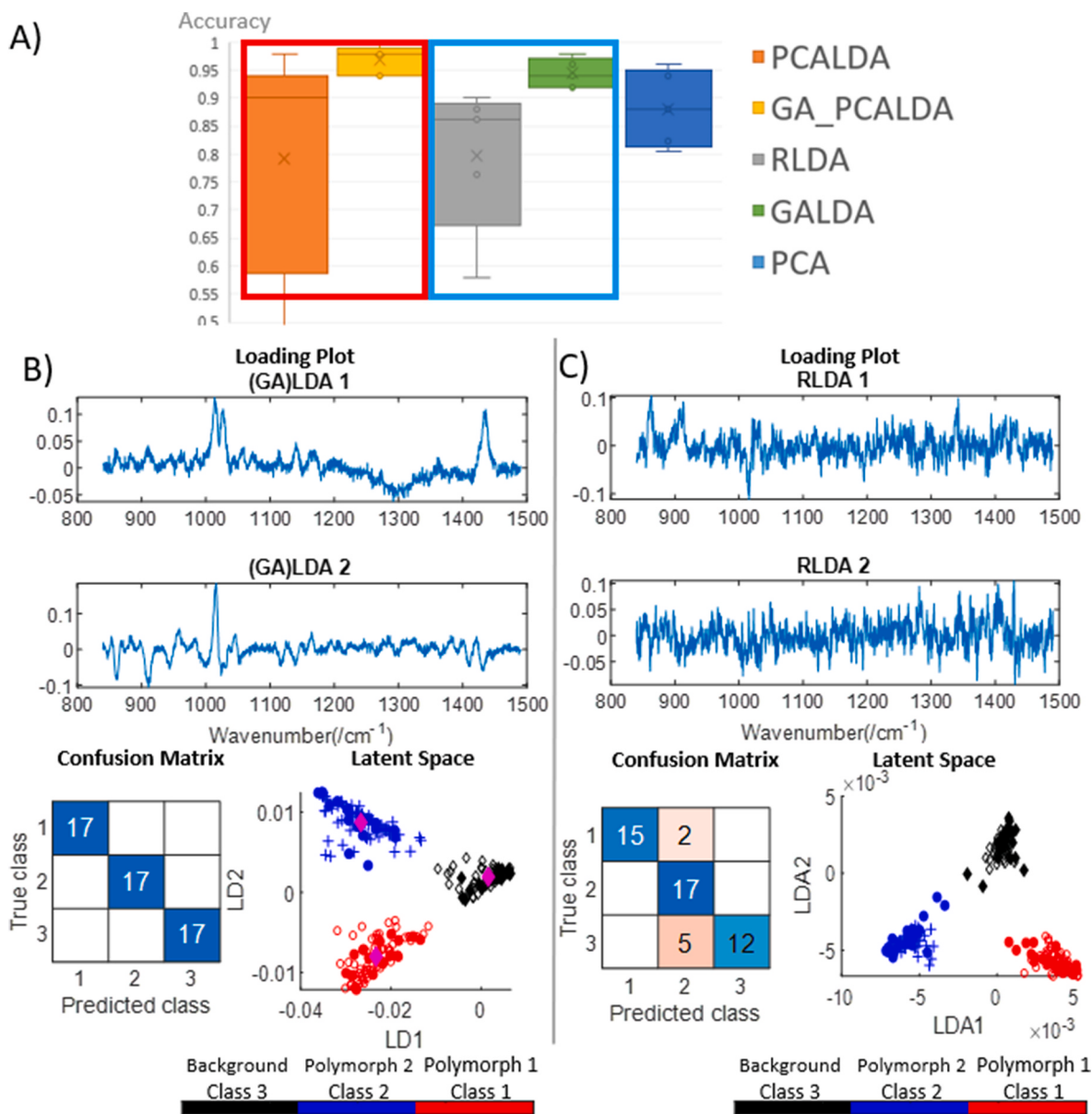


Fig. 5. Analysis of Raman spectra acquired for clopidogrel bisulfate polymorphs, comparing generative adversarial methods with common well-established linear dimension reduction/classification approaches. A) Cross-validated classification accuracy of GA methods with their analogous non-GA counterparts. The red rectangle groups PCALDA with GA-PCALDA and the blue rectangle groups LDA with GALDA. B) Spectral loading plots, projections in latent space, and confusion matrices for GALDA. C) Spectral loading plots, projections in latent space, and confusion matrices for RLDA. In latent spaces, hollowed symbols and crosses correspond to training and solid symbols to testing sets. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

(indicated in the figures) and tentative initial assignments of composition were made based on differences in crystal habits and size distributions, with red border indicating sucrose and blue aspirin. THz Raman imaging was performed in the smaller region contained within the black border in the bright field image. Each pixel in the Raman image was assigned to aspirin, sucrose, or magnesium stearate using the reduced dimensional projections and decision boundaries for each method described in Fig. 6.

Inspection of the resulting composition maps for RLDA, GALDA, PCALDA and GA-PCALDA in Fig. 7 yields clear differences in per-pixel classification assignments. Most notably, both GA methods classify the region in the upper left as predominantly magnesium stearate, while RLDA assigns the same region to sucrose and PCALDA is ambiguous containing all three class assignments. The relatively large domain indicated by the red outline in Fig. 7A is classified as sucrose with well-defined borders in both GA methods. In contrast, RLDA provides no

significant border between the upper left and central bottom regions and PCALDA yields assignments of both sucrose and aspirin for those pixels, complicating definitive assignment. GA-PCALDA includes both assignments of aspirin and magnesium stearate for the blue particles in the top right corner, which may potentially arise from the apparent ubiquity of small magnesium stearate microparticles adhered to the surfaces of the larger crystalline particles. Magnesium stearate is often added to powder blends to reduce friction and aid in particle flow. While no independent ground truth results are available for classification of these individual pixels, the assignments by GALDA and GA-PCALDA are arguably in best agreement with the preliminary assessments of composition based on crystal habit and size.

4. Conclusion

We describe an iterative generative adversarial approach to

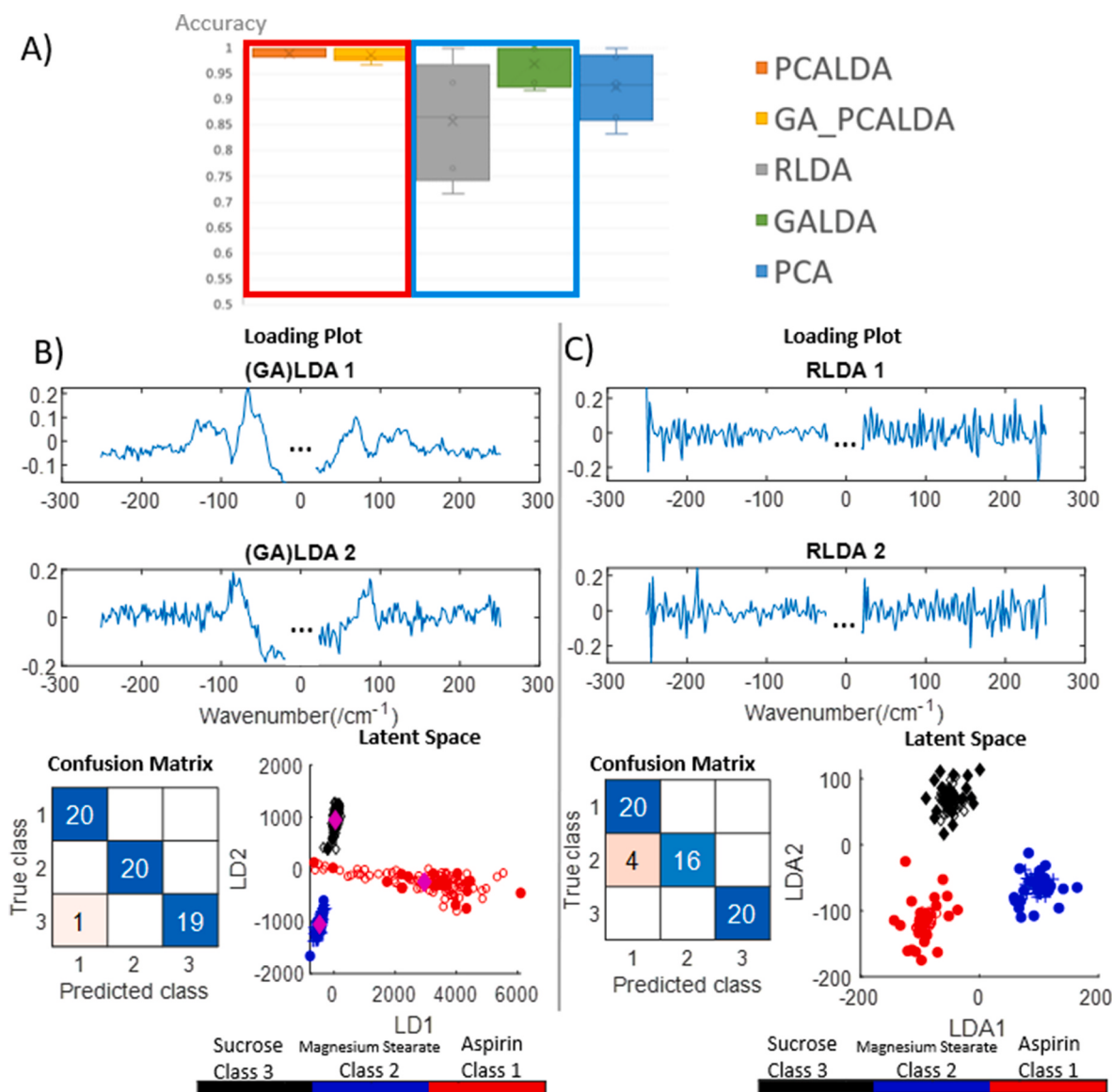


Fig. 6. Analysis of THz Raman spectra acquired for aspirin, sucrose, and magnesium stearate particulate samples, comparing generative adversarial methods with common well-established linear dimension reduction/classification approaches. A) Cross-validated classification accuracy of GA methods with their analogous non-GA counterparts. The red rectangle groups PCALDA with GA-PCALDA and the blue rectangle groups LDA with GALDA. B) Spectral loading plots, projections in latent space, and confusion matrices for GALDA. C) Spectral loading plots, projections in latent space, and confusion matrices for RLDA. (In the loading plots, wavenumber close to zero are not included in the analysis, because of Rayleigh scattering. In latent spaces, hollowed symbols and crosses correspond to training and solid symbols to testing sets. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

minimize overfitting propensities inherent in many dimension reduction/feature extraction algorithms for spectrochemical analysis and improve classification accuracy. GA methods were implemented for PCA-LDA, RLDA, and for LDA performed at full spectral rank, then compared with PCA, PCA-LDA, and RLDA as stand-alone methods using both simulated and experimentally measured Raman spectral data sets. Studies of simulated data with a known ground-truth optimum demonstrate results from GA methods achieved higher classification accuracy comparing to non-GA methods. Consistent with applications using generative adversarial networks in artificial intelligence engines, GA methods are anticipated to be compatible with complex and structured probability density functions underpinning data distributions. The insensitivity to initial conditions also suggests GALDA may find use as general-purpose tools for refinement of initial outcomes produced from other established dimension reduction followed by classification methods (e.g., PCA-LDA, RLDA), in lieu of potentially time-consuming hyperparameter optimizations. Following validation of the method,

GA-PCALDA and GALDA were used to map composition through spectral analysis of THz Raman imaging of pharmaceutically relevant powdered blends.

CRediT authorship contribution statement

Ziyi Cao: Simulations, Formal analysis, Visualization, Writing. **Shijie Zhang:** Data collection, Instrumentation. **Youlin Liu:** Establishing cross validated framework and comparison methods, writing. **Casey J. Smith:** Method Establishment, writing. **Alex M. Sherman:** Method Establishment, Writing. **Yechan Hwang:** Writing, Visualization. **Garth J. Simpson:** Conceptualization, Funding acquisition, Supervision, Writing, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial

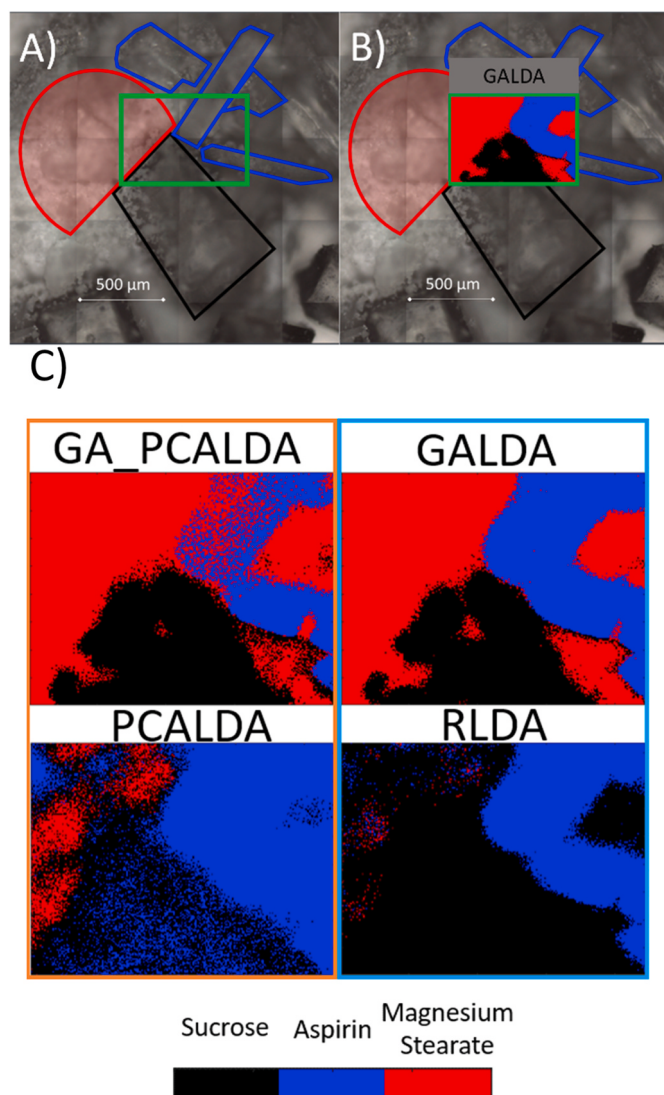


Fig. 7. Classification maps of THz Raman hyperspectral images based on GA-methods (top row) compared with their non-GA counterparts (bottom row). A) Bright-field images with preliminary crystal classification based on habit, with black, blue and red outlines indicating sucrose, aspirin and magnesium stearate respectively. B) Overlay of classified image indicating the region for which THz hyperspectral images were acquired. C) Classification maps of GA methods with their analogous non-GA counterparts. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data are publicly available on <https://github.itap.purdue.edu/Simpson-Laboratory-for-Nonlinear-Optics/GALDA-dev>

Acknowledgment

The authors gratefully acknowledge financial support for this work from the United States National Science Foundation (NSF-CHE-GOALI-1710475, NSF-CHE 2004046, NSF-CCF-1763896, and the NSF Center for Bioanalytic Metrology, an Industry/University Collaborative Research NSF-IIP-1916691). The THz Raman measurements were performed at Analytical Development at Takeda Pharmaceuticals

International Co.

Appendix ASupplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.aca.2023.341129>.

References

- [1] C.L. Philip Chen, C.Y. Zhang, Data-intensive applications, challenges, techniques and technologies: a survey on big data, *Inf. Sci.* 275 (2014) 314–347, <https://doi.org/10.1016/j.ins.2014.01.015>.
- [2] I.A.T. Hashem, I. Yaqoob, N.B. Anuar, S. Mokhtar, A. Gani, S. Ullah Khan, The rise of “big data” on cloud computing: review and open Research issues, *Inf. Syst.* 47 (2015) 98–115, <https://doi.org/10.1016/j.is.2014.07.006>.
- [3] M.G.B. Blum, M.A. Nunes, D. Prangle, S.A. Sisson, A comparative review of dimension reduction methods in approximate bayesian computation, *Stat. Sci.* 28 (2) (2013) 189–208, <https://doi.org/10.1214/12-STS406>.
- [4] M. Bahri, A. Bifet, S. Maniu, H.M. Gomes, Survey on feature transformation techniques for data streams, *Proc. Twenty-Ninth Inter. Joint Con. Artificial Intell.* (2020) 4796–4802.
- [5] M. Espadoto, R.M. Martins, A. Kerren, N.S.T. Hirata, A.C. Telea, Toward a quantitative survey of dimension reduction techniques, *IEEE Trans. Vis. Comput. Graph.* 27 (3) (2021) 2153–2173, <https://doi.org/10.1109/TVCG.2019.2944182>.
- [6] G. Chao, Y. Luo, W. Ding, Recent advances in supervised dimension reduction: a survey, *Machine Learn. Knowledge Extract.* 1 (1) (2019) 341–358, <https://doi.org/10.3390/MAKE1010020>.
- [7] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning*, second ed., Springer, 2009.
- [8] S. Wold, Pattern recognition by means of disjoint principal components models, *Pattern Recogn.* 8 (3) (1976) 127–139, [https://doi.org/10.1016/0031-3203\(76\)90014-5](https://doi.org/10.1016/0031-3203(76)90014-5).
- [9] A. Rácz, A. Gere, D. Bajusz, K. Héberger, Is soft independent modeling of class analogies a reasonable choice for supervised pattern recognition? *RSC Adv.* 8 (1) (2018) 10–21, <https://doi.org/10.1039/C7RA08901E>.
- [10] Y. Lecun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (7553) (2015) 436–444, <https://doi.org/10.1038/nature14539>, 7553.
- [11] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn.* (2016) 770–778, <https://doi.org/10.1109/CVPR.2016.90>, 2016-December.
- [12] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, T. Chen, Recent advances in convolutional neural networks, *Pattern Recogn.* 77 (2018) 354–377, <https://doi.org/10.1016/j.patcog.2017.10.013>.
- [13] G. Hinton, L. Deng, D. Yu, G. Dahl, A.R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. Sainath, B. Kingsbury, Deep neural networks for acoustic modeling in speech recognition: the shared views of four Research groups, *IEEE Signal Process. Mag.* 29 (6) (2012) 82–97, <https://doi.org/10.1109/MSP.2012.2205597>.
- [14] A.B. Nassif, I. Shahin, I. Attali, M. Azzeh, K. Shaalan, Speech recognition using deep neural networks: a systematic review, *IEEE Access* 7 (2019) 19143–19165, <https://doi.org/10.1109/ACCESS.2019.2896880>.
- [15] I. Pardoe, R.D. Cook, X. Yin, Graphical tools for quadratic discriminant analysis, *Technometrics* 49 (2) (2007) 172–183, <https://doi.org/10.1198/004017007000000074>.
- [16] S. Velilla, A method for dimension reduction in quadratic classification problems, *J. Comput. Graph Stat.* 17 (3) (2008) 572–589, <https://doi.org/10.1198/106186008341462>.
- [17] A.K. Jain, R.P.W. Duin, J. Mao, Statistical pattern recognition: a review, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (1) (2000) 4–37, <https://doi.org/10.1109/34.824819>.
- [18] A.M. Martinez, A.C. Kak, PCA versus LDA, *IEEE Trans. Pattern Anal. Mach. Intell.* 23 (2) (2001) 228–233, <https://doi.org/10.1109/34.908974>.
- [19] S. Karamizadeh, S.M. Abdullah, A.A. Manaf, M. Zamani, A. Hooman, S. Karamizadeh, S.M. Abdullah, A.A. Manaf, M. Zamani, A. Hooman, An overview of principal component analysis, *J. Signal Inf. Process.* 4 (3) (2013) 173–175, <https://doi.org/10.4236/JSIP.2013.43B031>.
- [20] O. Ledoit, M. Wolf, A well-conditioned estimator for large-dimensional covariance matrices, *J. Multivariate Anal.* 88 (2) (2004) 365–411, [https://doi.org/10.1016/S0047-259X\(03\)00096-4](https://doi.org/10.1016/S0047-259X(03)00096-4).
- [21] A. Sharma, K.K. Paliwal, Linear discriminant analysis for the small sample size problem: an overview, *Inter. J. Machine Learn. Cyber* (2014), <https://doi.org/10.1007/s13042-013-0226-9>.
- [22] R. Huang, Q. Liu, H. Lu, S. Ma, Solving the small sample size problem of LDA, *Int. Conf. Patt. Recognit.* 16 (3) (2002) 29–32, <https://doi.org/10.1109/ICPR.2002.1047787>.
- [23] Chen, L.-F.; Liao, H.-Y. M.; Ko, M.-T.; Lin, J.-C.; Yu, G.-J. A New LDA-Based Face Recognition System Which Can Solve the Small Sample Size Problem..
- [24] A. Tharwat, T. Gaber, A. Ibrahim, A.E. Hassanien, Linear discriminant analysis: a detailed tutorial, *AI Commun.* 30 (2) (2017) 169–190, <https://doi.org/10.3233/AIC-170729>.
- [25] S.J. Dixon, R.G. Brereton, Comparison of performance of five common classifiers represented as boundary methods: euclidean distance to centroids, linear discriminant analysis, quadratic discriminant analysis, learning vector

- quantization and support vector machines, as dependent on data structure, *Chemometr. Intell. Lab. Syst.* 1 (95) (2009) 1–17, <https://doi.org/10.1016/J.CHEMOLAB.2008.07.010>.
- [26] Y. Guo, T. Hastie, R. Tibshirani, Regularized linear discriminant analysis and its application in microarrays, *Biostatistics* 8 (1) (2007) 86–100, <https://doi.org/10.1093/BIOSTATISTICS/KXJ035>.
- [27] J.H. Friedman, Regularized discriminant analysis, *J. Am. Stat. Assoc.* 84 (405) (1989) 165–175, <https://doi.org/10.2307/2289860>.
- [28] A.W. Mehay, C. Cai, P.D.B. Harrington, Regularized linear discriminant analysis of wavelet compressed ion mobility spectra, *Appl. Spectrosc.* 56 (2) (2002) 223–231, <https://doi.org/10.1109/tgrs.2008.2005729>.
- [29] T. v Bandos, L. Bruzzone, G. Camps-Valls, Classification of hyperspectral images with regularized linear discriminant analysis, *IEEE Trans. Geosci. Rem. Sens.* 47 (3) (2009) 862–873, <https://doi.org/10.1109/tgrs.2008.2005729>.
- [30] R.G. Brereton, G.R. Lloyd, Partial least squares discriminant analysis: taking the magic away, *J. Chemometr.* 28 (4) (2014) 213–225, <https://doi.org/10.1002/CEM.2609>.
- [31] L.C. Lee, C.Y. Liong, A.A. Jemain, Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps, *Analyst* 143 (15) (2018) 3526–3539, <https://doi.org/10.1039/C8AN00599K>.
- [32] Y. Mallet, D. Coomans, O. deVel, Recent developments in discriminant analysis on high dimensional spectral data, *Chemometr. Intell. Lab. Syst.* 35 (2) (1996) 157–173, [https://doi.org/10.1016/s0169-7439\(96\)00050-0](https://doi.org/10.1016/s0169-7439(96)00050-0).
- [33] M.H. Kutner, C. Nachtsheim, J. Neter, W. Li, *Applied Linear Statistical Models*, McGraw-Hill Irwin, 2004, p. 1396.
- [34] L.N. Rgaard, A. Saudland, J. Wagner, J.P. Nielsen, L.M. Unck, S.B. Engelsen, Interval partial least-squares regression (IPLS): a comparative chemometric study with an example from near-infrared spectroscopy, *Appl. Spectrosc.* 54 (3) (2000).
- [35] H.M. Heise, U. Damm, P. Lampen, A.N. Davies, P.S. McIntyre, Spectral variable selection for partial least squares calibration applied to authentication and quantification of extra virgin olive oils using fourier transform Raman spectroscopy, *Appl. Spectrosc.* 59 (10) (2005) 1286–1294, <https://doi.org/10.1366/000370205774430927>.
- [36] X. Zou, J. Zhao, H. Mao, J. Shi, X. Yin, Y. Li, Genetic algorithm interval partial least squares regression combined successive projections algorithm for variable selection in near-infrared quantitative analysis of pigment in cucumber leaves, *Appl. Spectrosc.* 64 (7) (2010) 786–794, <https://doi.org/10.1366/000370210791666246>.
- [37] M.J. McShane, G.L. Coté, C. Spiegelman, Variable selection in multivariate calibration of a spectroscopic glucose sensor, *Appl. Spectrosc.* 51 (10) (1997) 1559–1564, <https://doi.org/10.1366/0003702971939118>.
- [38] Y. Kano, A. Harada, Stepwise variable selection in factor analysis, *Psychometrika* 65 (1) (2000) 7–22, <https://doi.org/10.1007/BF02294182>.
- [39] J.J.A. Workman, Review of calibration transfer practices and instrument differences in spectroscopy, *Appl. Spectrosc.* 72 (3) (2018) 340–365, <https://doi.org/10.1177/0003702817736064>.
- [40] R.N. Feudale, N.A. Woody, H. Tan, A.J. Myles, S.D. Brown, J. Ferré, Transfer of multivariate calibration models: a review, *Chemometr. Intell. Lab. Syst.* 2 (64) (2002) 181–192.
- [41] S.F.C. Soares, A.A. Gomes, M.C.U. Araujo, A.R.G. Filho, R.K.H. Galvão, The successive projections algorithm, *Trends Anal. Chem.* 42 (2013) 84–98, <https://doi.org/10.1016/J.TRAC.2012.09.006>.
- [42] L. Li, X. Pan, W. Chen, M. Wei, Y. Feng, L. Yin, C. Hu, H. Yang, Multi-manufacturer drug identification based on near infrared spectroscopy and deep transfer learning, *J. Innov. Optical Health Sci.* 13 (4) (2020), 2050016, <https://doi.org/10.1142/S1793545820500169>.
- [43] R. Zhang, H. Xie, S. Cai, Y. Hu, G. kun Liu, W. Hong, Z.Q. Tian, Transfer-learning-based Raman spectra identification, *J. Raman Spectrosc.* 51 (1) (2019) 176–186, <https://doi.org/10.1002/JRS.5750>.
- [44] L. Zhu, Y. Chen, P. Ghamisi, J.A. Benediktsson, Generative adversarial networks for hyperspectral image classification, *IEEE Trans. Geosci. Rem. Sens.* 56 (9) (2018) 5046–5063, <https://doi.org/10.1109/TGRS.2018.2805286>.
- [45] S. Yu, H. Li, X. Li, Y.V. Fu, F. Liu, Classification of pathogens by Raman spectroscopy combined with generative adversarial networks, *Sci. Total Environ.* 726 (2020), <https://doi.org/10.1016/J.SCITOTENV.2020.138477>.
- [46] I.v. Tetko, D.J. Livingstone, A.I. Luik, Neural network studies. 1. Comparison of overfitting and overtraining, *J. Chem. Inf. Comput. Sci.* 35 (5) (1995) 826–833, https://doi.org/10.1021/Ci00027A006/SUPPL_FILE/Ci826.PDF.
- [47] C. Zhang, S. Bengio, M. Hardt, B. Recht, O. Vinyals, Understanding deep learning requires rethinking generalization, *Commun. ACM* 64 (3) (2016) 107–115, <https://doi.org/10.1145/3446776>.
- [48] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, *Commun. ACM* 60 (6) (2017) 84–90, <https://doi.org/10.1145/3065386>.
- [49] W. Zhao, Research on the deep learning of the small sample data based on transfer learning, *AIP Conf. Proc.* 1864 (1) (2017), 020018, <https://doi.org/10.1063/1.4992835>.
- [50] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. C. Courville, Y. Bengio, Generative adversarial nets, *Adv. Neural Inf. Process. Syst.* 3 (2014) 2672–2680.
- [51] J. Yosinski, J. Clune, Y. Bengio, H. Lipson, How transferable are features in deep neural networks? *Adv. Neural Inf. Process. Syst.* (2014) 3320–3358.
- [52] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.* 378 (C) (2018) 686–707, <https://doi.org/10.1016/J.JCP.2018.10.045>.
- [53] I. Goodfellow, Tutorial: generative adversarial networks, *Proc. Neural Inform. Process. Syst.* (2017).
- [54] T. Karras, T. Aila, S. Laine, J. Lehtinen, Progressive growing of GANs for improved quality, stability, and variation, *Int. Conf. Learn. Represent.* (2018).
- [55] R. Webster, J. Rabin, L. Simon, F. Jurie, Detecting overfitting of deep generative networks via latent recovery, *IEEE Conf. Computer Vision Patt. Recognit.* (2019) 11265–11274, <https://doi.org/10.1109/CVPR.2019.01153>, 2019-June.
- [56] N. Buzgar, A.I. Apopei, Romanian database of Raman spectroscopy. <http://Rdrs.Ro>, 2009.
- [57] S. Zhang, Z. Song, G.M.D.P. Godaliyadda, D.H. Ye, A.U. Chowdhury, A. Sengupta, G.T. Buzzard, C.A. Bouman, G.J. Simpson, Dynamic sparse sampling for confocal Raman microscopy, *Anal. Chem.* 90 (7) (2018) 4461–4469, <https://doi.org/10.1021/ACS.ANALCHEM.7B04749>.
- [58] A. Savitzky, M.J.E. Golay, Smoothing and differentiation of data by simplified least squares procedures, *Anal. Chem.* 36 (8) (1964) 1627–1639, <https://doi.org/10.1021/AC60214A047>.
- [59] A. Biancolillo, F. Marini, Chemometric methods for spectroscopy-based pharmaceutical analysis, *Front. Chem.* 6 (NOV) (2018) 576, <https://doi.org/10.3389/FCHEM.2018.00576>.
- [60] I.T. Jolliffe, J. Cadima, Principal component analysis: a review and recent developments, *Phil. Trans. Math. Phys. Eng. Sci.* 374 (2065) (2016), <https://doi.org/10.1098/RSTA.2015.0202>.
- [61] I.J. Goodfellow, J. Shlens, C. Szegedy, Explaining and harnessing adversarial examples, 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc. (2014).
- [62] E. Pavlou, N. Kourkoumelis, Deep adversarial data augmentation for biomedical spectroscopy: application to modelling Raman spectra of bone, *Chemometr. Intell. Lab. Syst.* 228 (2022), 104634.