Chemometric Strategies for the Detection of Bromazolam and Xylazine in Illicit Opioids Using Surface-Enhanced Raman and Infrared Spectroscopy

by

Rebecca Robinton Martens
B.Sc., University of Victoria, 2021

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Chemistry

University of Victoria

Chemometric Strategies for the Detection of Bromazolam and Xylazine in Illicit Opioids Using Surface-Enhanced Raman and Infrared Spectroscopy

by

Rebecca Robinton Martens
B.Sc., University of Victoria, 2021

Supervisory committee

Dr. Dennis K. Hore, Supervisor
(Department of Chemistry)

Dr. Alexandre Brolo, Committee Member
(Department of Chemistry)

ABSTRACT

The detection of trace adulterants in opioid samples is an important aspect of drug checking, a harm reduction measure that is required as a result of the variability and unpredictability of the illicit drug supply. While many analytical methods are suitable for such analysis, community-based approaches require techniques that are amenable to point-of-care applications with minimal sample preparation and automated analysis. We demonstrate that surface-enhanced Raman spectroscopy, combined with a random forest classifier, is able to detect the presence of two common sedatives, bromazolam (0.32–36% w/w) and xylazine (0.15–15% w/w), found in street opioid samples collected as a part of a community drug checking service. The Raman predictions, benchmarked against mass spectrometry results, exhibited high specificity for the compounds of interest (88% for bromazolam, 96% for xylazine) and sensitivity (88% for bromazolam, 92% for xylazine). We additionally provide evidence that this exceeds the performance of a more conventional approach using infrared spectral data acquired on the same samples. This demonstrates the feasibility of surface-enhanced Raman spectroscopy for point-of-care analysis of challenging multi-component samples containing trace adulterants.

Surface-enhanced Raman spectroscopy and infrared spectroscopy were integrated into two data fusion strategies - hybrid (concatenated spectra) and high level (fusion of high outputs from both models) - to enhance the predictive accuracy for xylazine detection. Three advanced chemometric approaches - random forest, support vector machine, and *k*-nearest neighbor algorithms - were employed and optimized using a 5-fold cross-validation grid search for both fusion strategies. Validation results identified the random forest classifier as the optimal model for both fusion strategies, achieving high sensitivity (88% for hybrid, 84% for high level) and specificity (88% for hybrid, 92% for high level). We demonstrate the enhanced practicality of the high level fusion approach, effectively leveraging the surface-enhanced Raman data with a 90% voting weight, without compromising prediction accuracy when combined with infrared spectral data. This highlights the viability of a multi-instrumental approach using data fusion and random forest classification to improve the detection of various components in complex opioid samples for community-based drug checking.

Contents

Supervisory Committee ii
Abstract iv
Contents v
List of Figures viii
List of Tables xii
List of Symbols and Definitions xvi
Acknowledgements xix

1 Introduction **1**
- 1.1 The Overdose Crisis and the Need for a Harm Reduction Approach 1
  - 1.1.1 The Role of Drug Checking in Harm Reduction 2
  - 1.1.2 Challenges in an Unregulated Supply and Growing Presence of Adulterants 2
    - 1.1.2.1 Two Adulterants of Concern 3
- 1.2 Analytical Techniques for Drug Checking 4
  - 1.2.1 Fourier Transform Infrared (FTIR) and Raman Spectroscopy 5
  - 1.2.2 Surface-Enhanced Raman Spectroscopy (SERS) 5
- 1.3 Machine Learning in Spectral Analysis 6
- 1.4 Aims and Objectives 7

2 Methods **8**
- 2.1 Materials and Sample Selection 8

2.2 Sample Preparation and Data Acquisition 10
  2.2.1 SERS Measurements 10
  2.2.2 FTIR Measurements 10
  2.2.3 Mass Spectrometry Measurements 11
  2.2.4 Spectral Data Handling 11
2.3 Outlier Detection Method 13
2.4 Models 13
  2.4.1 Random Forest 13
  2.4.2 Support Vector Machine 16
  2.4.3 *k*-Nearest Neighbors 16
2.5 Model Optimization 17

3 Trace Detection of Adulterants in Illicit Opioid Samples Using Surface-Enhanced Raman Scattering and Random Forest Classification 20

3.1 Introduction 20
3.2 Model Development 21
  3.2.1 Hyperparameter Tuning for a Random Forest Binary Classifier . . . 21
  3.2.2 Targeted Secondary Model Development for SERS Data 21
    3.2.2.1 Low Confidence Range 23
    3.2.2.2 Model Tuning 23
    3.2.2.3 Model Application 23
3.3 Bromazolam Detection 25
  3.3.1 Model Optimization 25
  3.3.2 SERS Model Visualization 28
  3.3.3 Model Application and Performance 29
3.4 Xylazine Detection 34
  3.4.1 Model Optimization 34
  3.4.2 SERS Model Visualization 35

3.4.3 Model Application and Performance 38
3.5 Consequence for Drug Checking 40
3.6 Conclusions 43

4 Optimized Machine Learning Approaches to SERS–IR Data Fusion Strategies for Trace Detection of Xylazine Adulterated in Opioids 44
4.1 Introduction 44
4.2 Data Fusion Strategies 46
4.2.1 Hybrid Data Fusion 46
4.2.1.1 Low Level Data Fusion 46
4.2.1.2 Mid Level Data Fusion 46
4.2.2 High Level Data Fusion 48
4.3 Optimization of RF, SVM, and KNN Models 56
4.3.1 Hybrid Data Fusion Model Development 56
4.3.2 High Level Data Fusion Model Development 57
4.3.2.1 Random Forest Models 57
4.3.2.2 Support Vector Machine Models 57
4.3.2.3 *k*-Nearest Neighbors Models 58
4.3.2.4 Weighted Voting Ensemble 58
4.4 Analysis of Models 60
4.5 Conclusions 65

5 Conclusions 68
5.1 Summary of Work 68
5.2 Future Work 70

References 72

List of Figures

1.1 Chemical structures of (a) bromazolam and (b) xylazine. 4

2.1 Sample preparation workflow for SERS measurements. 10

2.2 Mean spectra of samples in the training set ($n = 168$) for (a) SERS, (b) IR, and (c) concatenated SERS–IR fused data. SERS spectra in the stokes region of 300–2000 cm^{-1} is represented in blue and IR spectra in the wavenumber region of 650–4000 cm^{-1} is represented in red. 12

2.3 Outlier detection using PCA and robust Mahalanobis distance for (a) SERS, (b) IR, and (c) concatenated SERS–IR fused spectral data. Each subplot illustrates the scatter plot of the first two principal components, colour-coded by Mahalanobis distance. Overlaid contour plots represent the Mahalanobis distances, with identified outliers marked with x. Gradient colours of the data points range from dark blue (low variance) to red (high variance). 14

3.1 Receiver operator characteristic (ROC) curve of secondary SERS RF binary classification models for (a) bromazolam on LCR subset ($n = 11$) and (b) xylazine on LCR subset ($n = 10$). The curves illustrate the true positive rate and false positive rate across various classification thresholds for adulterant detection. Optimal threshold selection for balancing sensitivity and specificity and area under the curve (AUC) is highlighted. 26

3.2 Confusion matrices for SERS secondary RF model prediction results of LCR samples for (a) bromazolam ($n = 11$) and (b) xylazine ($n = 10$). . . . 27

3.3 (a) Mean spectral data for bromazolam-positive and bromazolam-negative samples in the training set overlaid with vertical blue lines representing the top 20 features extracted by the model. Spectral regions with feature presence are highlighted. (b) Detailed view of the 650–750 cm^{-1} region with stacked mean spectra for different opioid drug combinations: opioid-only (black), adulterants other than bromazolam (red), bromazolam and additional adulterants (green), and bromazolam-only (blue). 30

3.4 Mean SERS spectral data for bromazolam-positive (blue) and bromazolam-negative (red) samples in the training set compared to a single component bromazolam street sample (black). Vertical lines represent the top 20 features extracted by the bromazolam SERS model. The characteristic spectral region of 650–750 cm^{-1} is highlighted. 31

3.5 A comparison between the bromazolam SERS model and bromazolam IR model prediction results on the test set ($n = 50$). The ROC curves illustrate the true positive rate and false positive rate across various classification thresholds for bromazolam detection. Optimal threshold selection for balancing sensitivity and specificity and area under the curve (AUC) is highlighted for both models. 31

3.6 (a) Mean IR spectral data for bromazolam-positive (blue) and bromazolam-negative (red) samples in the training set compared to a single component bromazolam street sample (black). Vertical lines represent the top 20 features extracted by the bromazolam IR model. (b) An enhanced view of the stacked spectra with the spectral region of 750–850 cm^{-1} highlighted to illustrate the characteristic peaks of bromazolam. 33

3.7 (a) Mean spectral data for xylazine-positive and xylazine-negative samples in the training set overlaid with vertical blue lines representing the top 20 features extracted by the model. (b) Detailed view of the 1050–1150 cm^{-1} region with overlaid mean spectra for different opioid drug combinations: opioid-only (black), adulterants other than xylazine (red), xylazine and additional adulterants (green), and xylazine-only (blue). 36

3.8 Detailed view of the 1350–1600 cm^{-1} region with overlaid mean spectra for different opioid drug combinations: opioid-only (black), adulterants other than xylazine (red), xylazine and additional adulterants (green), and xylazine-only (blue). Vertical blue line representing the 1495 cm^{-1} feature extracted in the region by the model. 37

3.9 Mean SERS spectral data for xylazine-positive (blue) and xylazine-negative (red) samples in the training set compared to a single component xylazine street sample (black). Vertical lines represent the top 20 features extracted by the xylazine SERS model. The characteristic spectral region of 1050–1150 cm^{-1} is highlighted. 37

3.10 A comparison between the xylazine SERS model and xylazine IR model prediction results of xylazine on the test set ($n = 50$). The ROC curves illustrate the true positive rate and false positive rate across various classification thresholds for xylazine detection. Optimal threshold selection for balancing sensitivity and specificity and area under the curve (AUC) is highlighted for both models. 38

3.11 (a) Mean IR spectral data for xylazine-positive (blue) and xylazine-negative (red) samples in the training set compared to a single component xylazine street sample (black). Vertical lines represent the top 20 features extracted by the xylazine IR model. (b) An enhanced view of the stacked spectra in the 650–1800 cm^{-1} spectral region with all features extracted by the model highlighted in green. 41

4.1 Flowchart of hybrid and high level fusion strategies. 47

4.2 A comparison between the RF model, SVM model and KNN model prediction results of xylazine on the test set ($n = 50$) for the (a) hybrid and (b) high level data fusion strategies. The ROC curves illustrate the true positive rate and false positive rate across various classification thresholds for xylazine detection. Optimal threshold selection for balancing sensitivity and specificity and area under the curve (AUC) is highlighted for all models. 61

List of Tables

2.1 Compounds identified by paper spray mass spectrometry and benzodiazepine immunoassay test strips in training ($n = 168$) and testing ($n = 50$) sets. 9

2.2 Investigation of hyperparameters using a 5-fold cross-validation grid search for development of random forest (RF), support vector machine (SVM), and k-nearest neighbors (KNN) classification models. Showing hyperparameter descriptions and the parameter combinations tested for all models. . 19

3.1 Investigation of RF hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search for (a) bromazolam and (b) xylazine, showing the combinations tested and the optimal set of parameters selected based on F1 score. 22

3.2 Investigation of RF hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search on the restricted SERS spectral region training libraries for bromazolam (650–750 cm^{-1}) and xylazine (1050–1150 cm^{-1}). This highlights the combinations tested and the optimal set of parameters selected for the targeted secondary models based of F1 score. 24

3.3 Bromazolam prediction results for samples in the test set ($n = 50$) for SERS and IR models. 32

3.4 Summary of accuracy, precision, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for opioid samples ($n = 50$) tested with RF models. Model prediction results for IR and SERS data are shown for (a) bromazolam detection and (b) xylazine detection. 34
3.5 Xylazine prediction results of the test set ($n = 50$) for the SERS and IR models. 39
3.6 Complete breakdown of validation set ($n = 50$) sample composition as identified by paper spray mass spectrometry and benzodiazepine immunoassay test strips. Concentration values of bromazolam (green) and xylazine (blue) are described next to true positive (TP), true negative (TN), false positive (FP), and false negative (FN) prediction results of the SERS and IR models. Correct predictions by the respective models are shown in white, while incorrect predictions are in grey. 42

4.1 Investigation of RF hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search for the detection of xylazine using the hybrid SERS–IR data fusion method. Showing the combinations tested and the optimal set of parameters selected based on F1 score. 49
4.2 Investigation of SVM hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search for the detection of xylazine using the hybrid SERS–IR data fusion method. Showing the combinations tested and the optimal set of parameters selected based on F1 score. 50
4.3 Investigation of KNN hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search for the detection of xylazine using the hybrid SERS–IR data fusion method. Showing the combinations tested and the optimal set of parameters selected based on F1 score. 51

4.4 Investigation of RF hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search for the detection of xylazine using SERS and IR data. Showing the combinations tested and the optimal set of parameters selected based on F1 score. 53

4.5 Investigation of SVM hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search for the detection of xylazine using SERS and IR data. Showing the combinations tested and the optimal set of parameters selected based on F1 score. 54

4.6 Investigation of KNN hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search for the detection of xylazine using SERS and IR data. Showing the combinations tested and the optimal set of parameters selected based on F1 score. 55

4.7 Optimized parameters and F1 scores for all models built with SERS, IR, and SERS–IR training data for the detection of xylazine. Optimal weights selected for SERS and IR model contribution and their respective F1 scores are detailed for the construction of high level classification models. 59

4.8 Xylazine prediction results of the test set ($n = 50$) for the RF, SVM, and KNN models developed using the (a) hybrid SERS–IR data fusion method and (b) high level SERS–IR data fusion method. 62

4.9 Performance metrics summary for RF, SVM, and KNN models using hybrid SERS–IR data fusion method for xylazine detection. Summary of the Area Under the Curve (AUC), optimal threshold, accuracy, precision, sensitivity, and specificity for opioid samples ($n = 50$) tested with RF, SVM, and KNN models. 63

4.10 Performance metrics summary for RF, SVM, and KNN models using high level SERS–IR data fusion method for xylazine detection. Presenting the selected voting classifier and the weights of the individual SERS and IR models that created the high level fused RF, SVM, and KNN models. Summary of the area under the curve (AUC), optimal threshold, accuracy, precision, sensitivity, and specificity for opioid samples ($n = 50$) tested with high level fused models. 64

4.11 Complete breakdown of validation set ($n = 50$) sample composition as identified by paper spray mass spectrometry and benzodiazepine immunoassay test strips. Concentration values of xylazine (blue) are described next to true positive (TP), true negative (TN), false positive (FP), and false negative (FN) prediction results of the hybrid and high level SERS–IR data fusion models. Correct predictions by the respective RF, SVM, and KNN models are shown in white, while incorrect predictions are in grey. 66

List of Symbols and Definitions

symbol	**definition**
PWUD	people who use drugs
MS	mass spectrometry
PS-MS	paper spray mass spectrometry
ATR	attenuated total reflection
IR	infrared
FTIR	Fourier-transform infrared
SERS	surface-enhanced Raman scattering
BC	British Columbia
USA	United States of America
PCA	principal component analysis
MCD	minimum convariance determinant
Deriv	order of derivatives
Norm	normalization
SNV	standard normal variate
Min-Max	min-max normalization
ROC	receiver operating characteristic
AUC	area under the curve
ML	machine learning
RF	random forest
SVM	support vector machine
KNN	*k*-nearest neighbors
% w/w	ratio of the target drug to the total sample weight
% error	relative error
χ^2	chi-squared

FIV	feature importance variable
LCR	low confidence range
LoD	limit of detection

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Dennis Hore, for all his support and instruction throughout my degree. He has been a fierce advocate for both graduate research and the project as a whole. I would also like to extend my appreciation to Dr. Bruce Wallace and Dr. Marilou Gagnon for introducing me to harm reduction and connecting me with this project. Thank you to Dr. Collin Kielty, Zoe Riell, Oscar Sandford, and Derek Robinson for all their help during my time at the University of Victoria, they have all been instrumental to facilitating my research and figuring out how to **successfully** run my code. Dr. Lea Gozdzialski has been one of the most patient and intelligent people I have had the pleasure of working with, she has helped me navigate the transition into research and I am deeply thankful for her invaluable coding assistance and thoughtful instruction. Allie Miskulin has been another huge part of my degree, her encouragement, support, and kindness has made this entire journey possible. I will always be in awe of her skills as a researcher and I am eternally grateful for her friendship (thank you for learning what fluorescence is with me). I would like to extend another huge acknowledgment to Ella Newman for her contributions to data acquisition and for making this research possible. I know she will approach her own research with the same care and intelligence she has demonstrated in mine. In addition to the research side of things, I want to thank both Vivi Le and Kavin Bhuvan for their help towards this project and the genuine kindness they have shown me during my time here. A huge shout out to the entire SFG group at Petch for helping me manage those first few months on the project, they are all amazing.

Other supporters, friends, and family I would like to thank:

First and foremost, I would like to thank my big sister, Thea Robinton, who has been my steadfast supporter since my first day on Earth. Her intelligence and skill are unmatched and she has been a guiding light throughout my entire schooling. She has always been

my biggest advocate, ready to help at a moment's notice. I would not be where I am today without her, and I am forever in her debt. I also wish to thank my grandfather, Dr. Alan Arthur Boulton, the catalyst for my interest in the sciences. He, along with my grandmother Anne Boulton, have been a huge inspiration and a driving force behind my continued education. I am immensely grateful to my mom, Lana Robinton, and my dad, Gary Martens. I wouldn't be here without them, and I am beyond thankful for their support, interest, and help in everything I have done in life. I also want to thank the remainder of my family, including my brother Aaron Martens, for their continued support. Additionally, I extend my gratitude to Corinne, Raouf, and Nada Saad, who have always shown interest in my work and kindness to me throughout the time I have known them. To my amazing friend, Chelsea Pacheco, thank you for your unwavering friendship. You have been my rock, and I cannot thank you enough. Thank you to Daniel Barker-Rothschild for starting me on my master's journey and for allowing me to pester you with questions about chemometrics. I am also so grateful to Alexander Birch, Ross McBee, and Ryan Moro for their friendship and support. And a huge thank you to all three of my beautiful cats.

I wish to give special thanks to Rami Saad. He has been instrumental in connecting me to the field of harm reduction, educating me, and being a significant source of my confidence. I have nothing but the utmost respect and adoration for him, both as a partner and as a professional. The effort and care he brings to his work have resulted in so many lives being saved and improved. I strive to reflect just an ounce of his character in my life and I hope to continue working in this field and learning more from him.

Finally, I would like to thank everyone at the Vancouver Island Drug Checking Project and the community we serve. I dedicate this work to all the lives lost to the overdose crisis and to those who loved and mourn them. I hope this work can contribute in any way to the prevention of future needless deaths and provide further evidence toward the lifesaving impacts of harm reduction policies.

Chapter 1

Introduction

1.1 The Overdose Crisis and the Need for a Harm Reduction Approach

Over the past decade, there has been a rise in deaths attributed to drug overdose, as communities across the globe struggle to understand and respond to variability in an unregulated illicit drug supply. Since the government of British Columbia, Canada declared the overdose epidemic a public health emergency in 2016, more than 13,000 people have died, making it the leading cause of unnatural death for persons aged 10–59 [1]. The illicit drug market experiences a continual evolution of active components and adulterants [2–4], making it increasingly difficult for people who use drugs (PWUD) to make informed decisions and manage their dose. Synthetic opioids, specifically fentanyl and its analogues, have received a significant amount of attention since they are often the primary active in the opioid supply [5, 6], and have been involved in 85% of unregulated drug deaths in 2023 in British Columbia (BC) [1]. The main combatant to opioid related overdoses has been naloxone, a lifesaving medication for the reversal of acute opioid intoxication [7–9]. In response to the current drug market and the increase in opioid-related overdoses, the injectable form of naloxone has been made freely and publicly accessible in Canada. Naloxone reverses opioid overdoses by acting as a pure competitive antagonist at the (μ)-opioid receptor, displacing opioids and rapidly reversing their effects [7], however it does not relieve overdose effects from any non-opioid drugs still present in the body. As the unregulated market evolves, it has continued to introduce powerful and potent adulterants into the drug supply, which do not respond to naloxone, translating to complicating overdose response protocols and posing a substantial safety concern to PWUD.

1.1.1 The Role of Drug Checking in Harm Reduction

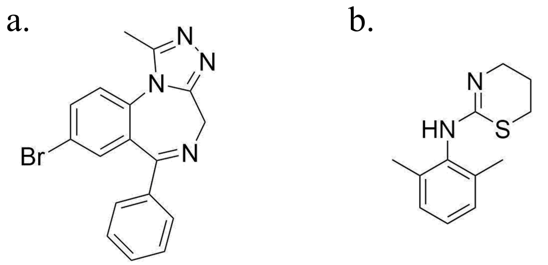
The response to the growing overdose crisis in BC and Canada has led to the adoption of several harm reduction practices. Harm reduction as a concept, aims to mitigate the negative issues surrounding drug use experienced by PWUD through compassionate programs and strategies. In recent years in BC, this has included the use of safe injection sites, outreach and support programs, fentanyl test strips, naloxone distribution and training, and of course, drug checking. Drug checking has a valuable role within responses to the overdose crisis as a harm reduction measure [10–13]. It aims to use analytical instruments and techniques in a point-of-care setting to provide qualitative and sometimes quantitative information to service users about the contents of their samples in a safe, anonymous, and judgement free manner. This service encourages de-stigmatisation of PWUD and enables informed consent of drug use, granting autonomy and enhancing safety. Given the irregularities in sample composition within an ever evolving drug market, detecting novel and potentially co-competing drug compounds in complex mixtures is a fundamental area of research for drug checking.

1.1.2 Challenges in an Unregulated Supply and Growing Presence of Adulterants

Currently, there is a growing focus on detecting potent and therefore low concentration sedatives that appear as adulterants in opioid samples. Of particular importance is the presence of benzodiazepines, a broad class of depressant drugs that slow activity in the brain and nervous system [14], making them a popular addition to down samples which have a similar psychoactive effect. Opioids, specifically fentanyl, have a very potent effect but a short duration of action [15], meaning that many PWUD will continue to redose upon exhaustion of active effects. Conversely, the longer half-lives of benzodiazepine adulterants results in continuous and accumulating presence within the user. Thus, the co-consumption of benzodiazepines and opioids has been found to significantly increase the risk of respiratory depression, overdose, and death [1, 16, 17]. Benzodiazepines have quickly grown to being the most common adulterants found in the down supply, appearing in 47% of opioid-down samples checked by Substance, the Vancouver Island Drug Checking Project [18], and contributing to 52% of complicated overdoses in 2023 in BC [1]. The ability to detect and differentiate benzodiazepine sub-classes is, therefore, a primary concern in dose preparation for PWUD.

1.1.2.1 Two Adulterants of Concern

Bromazolam (Figure 1.1a) is a designer benzodiazepine [19] and the brominated version of alprazolam (Xanax). It was detected in 33.9% of down samples in Victoria, BC (2023) [18], making it the most common adulterant in opioid-down samples for the past two years. Due to its extended half-life (that is comparable to its structurally similar counterpart alprazolam) and its growing prevalence worldwide [20, 21], there is substantial cause for research focused on identifying bromazolam in complex mixtures. Another growing concern in the illicit drug supply is xylazine (Figure 1.1b), a veterinary tranquilliser with no human studies documenting its long term impact or relative experience when combined with opioids [22–24]. This powerful sedative has been linked to skin sores and infections at the injection point and has been found to adulterate opioid samples across BC and even more pervasively in the United States and Puerto Rico [24–27]. Consequently, high confidence detection of adulterants like bromazolam and xylazine in trace amounts is a crucial area of interest for drug checking techniques.



Bromazolam

Xylazine

Figure 1.1: Chemical structures of (a) bromazolam and (b) xylazine.

1.2 Analytical Techniques for Drug Checking

Although the analysis methods employed for drug checking range from simple antibody-based test strips [14, 28, 29] to lab-based mass spectrometry [30–32], drug checking using analytical instruments aims to detect all components present in street samples. The use of benzodiazepine and newly emerging xylazine immunoassay test strips can provide rapid qualitative information at a cut off concentration of 300 ng/mL and 1000 ng/mL, respectively [33]. While they are a practical addition in a point-of-care setting, they are not capable of providing any relative sample mixture information or differentiating between benzodiazepine sub-classes which have varied duration and potencies. Other more advanced and discriminate techniques, such as mass spectrometry, can provide meaningful quantitative information and is considered the gold standard in drug analysis [34]. However, the associated costs of the instrument, maintenance, and method development make it less feasible for integration into most point-of-care drug checking sites. Techniques capable of trace analysis specifically in point-of-care settings are required. However, it remains challenging for a single instrument to exclusively meet the accessibility requirements for community drug checking [35–37], namely a simple, low cost, field-portable method that requires minimal operator training.

1.2.1 Fourier Transform Infrared (FTIR) and Raman Spectroscopy

Fourier transform infrared (FTIR) spectroscopy has largely met those requirements, revealing the bulk make-up of samples as well as many high concentration active components [38–40]. However, it is widely regarded as having a limit of detection of 5% w/w for many drugs of interest [41–43], therefore making it unsuitable for trace detection. Portable Raman spectroscopy has also found application in drug checking [39, 44–46], but with roughly the same limit of detection of FTIR and the additional complexity of interfering fluorescence, particularly from opioids [43, 47, 48]. The inability to detect potentially harmful adulterants and a heavy reliance on manual interpretation present significant limitations to the widespread application of FTIR or Raman as exclusive drug checking methods.

1.2.2 Surface-Enhanced Raman Spectroscopy (SERS)

Surface-enhanced Raman spectroscopy (SERS) has a potential to meet this need, as the amplification of Raman signals through localized electromagnetic fields can provide access to suitably low concentrations of interest for drug analysis [49–53]. SERS has demonstrated success for trace detection of adulterants in complex biological samples [54] and laboratory mixtures of opioid samples [55–58]. Despite the promise of SERS, significant method development is required for it to reliably detect and differentiate trace adulterants in street drugs, often involving interferences in complex mixtures [3, 59]. Progress in this area has simultaneously focused on the platform itself, including optimizing the design of SERS substrate, both in general and for specific analytes [60–65], and the associated analysis methods [60, 66, 67]. The success of manual spectral interpretation is often limited for SERS [68], and there is a significant interest in the application of chemometric approaches [52, 69–72].

1.3 Machine Learning in Spectral Analysis

Unsupervised machine learning (ML) methods such as principal component analysis (PCA), have been demonstrated for the detection of fentanyl using SERS spectra of binary drug mixtures [73]. However, given the variability of SERS measurements, these unsupervised ML techniques often struggle to differentiate trace components as the complexity of the sample matrix increases [74]. Supervised ML techniques have more advanced decision-making abilities and are therefore useful for more complex samples with multiple components [75–77]. Among these, random forest (RF) classification [78] employs multiple decision trees to make predictions based on characteristic features, thereby improving overall accuracy and generalization capabilities. Support vector machine (SVM) is another powerful supervised learning method that can effectively classify complex SERS spectra by finding the optimal hyperplane that separates different classes [79, 80], enhancing the detection of trace substances even in complex mixtures. Similarly, *k*-nearest neighbors (KNN) leverages the similarity between spectra by assigning class labels based on the most common labels of the nearest neighbours, making it particularly useful for non-linear classification tasks with complex sample matrices [80,81]. Previous work has centered on the analysis of laboratory drug mixtures, showcasing notable success when using SERS for identifying fentanyl in heroin down to sub-microgram concentrations [55]. Machine learning algorithms including RF, SVM and KNN have proven to enhance the identification of components within increasingly complicated drug mixtures with high accuracy when employing traditional Raman spectroscopy techniques [82]. This illustrates the synergy of combining the heightened sensitivity of SERS with the robust predictive capabilities of RF, SVM, and KNN for the detection of trace adulterants within real-world samples in the illicit supply.

1.4 Aims and Objectives

The primary objective of this work is to develop and optimize supervised machine learning models using SERS and IR spectral data from real-world opioid samples to effectively detect trace adulterants in illicit drug mixtures. The training metrics and prediction results of models built with SERS, IR, and two data fusion strategies are examined. This includes investigating hyperparameter tuning and spectral preprocessing techniques to optimize SERS and IR spectral information for detecting bromazolam or xylazine in complex opioid samples.

Chapter 3 presents optimized random forest classifiers trained with SERS and IR spectral data for detecting bromazolam and xylazine, demonstrating the superiority of the sensitive SERS technique combined with a robust RF classifier. This chapter details model development, visualizes RF features in SERS data, and implements restricted SERS spectral regions to increase model sensitivity and improve performance.

Chapter 4 explores two data fusion strategies to combine prediction information from the high-performing SERS platform with IR spectroscopy. This chapter assesses xylazine prediction results using RF, SVM, and KNN classifiers. It highlights the practical combination of a commonly used drug-checking technology (FTIR) with the sensitivity of an enhanced spectroscopic platform (SERS), demonstrating the applicability of both platforms within point-of-care drug-checking services.

Chapter 2

Methods[1]

2.1 Materials and Sample Selection

50 nm gold nanoparticle (AuNP) solution (BBI Solutions, UK), magnesium sulfate anhydrous certified powder ($MgSO_4$, Fisher Chemical), and deionized water (18.2 MΩ·cm, Barnstead Nanopure water, Thermo Fisher Scientific) were used for SERS sample preparation. All drug samples used in this study were acquired from Substance, the Vancouver Island Drug Checking Project [40, 83], as part of our ongoing drug-checking service in Victoria BC, Canada. All samples were thoroughly crushed and mixed using a spatula before being analyzed. The samples were analyzed to determine drug composition through benzodiazepine immunoassay test strips (Rapid Response, BTNX) and paper-spray mass spectrometry (TSQ Fortis^{TM}, Thermo Fisher Scientific). The samples selected for this study ($n = 218$) were received in powdered form and were identified as opioid samples, with either fentanyl, fluorofentanyl, or heroin as the primary active components. In addition to the main active opioids, the samples chosen had diverse combinations of cutting agents, opioid analogues, and adulterants. The breakdown of the components identified in the selected samples is shown in Table 2.1.

A subset of samples ($n = 50$) was selected as the test set for model validation using both

---

[1] Parts of this chapter have appeared in Martens, R. R.; Gozdzialski, L.; Newman, E.; Gill, C.; Wallace, B.; Hore, D. K. ”Trace Detection of Adulterants in Illicit Opioid Samples Using Surface-Enhanced Raman Scattering and Random Forest Classification.“ *Anal. Chem.* **96**, 12277–12285 (2024). Copyright 2024 American Chemical Society.

Compound	Training set ($n = 168$)	Testing set ($n = 50$)
	($n =$)	($n =$)
Fentanyl	134	46
Caffeine	143	45
Fluorofentanyl	106	33
Bromazolam	65	25
Xylazine	41	25
ANPP	19	10
Chloroisobutyryl fentanyl	14	2
Benzodiazepine (undifferentiated)	7	4
Etizolam	6	3
Flubromazepam	5	2
Heroin	4	2
Flualprazolam	4	2
Isobutyryl fentanyl	3	2
Acetylfentanyl	3	1
Acetylcodeine	2	0
Carfentanil	2	0
Metonitazene	1	0
N-desethyl isotonitazene	1	0
Flubromazolam	1	2
Morphine	1	0
Acetylmorphine	1	0
4-Anilino-boc-piperidine	0	1

Table 2.1: Compounds identified by paper spray mass spectrometry and benzodiazepine immunoassay test strips in training ($n = 168$) and testing ($n = 50$) sets.

SERS and FTIR-based platforms. Among these, $n = 25$ samples contained bromazolam (median 6.59% w/w) and $n = 25$ samples contained xylazine (median 4.35% w/w). These concentrations refer to the ratio of the target drug to the total sample weight. The specific makeup of the testing group was organized to maintain the complexity of these multi-component opioid samples as seen in Table 2.1, with each sample of the $n = 50$ subset categorized into one of four groups: containing both bromazolam and xylazine ($n = 13$), only bromazolam ($n = 12$), only xylazine ($n = 12$), and those containing neither adulterant ($n = 13$).

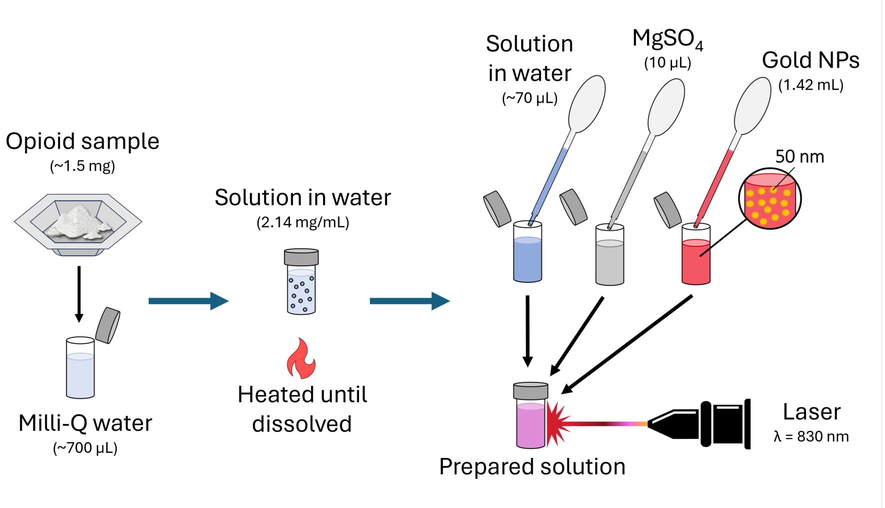


Figure 2.1: Sample preparation workflow for SERS measurements.

2.2 Sample Preparation and Data Acquisition

2.2.1 SERS Measurements

SERS measurements were performed using a portable Raman spectrometer (Resolve, Agilent Technologies, Santa Clara, USA) equipped with an 830 nm excitation source at 100–475 mW variable power as described previously [59]. Approximately 1.5 mg of powdered opioid sample was added to ≈ 700 μL of deionized water. The solution was then vortexed and heated until dissolved. 70 μL of the sample solution was added to 1.42 mL of the 50 nm gold nanoparticle solution. Solutions were vortexed for 30 s and 10 μL of 1 M $MgSO_4$ was added as an aggregating agent. The final 1.5 mL solution was vortexed a final time for approximately 10 s, and the complete work flow is detailed visually in Figure 2.1. Spectra were recorded between 200–2000 cm^{-1} at a constant laser power, 2 s integration time, and 10 averages. Automated baseline, fluorescence, and cosmic ray correction procedures were applied by instrument firmware [84].

2.2.2 FTIR Measurements

FTIR measurements were recorded with a portable FTIR spectrometer (Agilent 4500a, Santa Clara, USA) fitted with a single-bounce 45° diamond attenuated total reflectance

(ATR) sampling system. The infrared (IR) spectra for the powdered opioid samples were recorded in the 650–4000 cm^{-1} region using a resolution of 4 cm^{-1}.

2.2.3 Mass Spectrometry Measurements

All samples used in this study were analyzed using a TSQ Fortis^{TM} triple quadrupole mass spectrometer and a VeriSpray^{TM} Paper Spray ion source (Thermo Fisher Scientific, San Jose, CA, USA). Information regarding the method, calibration and data analysis can be found in recent publications [30,31,85,86]. Briefly, 0.5–2.1 mg of the sample was dissolved into 1.2 mL of methanol and vortexed to prepare a solution in the range of 1 mg/mL. 1 μL of the solution was diluted in an internal standard solution to acquire a final concentration of 6 μg/mL. 10 μL of the spiked solution was spotted on the VeriSpray^{TM} sample plate for MS analysis.

2.2.4 Spectral Data Handling

SERS (300–2000 cm^{-1}) and IR (650–4000 cm^{-1}) spectral data for all samples in the training ($n = 168$) and testing ($n = 50$) sets were organized into dataframes using the pandas python package. Each sample was labeled as true or false for all components determined via paper spray mass spectrometry (PS-MS). The wavenumber values were used as column titles, and the corresponding intensity values were organized into rows to create a python data object, hereafter referred to as a dataframe.

Before creating the concatenated SERS–IR dataset, the spectral data were scaled to a common reference point to ensure meaningful integration and to eliminate discrepancies caused by differing measurement scales and intensities. All SERS and IR data were preprocessed using global max normalization, setting the maximum intensity value across all data sets to 1.

Data fusion of the SERS and IR normalized spectra was performed by concatenating the IR spectral values to the end of the SERS data in a single matrix (x), forming a new series dataframe. Each sample in both the training ($n = 168$) and testing ($n = 50$) sets

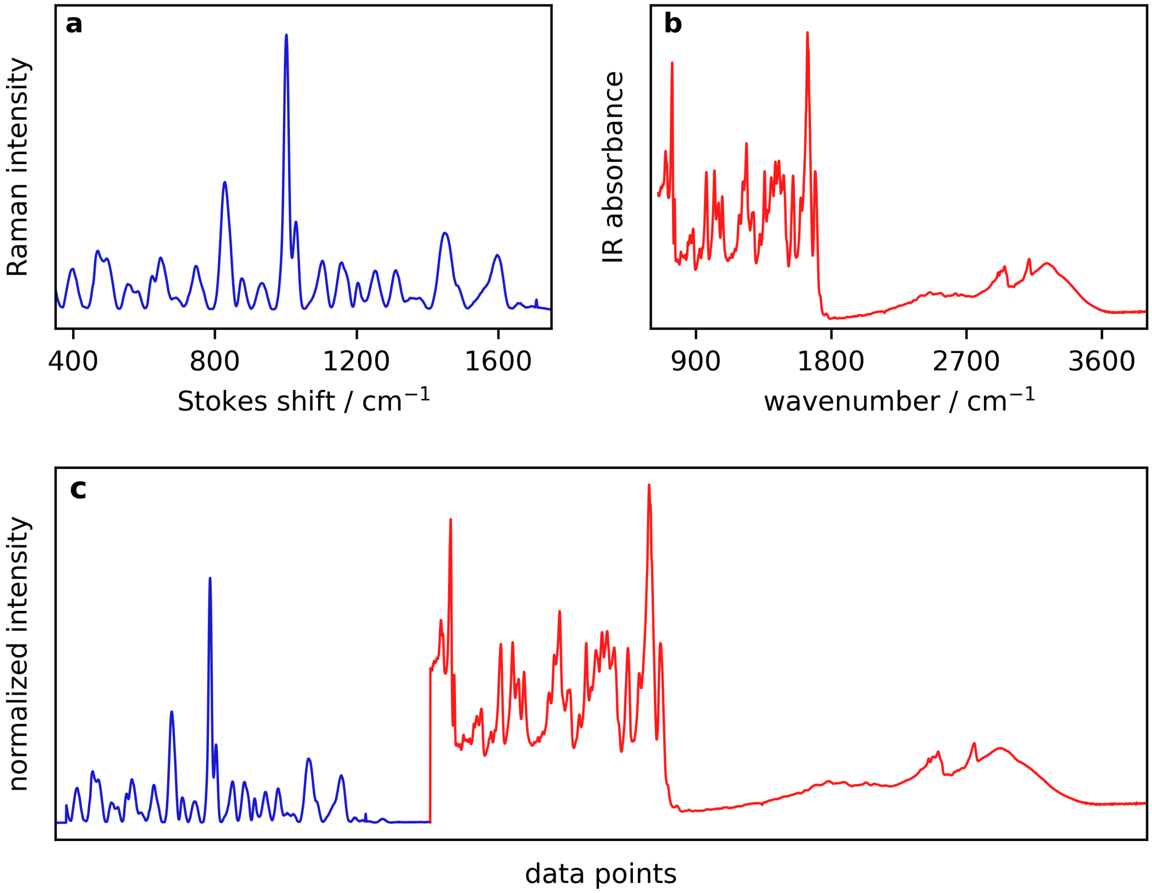


Figure 2.2: Mean spectra of samples in the training set ($n = 168$) for (a) SERS, (b) IR, and (c) concatenated SERS–IR fused data. SERS spectra in the stokes region of 300–2000 cm^{-1} is represented in blue and IR spectra in the wavenumber region of 650–4000 cm^{-1} is represented in red.

had corresponding SERS (1601 x-values), IR (4000 x-values), and concatenated SERS–IR (5601 x-values) labeled spectral data.

Figure 2.2 illustrates the mean spectra of the training set ($n = 168$) for SERS, IR, and concatenated SERS–IR data used in this study. The scikit-learn Python package was used to facilitate the outlier detection and model development on the labeled spectral data. These analyses included: principal component analysis (PCA), minimum covariance determinant (MCD) method, and Mahalanobis distances for outlier detection. Additionally, classification models were developed and optimized using random forest (RF), support vector machine (SVM), and k-nearest neighbors algorithm (KNN).

2.3 Outlier Detection Method

The spectral data for all samples in the training set ($n = 168$) were preprocessed by first subtracting the mean and scaling to unit variance using the scikit-learn package [87] in python. Principal component analysis (PCA), was used to distinguish high variance outliers that could reduce model performance. The selection of outliers was then performed using the minimum covariance determinant (MCD) method, and Mahalanobis distances were computed. Outliers were identified based on a threshold derived from a χ^2 distribution. A visualization of the outlier detection method using the first 2 principal components and robust Mahalanobis distances in SERS, IR, and concatenated SERS–IR spectral data is represented in Figure 2.3. This outlier detection method restricts the training libraries to exclusively complex multi-component samples that share commonality within the first two principal components. The outliers detected from the SERS, IR, and concatenated SERS–IR training libraries exhibited irregularities in the spectral features as compared to the other samples in PCA. These outliers corresponded to samples with high benzodiazepine concentration and unique bulk cutting agents. The data determined to be outliers were removed, and the spectra of the established training set for SERS ($n = 151$), IR ($n = 153$), and SERS–IR ($n = 154$) were used for subsequent hyperparameter tuning.

2.4 Models

2.4.1 Random Forest

Surface-enhanced Raman scattering (SERS), combined with chemometric techniques, provides powerful tools for the analysis and interpretation of chemical and biological data. Among these techniques, Random Forest (RF) stands out as a popular method for classifying compounds. RF classification is an ensemble learning algorithm that builds multiple decision trees during the training phase and outputs the majority class from the individual trees [78].

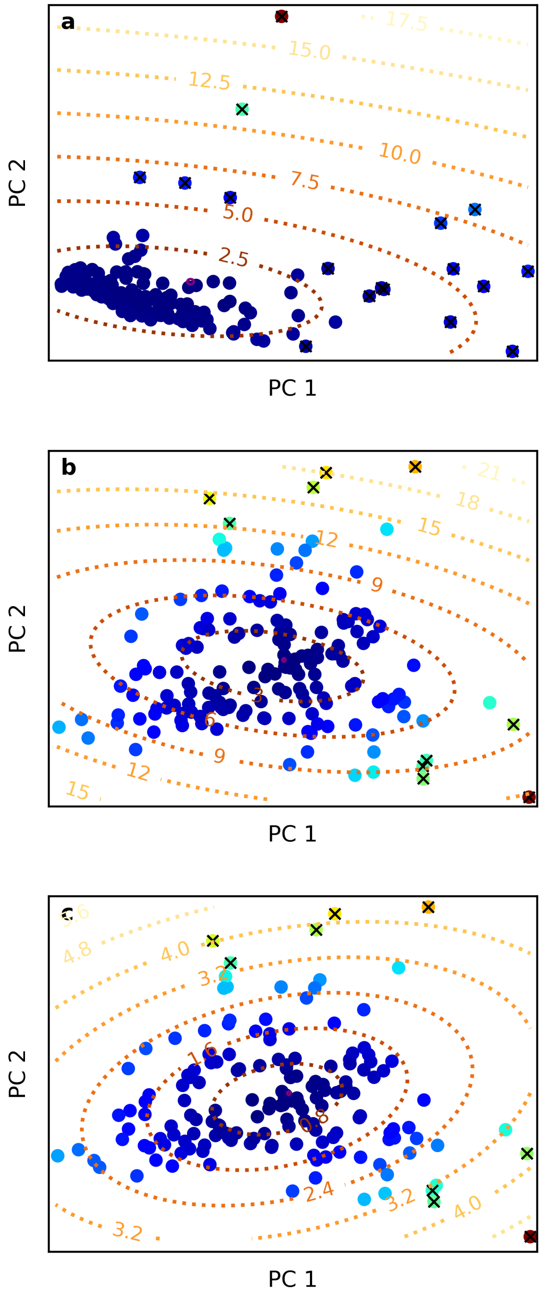


Figure 2.3: Outlier detection using PCA and robust Mahalanobis distance for (a) SERS, (b) IR, and (c) concatenated SERS–IR fused spectral data. Each subplot illustrates the scatter plot of the first two principal components, colour-coded by Mahalanobis distance. Overlaid contour plots represent the Mahalanobis distances, with identified outliers marked with x. Gradient colours of the data points range from dark blue (low variance) to red (high variance).

The RF algorithm leverages bagging and feature randomness to create an uncorrelated forest of trees whose prediction by committee is more accurate than that of any individual tree. Each tree in the forest is built from a sample drawn with replacement (i.e., a bootstrap sample) from the training set, and each node is split using the best among a subset of predictors randomly chosen at that node [78]. This approach results in a diverse set of decision trees, which collectively improve predictive performance.

RF binary classification utilizes labelling of true (y value of 1) or false (y value of 0) for a given compound in a sample and the decision trees are grown on the training set using random selection of spectral features (i.e., independent wavenumber bands) to split each node. The selected features deemed to be the most influential in RF decision making can allow for a visualization of the model which helps to minimize the black-box aspect of many classification models [71]. RF affords the ability to individualize component classification and improve predictive success by personalizing the amount of decision trees as well as modifying tree pruning parameters. Therefore, it is a powerful classification tool for handling specific component detection and quantification in complex mixtures [75] and has been demonstrated to have success in biological samples [77, 88] and simulated data [76] in combination with SERS.

SERS and RF have been implemented in various studies for drug analysis due to the trace capabilities of SERS and the feature selection of RF. However, these cases typically focus on binary or ternary mixtures and lab grade samples with known concentrations of analytes to determine limit of detection and train the model. Methods for trace detection of compounds in highly complex illicit drugs pose significant challenges using simple model development strategies. Therefore, chemometric approaches must incorporate spectral preprocessing measures along with a comprehensive analysis of RF hyperparameters.

2.4.2 Support Vector Machine

Support vector machine (SVM) is a supervised learning algorithm widely used for both classification and regression tasks [79, 80]. It has proven to be an effective classifier in numerous prediction models. For instance, it has been used for the non-invasive diagnosis of Crohn's disease utilizing PCA-SVM with SERS [89] and for detecting breast cancer by optimizing parameters to increase model performance [90].

SVM works by identifying the optimal hyperplane that best separates data points of different classes in a high-dimensional space. This hyperplane is defined by support vectors closest to the boundary. The goal of SVM is to maximize the margin between the support vectors of different classes, thereby improving the ability of the model to generalize to unseen data.

SVM is particularly effective in high-dimensional spaces and scenarios where the number of dimensions exceeds the number of samples. SVM is a suitable chemometric approach for analyzing spectral data [91], which is often high-dimensional and complex. It can handle non-linear relationships through the use of kernel functions such as radial basis function (RBF), polynomial, and sigmoid. These functions map the input features into higher-dimensional spaces where linear separation becomes possible. By tuning parameters like the regularization term (C) and the kernel parameters, SVM can be optimized to enhance performance and accuracy for complex data sets.

2.4.3 *k*-Nearest Neighbors

The *k*-nearest neighbors (KNN) algorithm is a straightforward yet powerful supervised learning method used for both classification and regression tasks [80, 81]. It operates on the principle that similar data points are likely to have similar output values. In classification, KNN assigns a class to a data point based on the majority class among its *k*-nearest neighbors in the feature space. As a non-parametric algorithm, KNN makes no assumptions about the underlying data distribution, making it highly intuitive and easy to implement [92]. The model structure is derived directly from the dataset, making KNN particularly effective for real-world data sets that do not adhere to strict theoretical assumptions [93].

KNN is particularly valuable in model development due to its simplicity and effectiveness in handling single- and multi-class classification problems. For example, it has been used to distinguish proteins in salivary Raman spectra by optimizing *k*-values [94], and for comparison to other machine learning techniques for component identification in advanced spectral preprocessing [95] and neural network strategies [91].

To optimize KNN performance, a hyperparameter grid search can be used to test various values for *k*, weighting functions (uniform or distance-based), algorithms (brute-force, *k*-d trees, or ball trees), and distance metrics (Euclidean, Manhattan, or Minkowski). Additionally, feature scaling and spectral preprocessing strategies can further enhance the model. By carefully tuning these hyperparameters and preprocessing techniques, KNN can become a highly effective tool for developing robust predictive models.

2.5 Model Optimization

Optimization of the models must take into account the way spectral data is introduced to the model in order to ensure compatibility of preproccesed spectra with the model parameters. Preprocessing of spectroscopic data included the normalization and order of derivatives of the training set to evaluate how they might improve or reduce model performance on the testing set. Proper consideration to the handling of training data is necessary before inputting it to ensure a maximized performance of the selected hyperparameters [96]. Therefore, preprocessing techniques appropriate for SERS and IR spectroscopy were tested alongside all model parameters. The derivative and normalization techniques implemented were area normalization, min-max normalization, standard normal variate (SNV), first order derivative, and second order derivative. Both derivatives were implemented with Savitzky-Golay smoothing (window size 5, polynomial order 2).

A cross-validated grid search is a model evaluation technique used to analyze several hyperparameters, allowing for a robust computational examination to determine the best performing combination [97–100]. A grid search allows for the examination of thousands of parameter combinations, and cross-validation is widely recognized as a reliable method for determining the optimal hyperparameters for any type of classifier [96]. A *k*-fold cross validation grid search was selected as the method for hyperparameter tuning of all models developed in this work due to its increased success at enhancing model generalization [97, 100]. The procedure involves *k* rounds of training and validation; in each round, a 1/*k* subset (1 fold) of the training library is used as the validation set while the model trains on the other *k*-1 folds of data. An accuracy and precision score is determined for the prediction results on the validation fold, and this process is repeated *k* times until each fold has served as a validation set. Once completed the model can be retrained on the entire training library in order to make predictions on the unseen data in the testing set ($n = 50$). In this work, we employed a 5-fold cross-validation approach to determine the optimal parameters needed to construct our models. The averaged accuracy and precision for all 5 folds of training is determined for every parameter combination and the final model parameters are selected by maximum F1 score. All subsequent model development described in Chapters 3 and 4 were conducted using this method. All models were tested with their respected model-specific hyperparameters as well as the spectral preprocessing techniques. Table 2.2 illustrates all hyperparameters used for random forest (RF), support vector machine (SVM), and *k*-nearest neighbors (KNN) model development, with descriptions for the role of the parameter and all tested hyperparameters listed.

Model Optimization

Parameter	Description	Grid Search Values
Random Forest		
n_estimators	The number of trees in RF	5, 10, 15, 20, 30, 40, 50, 100, 150, 200, 300, 400, 500, 800, 1000, 2000
max_depth	The maximum depth of a single tree	None, 1, 2, 3, 4, 5, 10, 20, 30, 50
min_samples_leaf	The minimum number of samples required to be at a leaf node	1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50
SVM		
C	Regularization parameter. The strength of the regularization is inversely proportional to C	1, 10, 100
kernel	Specifies the kernel type to be used in the algorithm	'linear', 'rbf', 'poly', 'sigmoid'
gamma	Kernel coefficient for 'rbf', 'poly', and 'sigmoid' kernels. It defines how far the influence of a single training example reaches	'auto’, ‘scale'
degree	Degree of the polynomial kernel function ('poly')	1, 2, 3, 4, 5
coef0	Independent term in the kernel function (Only significant for ‘poly’ and ‘sigmoid’ kernels)	0.0, 0.1, 0.5, 1.0
class_weight	Weights associated with classes	None, ‘balanced’
KNN		
n_neighbors	The number of neighbors to use for k-neighbors queries	1, 2, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 25, 30
weights	Weight function used in prediction	'uniform', 'distance'
algorithm	Algorithm used to compute the nearest neighbors	'auto', 'ball_tree', 'kd_tree', 'brute'
metric	How the algorithm calculates the distance between data points	'euclidean', 'manhattan', 'minkowski'

Table 2.2: Investigation of hyperparameters using a 5-fold cross-validation grid search for development of random forest (RF), support vector machine (SVM), and *k*-nearest neighbors (KNN) classification models. Showing hyperparameter descriptions and the parameter combinations tested for all models.

Chapter 3

Trace Detection of Adulterants in Illicit Opioid Samples Using Surface-Enhanced Raman Scattering and Random Forest Classification[1]

3.1 Introduction

Identifying adulterants in the unregulated market poses challenges for drug checking services which are often limited for differentiating trace components using current methods, namely test strips or vibrational spectroscopic instruments. Given the context of complex and co-competing compounds in illicit drug matrices, highly sensitive techniques are necessary to expand the efficacy of drug checking. The use of automated technologies and machine learning algorithms capable of trace detection can ease the burden of manual interpretation and facilitate more widespread implementation.

In this study, we employ random forest classifiers trained on FTIR and SERS spectral data from illicit drug samples acquired as a part of a community drug checking service. We focus on two common adulterants that are frequently present in opioid samples and typically in trace amounts, below the limit of detection of FTIR instruments. The

---

[1] Reproduced in part from Martens, R. R.; Gozdzialski, L.; Newman, E.; Gill, C.; Wallace, B.; Hore, D. K. ”Trace Detection of Adulterants in Illicit Opioid Samples Using Surface-Enhanced Raman Scattering and Random Forest Classification.“ *Anal. Chem.* **96**, 12277–12285 (2024). Copyright 2024 American Chemical Society.

first is bromazolam, a derivative of the benzodiazepine class known as triazolobenzodiazepine [101]. This is an important illustration of how SERS can be used to differentiate compounds with varying potency and pharmacological effects within the benzodiazepine class. The second example is the detection of xylazine, a more challenging compound to identify due to strong spectral interferences. We have found the overall accuracy and specificity of the SERS method in the classification of adulterants to be superior to what can be achieved using FTIR data, when both types of spectra are analyzed using optimized RF models. This illustrates the potential of a simple and sensitive SERS technique to be successfully implemented in a point-of-care community drug checking application.

3.2 Model Development

3.2.1 Hyperparameter Tuning for a Random Forest Binary Classifier

Binary classification for the prediction of bromazolam and xylazine in the SERS training sets was performed using a random forest classifier model (scikit-learn). The default criterion of Gini impurity was used. Optimization of the preprocessing methods (derivatives and normalization) and hyperparameters (number of decision trees, maximum depth, and minimum samples per leaf) for the RF classifier was conducted using a 5-fold cross-validation grid search. This process tested 11,880 parameter combinations for bromazolam and xylazine labeled independently. Optimal hyperparameter values were selected based on the highest F1 scores obtained through grid search. This tuning method was also performed using the IR training set to establish IR-based RF models with the same learning metrics. Results for the optimized model parameters for SERS and IR spectral data are shown for bromazolam in Table 3.1a and xylazine in Table 3.1b.

3.2.2 Targeted Secondary Model Development for SERS Data

Optimization of the prediction results from the bromazolam-SERS and xylazine-SERS models for the test set ($n = 50$) was performed in order to reduce the number of false

a. Bromazolam

Parameter Name	Description	Grid Search Values	SERS		IR	
			Best Parameter	F1 Score	Best Parameter	F1 Score
n_estimators	The number of trees in RF	5, 10, 15, 20, 30, 40, 50, 100, 150, 200, 300, 400, 500, 800, 1000	800	0.891	40	0.737
max_depth	The maximum depth of a single tree	None, 1, 2, 3, 4, 5	None		5	
min_samples_leaf	The minimum number of samples required to be at a leaf node	1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50	3		4	
deriv	The order of derivatives of the spectral data	0, 1, 2	1		2	
norm	The normalization process of the spectral data	‘min-max', 'snv', 'none', 'area'	'none'		'none'	

b. Xylazine

Parameter Name	Description	Grid Search Values	SERS		IR	
			Best Parameter	F1 Score	Best Parameter	F1 Score
n_estimators	The number of trees in RF	5, 10, 15, 20, 30, 40, 50, 100, 150, 200, 300, 400, 500, 800, 1000	40	0.805	5	0.499
max_depth	The maximum depth of a single tree	None, 1, 2, 3, 4, 5	3		None	
min_samples_leaf	The minimum number of samples required to be at a leaf node	1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50	4		5	
deriv	The order of derivatives of the spectral data	0, 1, 2	0		2	
norm	The normalization process of the spectral data	‘min-max', 'snv', 'none', 'area'	'snv'		'min-max'	

Table 3.1: Investigation of RF hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search for (a) bromazolam and (b) xylazine, showing the combinations tested and the optimal set of parameters selected based on F1 score.

positives.

3.2.2.1 Low Confidence Range

The predicted probability scores of all false positives in the test set ($n = 50$) from the bromazolam-SERS and xylazine-SERS models were used to establish their respective low confidence range (LCR). With the highest false positive prediction score marking the maximum end of the range and the determined threshold establishing the minimum. An LCR of 0.5–0.6 was determined for bromazolam-SERS and 0.25–0.45 for xylazine-SERS. The samples with prediction scores within the LCR developed for bromazolam ($n = 11$) and xylazine ($n = 10$) underwent another round of predictions using the secondary RF classification models.

3.2.2.2 Model Tuning

Model visualization of the top twenty features contributing to the classification basis of the SERS models established the 650–750 cm^{-1} region for bromazolam and 1050–1150 cm^{-1} region for xylazine to be the most significant for characteristic information. The training library data ($n = 168$) was restricted to only contain intensity values in the described spectral region for bromazolam and xylazine to establish their respective training libraries. An identical outlier detection and removal method described in the paper was applied to the training libraries creating with the restricted bromazolam-SERS library ($n = 140$) and the restricted xylazine-SERS library ($n = 152$). An additional hyperparameter tuning method using a 5-fold cross validation grid search was performed to determine the optimal parameters of the targeted secondary models for bromazolam-SERS and xylazine-SERS illustrated in Table 3.2.

3.2.2.3 Model Application

The LCR samples were run with their respective secondary RF classifier model and the optimal threshold determined by ROC analysis can be seen in Figure 3.1. Prediction results

Parameter Name	Description	Grid Search Values	Bromazolam-SERS Model 2		Xylazine-SERS Model 2	
			Best Parameter	F1 Score	Best Parameter	F1 Score
n_estimators	The number of trees in RF	5, 10, 15, 20, 30, 40, 50, 100, 150, 200, 300, 400, 500, 800, 1000	30	0.872	300	0.727
max_depth	The maximum depth of a single tree	None, 1, 2, 3, 4, 5	5		None	
min_samples_leaf	The minimum number of samples required to be at a leaf node	1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50	2		20	
deriv	The order of derivatives of the spectral data	0, 1, 2	1		0	
norm	The normalization process of the spectral data	‘min-max', 'snv', 'none', 'area'	'none'		'min-max'	

Table 3.2: Investigation of RF hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search on the restricted SERS spectral region training libraries for bromazolam (650–750 cm^{-1}) and xylazine (1050–1150 cm^{-1}). This highlights the combinations tested and the optimal set of parameters selected for the targeted secondary models based of F1 score.

of the secondary models on the LCR test sets are illustrated in Figure 3.2. The secondary bromazolam-SERS model correctly assigned 1 true negative sample whereas the secondary xylazine-SERS model correctly assigned 3 true negative samples and incorrectly assigned 1 false negative sample. The new prediction results for the selected samples are reflected in the final model prediction results of the study.

3.3 Bromazolam Detection

3.3.1 Model Optimization

The optimized RF binary classifiers for predicting bromazolam using SERS and IR spectral data demonstrated robust performance, yielding F1 scores of 0.891 and 0.737, respectively, during cross-validation on the training set (Table 3.1a). The higher F1 score of 0.891 obtained by the SERS model suggests superior precision and recall balance, indicating its effectiveness in accurately identifying positive instances within the SERS training set. Notably, the ideal parameter for optimizing SERS spectra emphasized the first-order derivative, which is particularly beneficial for improving the resolution of overlapping peaks, a common occurrence in samples with structurally similar components. This characteristic could elucidate the model's effectiveness at differentiating bromazolam features from other benzodiazepines in complex mixtures, enhancing model sensitivity to subtle spectral variations. Additionally, the large number of decision trees (800) selected for the SERS model could contribute to its high F1 score, as this helps to reduce overfitting and captures a more comprehensive representation of data patterns. The IR model, with an F1 score of 0.737, also exhibits relatively high initial predictive performance. These results underscore the capacity of the model to discern bromazolam patterns in the respective spectral data sets, with the SERS model showing a particular advantage in achieving a harmonized trade-off between precision and recall. The initial success observed in the predictive performance of the SERS model suggests that the features in SERS spectra allow for more accurate detection of bromazolam.

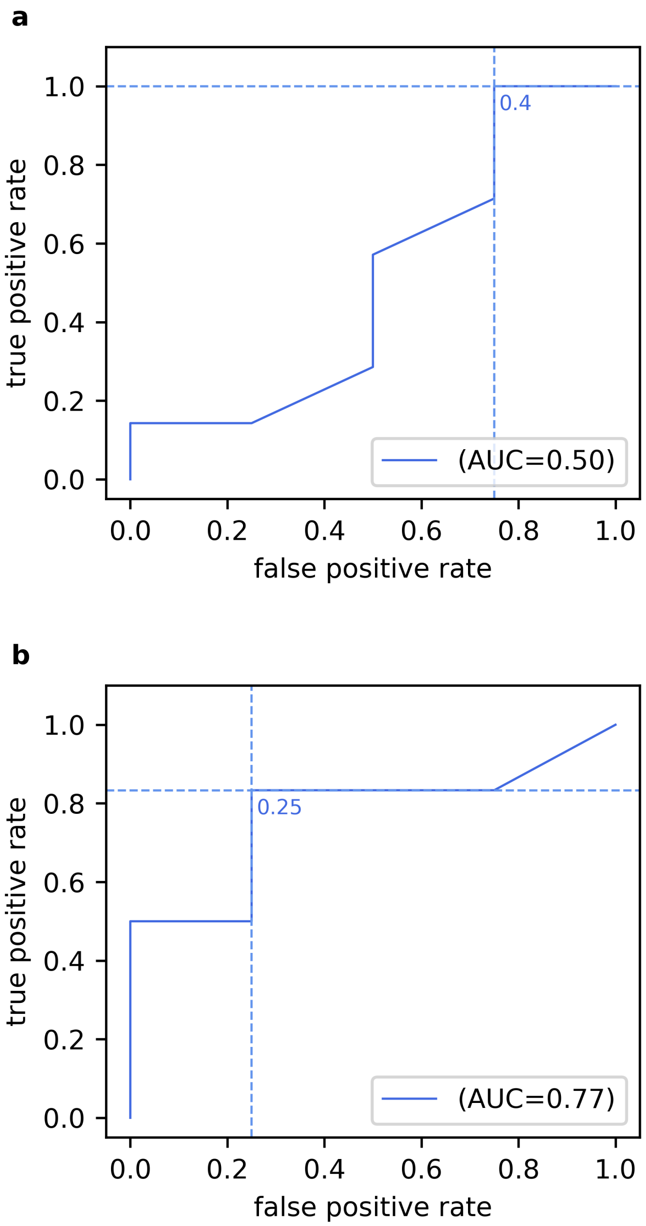


Figure 3.1: Receiver operator characteristic (ROC) curve of secondary SERS RF binary classification models for (a) bromazolam on LCR subset ($n = 11$) and (b) xylazine on LCR subset ($n = 10$). The curves illustrate the true positive rate and false positive rate across various classification thresholds for adulterant detection. Optimal threshold selection for balancing sensitivity and specificity and area under the curve (AUC) is highlighted.

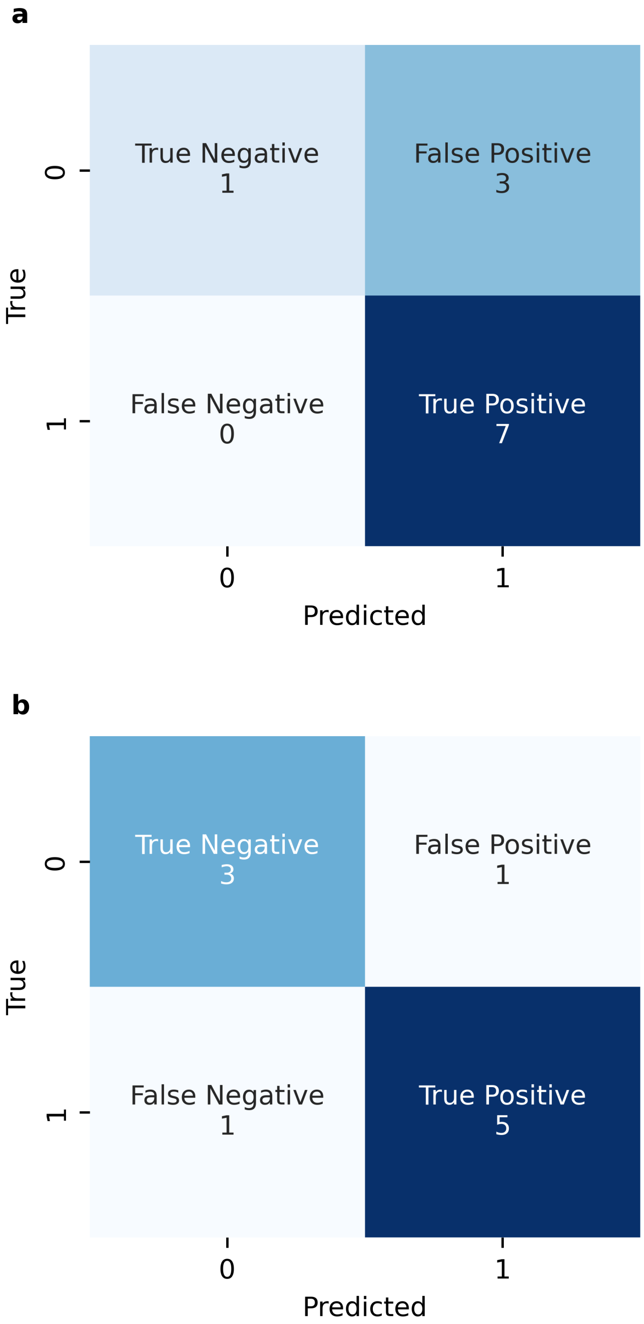


Figure 3.2: Confusion matrices for SERS secondary RF model prediction results of LCR samples for (a) bromazolam ($n = 11$) and (b) xylazine ($n = 10$).

3.3.2 SERS Model Visualization

The mean spectra of bromazolam-positive and bromazolam-negative samples in the training library were used to understand the classification basis of the SERS model, as depicted in Figure 3.3a. The blue bands overlaid on the mean spectra represent the top twenty features extracted by the model. The wavenumber bands selected constitute the features with the highest importance degree, as calculated by the Gini index, that contribute to the identification of bromazolam by the RF model. Among the identified feature importance variables (FIVs), 14 were notably concentrated within the 650–750 cm^{-1} region. These were evenly distributed, split between the initial rise (675–681 cm^{-1}) and subsequent decline (691–696 cm^{-1}) of intensity around the distinctive band observed in the mean bromazolam-positive spectrum at 690 cm^{-1}. These features seem to correlate with the rates of change both preceding and following the band, likely a reflection of the model's selection to use the first-order derivative as the optimal spectral preprocessing parameter. This band appears to correspond to a non-obstructed spectral region and aligns with a characteristic peak visible in the SERS spectrum of a single component bromazolam street sample shown in Figure 3.4, largely accredited to C–Br stretching [102, 103], unique to bromazolam that distinguishes it from other triazolobenzodiazepines. However, it is important to note that the peak highlighted by the FIVs is intended solely for illustrating the classification framework of the SERS model and is not meant to provide information that could be used for manual interpretation of the spectra. Characterization of the band in the concentrated FIV region is further evaluated by analyzing the mean spectra of various combinations that exist within the opioid sample training set illustrated in Figure 3.3b. In order to ascertain whether bromazolam is the primary contributor to the band, opioid samples adulterated exclusively with bromazolam were compared against samples containing a mixture of adulterants, samples without bromazolam adulteration, and unadulterated opioid samples. Both mean spectra for samples not containing bromazolam did not possess the characteristic band at 690 cm^{-1}, and while less pronounced, evidence of the band was still visually apparent for the mean spectrum of samples containing bromazolam and other adulterants.

Following this, a second model was developed where the spectral training data was restricted to the 650–750 cm^{-1} region, thereby focusing on the area with the most significant contribution to bromazolam characterization. This targeted approach aims to streamline the development of a secondary RF model (Table 3.2) for the detection of bromazolam in low confidence samples in the test set, as detailed in Section 3.2.2.

3.3.3 Model Application and Performance

A receiver operating characteristic (ROC) curve was used to assess classifier accuracy and determine optimal thresholds for classification using the predicted probability scores of the SERS and IR models on the test set. These ROC curves are plotted using varying threshold values as illustrated in Figure 3.5.

The area under the curve (AUC) was used to evaluate the performance of the binary classification models across the chosen range of thresholds. AUC values of 0.93 and 0.78 were determined for the SERS and IR models, respectively, indicating good discriminatory power. However, the high AUC value of 0.93 demonstrates the increased capability of SERS model in distinguishing true bromazolam from false bromazolam, irrespective of the classification threshold.

Optimal thresholds for the SERS (0.5) and IR (0.4) models were identified by balancing sensitivity and specificity using Youden's J statistic. Test set prediction results for bromazolam were generated using the corresponding threshold values. A low confidence range (LCR) between 0.5 and 0.6 was established for the SERS model based on the minimum and maximum probability scores derived from false positive samples in the test set. Subsequently, all samples with scores in the LCR ($n = 11$) were subject to additional testing using the targeted secondary bromazolam SERS model (Figure 3.1a and Figure 3.2a) trained on restricted spectral regions as detailed in Section 3.2.2. The LCR

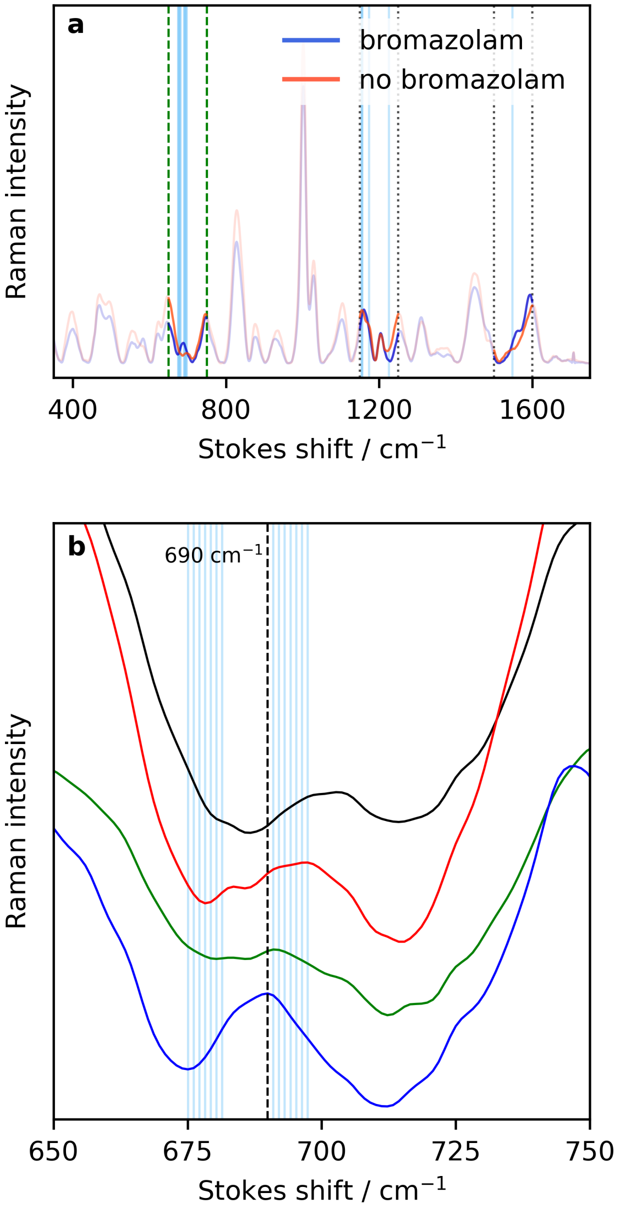


Figure 3.3: (a) Mean spectral data for bromazolam-positive and bromazolam-negative samples in the training set overlaid with vertical blue lines representing the top 20 features extracted by the model. Spectral regions with feature presence are highlighted. (b) Detailed view of the 650–750 cm^{-1} region with stacked mean spectra for different opioid drug combinations: opioid-only (black), adulterants other than bromazolam (red), bromazolam and additional adulterants (green), and bromazolam-only (blue).

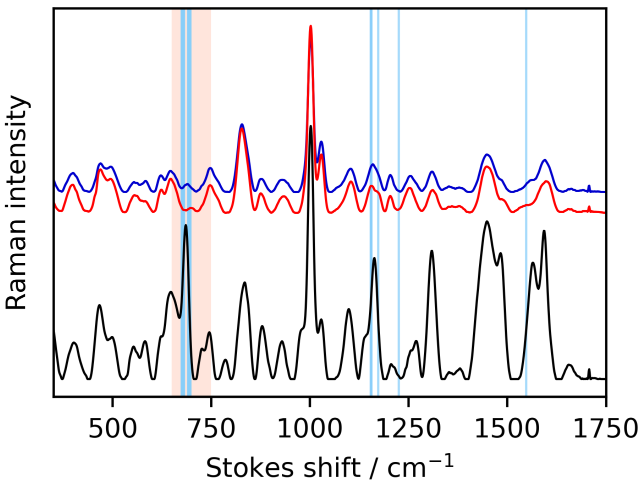


Figure 3.4: Mean SERS spectral data for bromazolam-positive (blue) and bromazolam-negative (red) samples in the training set compared to a single component bromazolam street sample (black). Vertical lines represent the top 20 features extracted by the bromazolam SERS model. The characteristic spectral region of 650–750 cm^{-1} is highlighted.

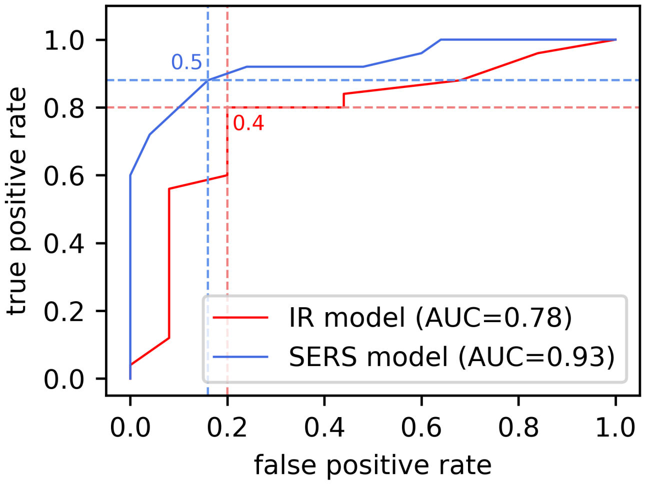


Figure 3.5: A comparison between the bromazolam SERS model and bromazolam IR model prediction results on the test set ($n = 50$). The ROC curves illustrate the true positive rate and false positive rate across various classification thresholds for bromazolam detection. Optimal threshold selection for balancing sensitivity and specificity and area under the curve (AUC) is highlighted for both models.

prediction results were then reintegrated with the test set for model evaluation.

The SERS and IR models prediction results of the test set used for model evaluation are illustrated in Table 3.3.

	IR Model on Test Set ($n = 50$)	SERS Model on Test Set ($n = 50$)
True positive (n)	20	22
True negative (n)	20	22
False positive (n)	5	3
False negative (n)	5	3

Table 3.3: Bromazolam prediction results for samples in the test set ($n = 50$) for SERS and IR models.

Given the complexity of opioid samples and the impact of misidentifying powerful sedatives in the sample, the ability to detect bromazolam with high sensitivity is of greater importance than specificity. While both remain integral to assessing the performance of the model, increasing the number of true positives and lowering the number of false negatives is a significant factor in the context of this work. Both the SERS and IR models demonstrated high predictive success in the positive and negative class, showing balanced sensitivity and specificity (Table 3.4a). The IR model correctly identified 12 out of the 17 bromazolam-positive samples below its widely accepted threshold of 5% w/w [41–43]. The expanded limit of detection and the relatively high accuracy of the IR model (80%) is likely a consequence of the distinct peaks of bromazolam in the 750–850 cm^{-1} region as shown in Figure 3.6. In comparison, the SERS model correctly identified 15 out of 17 samples with bromazolam concentrations of <5% w/w, with the lowest detected concentration being 0.32% w/w, the highest at 36% w/w. This level of performance illustrates the compatibility of the RF-based approach with both SERS and IR spectral data in trace detection of bromazolam.

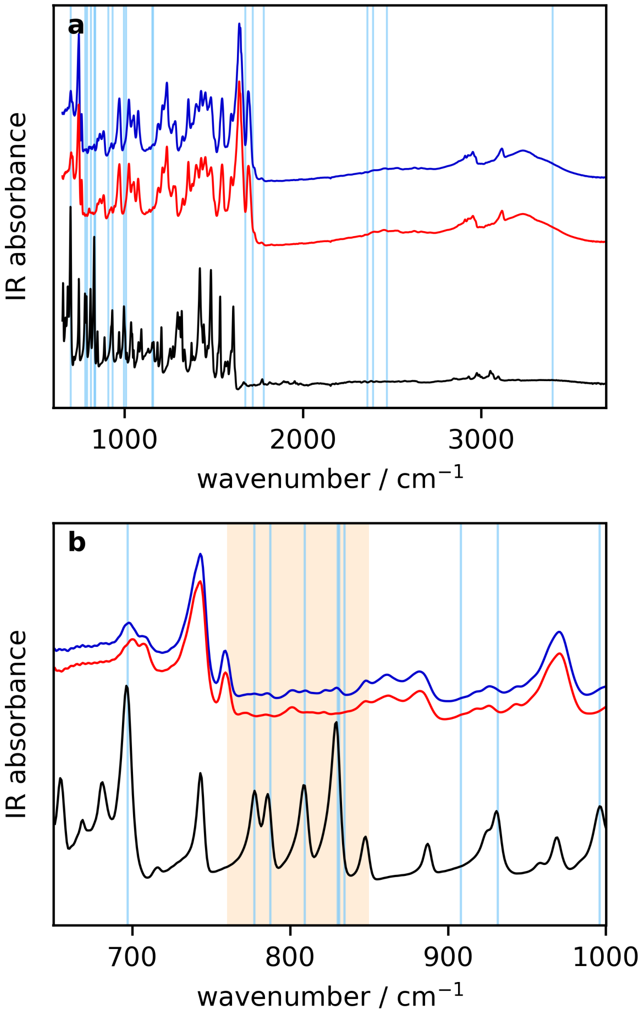


Figure 3.6: (a) Mean IR spectral data for bromazolam-positive (blue) and bromazolam-negative (red) samples in the training set compared to a single component bromazolam street sample (black). Vertical lines represent the top 20 features extracted by the bromazolam IR model. (b) An enhanced view of the stacked spectra with the spectral region of 750–850 cm^{-1} highlighted to illustrate the characteristic peaks of bromazolam.

a. Bromazolam

Model	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
IR	80	80	80	80	80	80
SERS	88	88	88	88	88	88

b. Xylazine

Model	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
IR	68	62	96	40	62	91
SERS	94	96	92	96	96	92

Table 3.4: Summary of accuracy, precision, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for opioid samples ($n = 50$) tested with RF models. Model prediction results for IR and SERS data are shown for (a) bromazolam detection and (b) xylazine detection.

3.4 Xylazine Detection

3.4.1 Model Optimization

The optimized RF binary classifiers for predicting xylazine using the SERS and IR spectral data demonstrated varied performance, yielding F1 scores of 0.805 and 0.499, respectively, during cross-validation (Table 3.1b). The resulting SERS model F1 score of 0.805 indicates a high precision and recall within the training set and notably prioritized normalization over derivative-based parameters for optimizing spectra. The standard normal variate (SNV) normalization selected for the model is a useful technique in reducing SERS intensity variations and in standardizing the scale of spectral features. The IR model demonstrated weak initial predictive performance, as evidenced by an F1 score of 0.499 and a particularly low mean sensitivity (recall) of 0.439, averaged across the five validation folds. The selection of 5 decision trees suggests the IR model has low complexity for the intricacies of the sample matrices it is analyzing and may be making simplistic assumptions of patterns regarding xylazine to improve its performance. The selection of min-max normalization is indicative of a high degree of spectral overlap, with more dominating components in the mixture presenting a challenge for the classification of xylazine using the IR method.

3.4.2 SERS Model Visualization

Mean xylazine-positive and xylazine-negative SERS spectra of samples in the training set and the top twenty features extracted by the model are illustrated in Figure 3.7a. The distinct peak observed at 1490 cm^{-1} in the mean xylazine-positive spectrum shown in Figure 3.8, accompanied by a nearby feature importance band at 1495 cm^{-1} (Figure 3.7a), was also detected in benzodiazepine-only mixtures. This spectral region is closely associated with aromatic ring C–C stretching vibrations and N–C vibrations, which are common in most benzodiazepines [104–107]. Specifically, etizolam exhibits a characteristic peak at 1491 cm^{-1} [59], and other triazolobenzodiazepines show significant band presence in the 1400–1500 cm^{-1} range [104], which are frequent adulterants found in the SERS training group. Therefore, the shoulder band at 1490 cm^{-1}, experiencing significant overlap within the matrix, was determined to lack specificity for xylazine characterization and was consequently disregarded in the development of the secondary model. However, Many of the FIVs accumulate in the 1050–1150 cm^{-1} region and 14 are particularly clustered around the 1075–1098 cm^{-1} band edge. Although this region also contains contributions from additional components in the sample matrix, there is evidence of broadening due to the presence of xylazine. The FIVs and increased intensity firmly align with the strong xylazine peak visible in the single component SERS spectrum seen in Figure 3.9. The mean spectra of different combinations of opioid samples in the training set, as shown in Figure 3.7b, further illustrate the intensity variations of the peak.

Following this, the spectral training data was restricted exclusively to the xylazine characteristic SERS region of 1050–1150 cm^{-1} and used for the development of a secondary RF model (Table 3.2) targeted at low confidence samples in the test set, as detailed in the Section 3.2.2.

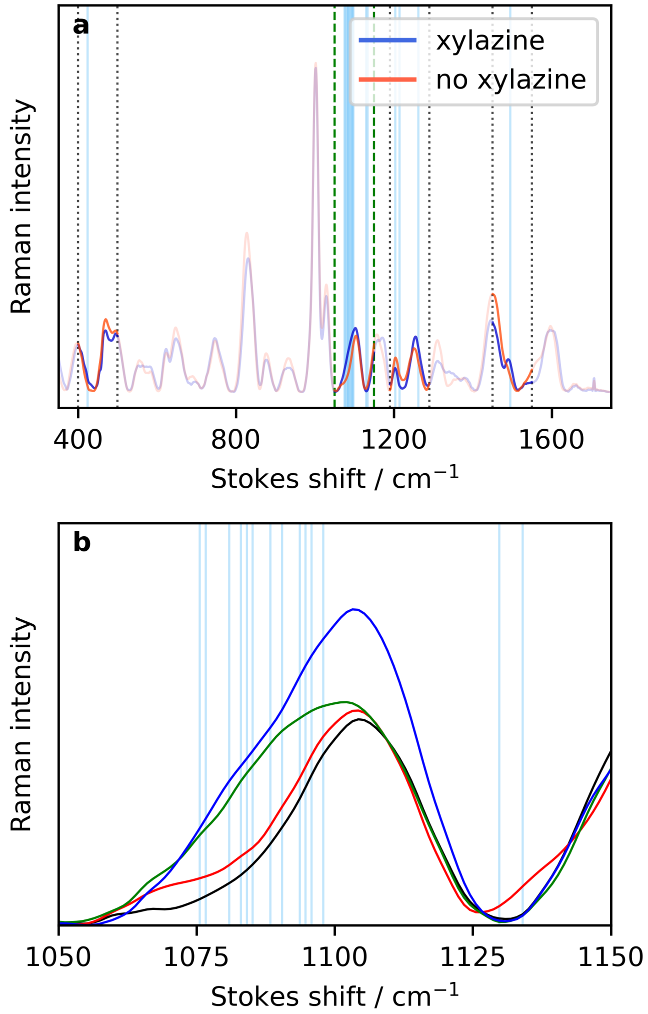


Figure 3.7: (a) Mean spectral data for xylazine-positive and xylazine-negative samples in the training set overlaid with vertical blue lines representing the top 20 features extracted by the model. (b) Detailed view of the 1050–1150 cm^{-1} region with overlaid mean spectra for different opioid drug combinations: opioid-only (black), adulterants other than xylazine (red), xylazine and additional adulterants (green), and xylazine-only (blue).

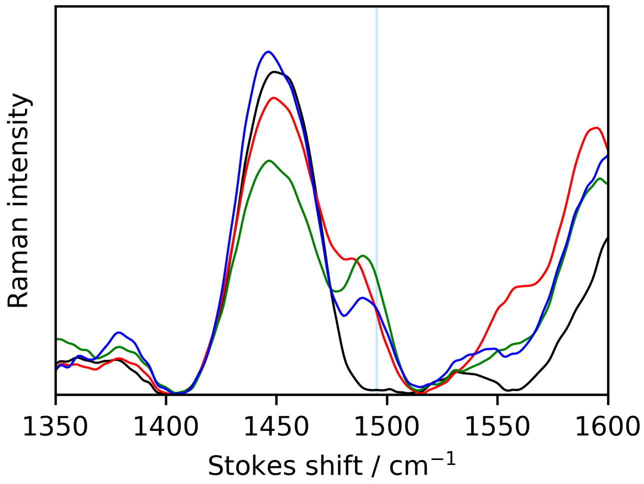


Figure 3.8: Detailed view of the 1350–1600 cm^{-1} region with overlaid mean spectra for different opioid drug combinations: opioid-only (black), adulterants other than xylazine (red), xylazine and additional adulterants (green), and xylazine-only (blue). Vertical blue line representing the 1495 cm^{-1} feature extracted in the region by the model.

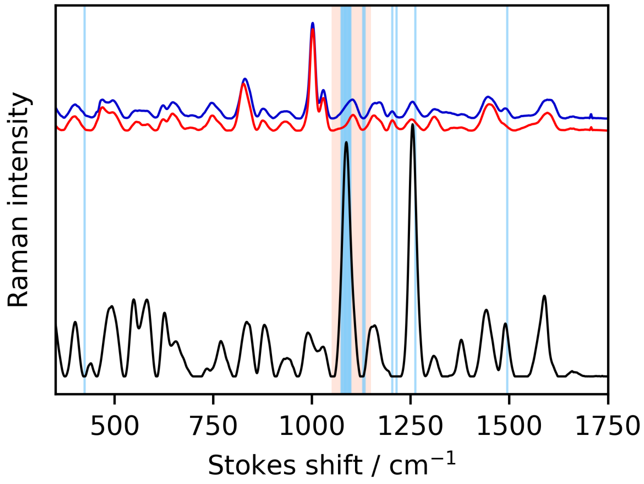


Figure 3.9: Mean SERS spectral data for xylazine-positive (blue) and xylazine-negative (red) samples in the training set compared to a single component xylazine street sample (black). Vertical lines represent the top 20 features extracted by the xylazine SERS model. The characteristic spectral region of 1050–1150 cm^{-1} is highlighted.

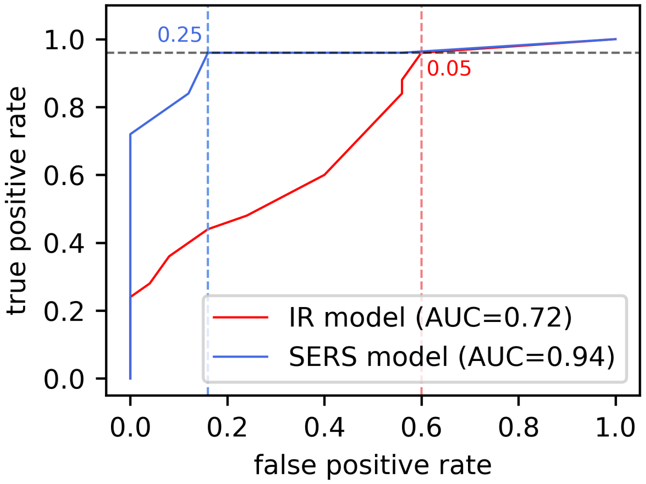


Figure 3.10: A comparison between the xylazine SERS model and xylazine IR model prediction results of xylazine on the test set ($n = 50$). The ROC curves illustrate the true positive rate and false positive rate across various classification thresholds for xylazine detection. Optimal threshold selection for balancing sensitivity and specificity and area under the curve (AUC) is highlighted for both models.

3.4.3 Model Application and Performance

ROC analysis was performed on the SERS and IR models for the test set as shown in Figure 3.10.

AUC values of 0.94 and 0.72 were calculated for the SERS and IR models, respectively. The SERS model reflects high discriminatory power across the range of thresholds, illustrating the utility of SERS for the accurate detection of xylazine with little influence from competing spectral components in the sample matrix. The IR model demonstrated poor performance across the range of predicted probabilities, with prediction results approaching that of a baseline classifier.

Optimal predicted probability score thresholds for the SERS (0.25) and IR (0.05) models were calculated and applied to the test set. The LCR for the SERS model was established to be between 0.25 and 0.45 and all samples with scores in the LCR ($n = 10$) underwent additional testing using the targeted secondary xylazine SERS model (Figure 3.1b and Figure 3.2b) detailed in Section 3.2.2. The LCR prediction results were then reintegrated with the test set for model evaluation.

Xylazine prediction results are shown in Table 3.5 and were used to evaluate the performance of the SERS and IR models. The SERS model correctly identified 23 out of the 25 xylazine-positive samples and 24 out of the 25 xylazine-negative samples. Two positive trace concentration samples (both below 1% w/w) were misidentified, yet the model successfully distinguished all other samples in the test set containing xylazine (0.15–15% w/w), demonstrating a robust prediction range.

	IR Model on Test Set ($n = 50$)	SERS Model on Test Set ($n = 50$)
True positive (n)	24	23
True negative (n)	10	24
False positive (n)	15	1
False negative (n)	1	2

Table 3.5: Xylazine prediction results of the test set ($n = 50$) for the SERS and IR models.

The results of the SERS and IR models shown in Table 3.4b coincide with the performance predictions of their respective F1 scores (0.801 and 0.499) and AUC values (0.94 and 0.72). The SERS model exhibited the ability to distinguish true negatives from true positives while minimizing false negatives and false positives (94% accuracy and 96% precision). By comparison, the IR model correctly identified 24 out of 25 true positive samples and incorrectly claimed 15 false positives, indicating superior sensitivity but lacking discriminatory power. This tendency to overpredict resulted in lower accuracy (68%) and precision (62%) for the IR model (Table 3.4b). Over prediction of the positive class could be attributed to underfitting during model training, further aggravated by a low prediction threshold of 0.05 that hampers performance for negative instances. The suboptimal performance of the IR model may be indicative of significant spectral obstruction, particularly in the 1600–1650 cm^{-1} region where xylazine's characteristic features [108] experience strong interference from the vibrational bands of opioids and common cutting agents [109]. Further evidence appears in the features selected for the xylazine IR model that do not correspond to the strong characteristic peaks in the single component xylazine reference spectrum, and the lack of distinguishable peaks in the xylazine-positive mean spectrum (Figure 3.11). This points to the limitations of IR for detecting low concentrations of xylazine in complex opioid mixtures, both in the case of manual interpretation and model development.

3.5 Consequence for Drug Checking

The implementation of a simple random forest classification method provides a valuable tool for standardizing spectral analysis in point-of-care services. However, despite the improved LoD of bromazolam afforded by the IR model, the poor xylazine performance underscores the limitations of techniques based on IR absorption in detecting low concentration adulterants. The detection of both bromazolam and xylazine was highly improved with the use of RF classification coupled with SERS. The complete breakdown of sample composition and model results on the validation set is illustrated in Table 3.6. The accurate differentiation of two common adulterants in concentrations less than 1% w/w using the SERS based models illustrates the superior sensitivity that can be achieved without sacrificing the benefits of affordable and accessible technology desired by drug checkers. The application of a straightforward and uniform SERS sample preparation method enables facile integration into drug checking training protocols. The success of these models in identifying bromazolam and xylazine using SERS data may be attributed to the fact that the training set was established using drugs collected from a community harm reduction service. Street samples of illicit opioids have inherent complexities that are often a significant challenge to accurately replicate for model development using laboratory standards or simulated data. This work showcases the predictive enhancements of binary classifiers particularly when the learning metrics of models reflect what they

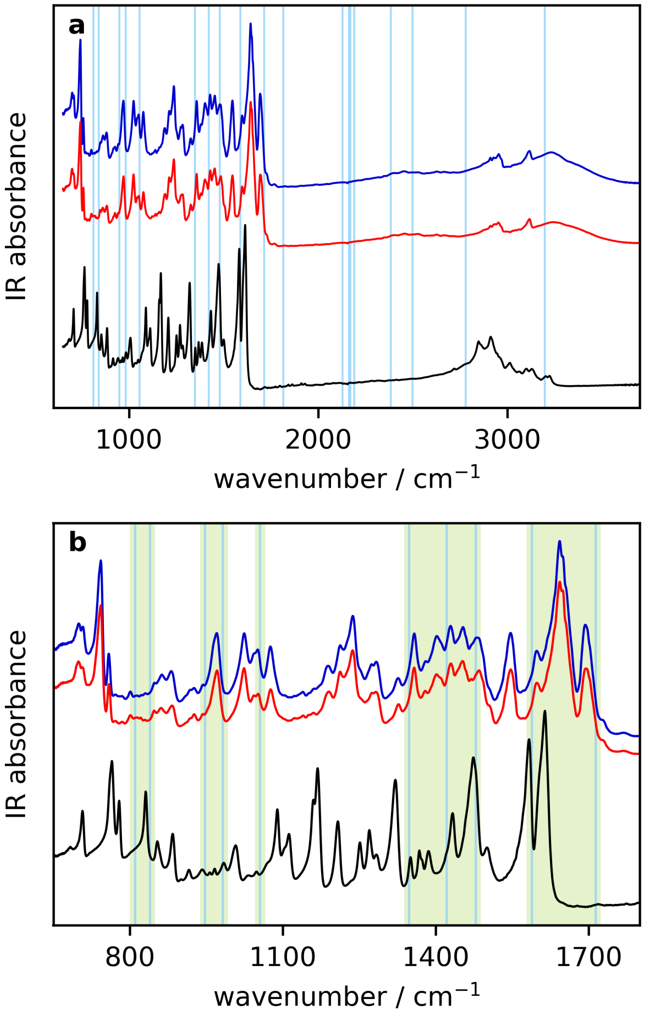


Figure 3.11: (a) Mean IR spectral data for xylazine-positive (blue) and xylazine-negative (red) samples in the training set compared to a single component xylazine street sample (black). Vertical lines represent the top 20 features extracted by the xylazine IR model. (b) An enhanced view of the stacked spectra in the 650–1800 cm^{-1} spectral region with all features extracted by the model highlighted in green.

AERL	Sample Composition	Bromazolam			Xylazine		
		Concentration (% w/w)	Model Results		Concentration (% w/w)	Model Results	
			SERS	IR		SERS	IR
5184	Fentanyl (6.95%) + Etizolam (1.07%) + Flubromazepam (0.73%) + Xylazine (0.5%) + Fluorofentanyl (0.46%) + Caffeine (nan%)	0	TN	TN	0.5	TP	TP
5404	Fentanyl (4.36%) + Bromazolam (0.99%) + Xylazine (0.64%) + Caffeine (nan%)	0.99	TP	FN	0.64	TP	TP
5405	Fentanyl (3.68%) + Bromazolam (0.95%) + Xylazine (0.5%) + Caffeine (nan%)	0.95	TP	FN	0.5	TP	TP
5416	Fentanyl (17.62%) + Fluorofentanyl (0.37%) + Xylazine (0.96%) + Caffeine (nan%)	0	TN	TN	0.96	TP	FN
5417	Fentanyl (6.45%) + ANPP (0.46%) + Fluorofentanyl (0.65%) + Xylazine (1.74%) + Caffeine (nan%)	0	TN	TN	1.74	TP	TP
5424	Fentanyl (6.21%) + Fluorofentanyl (0.71%) + Xylazine (1.64%) + Benzodiazepine (undifferentiated) (nan%) + Caffeine (nan%)	0	FP	TN	1.64	TP	TP
5452	Fentanyl (15.48%) + Etizolam (1.06%) + Flubromazepam (2.8%) + Xylazine (4.57%) + Caffeine (nan%)	0	TN	TN	4.57	TP	TP
5453	Fentanyl (10.53%) + Etizolam (1.06%) + Xylazine (0.88%) + Caffeine (nan%)	0	TN	TN	0.88	TP	TP
5464	Fentanyl (4.4%) + Fluorofentanyl (0.49%) + Bromazolam (4.32%) + Xylazine (0.98%) + Caffeine (nan%)	4.32	TP	TP	0.98	TP	TP
5493	Fentanyl (11.18%) + Fluorofentanyl (0.65%) + Xylazine (4.12%) + Benzodiazepine (undifferentiated) (nan%) + Caffeine (nan%)	0	TN	TN	4.12	TP	TP
5779	Fentanyl (6.63%) + Xylazine (14.88%) + Caffeine (nan%)	0	TN	TN	14.88	TP	TP
5934	Fentanyl (6.48%) + Xylazine (1.15%) + Fluorofentanyl (0.29%) + Benzodiazepine (undifferentiated) (nan%) + Caffeine (nan%)	0	TN	TN	1.15	TP	TP
10073	Fentanyl (16.21%) + Caffeine (nan%)	0	TN	TN	0	TN	FP
10543	Fentanyl (1.36%) + Fluorofentanyl (7.87%) + Bromazolam (3.15%) + Xylazine (3.17%) + Caffeine (nan%)	3.15	TP	TP	3.17	TP	TP
10549	Fentanyl (20.46%) + Fluorofentanyl (0.46%) + + Caffeine (nan%)	0	TN	TN	0	TN	FP
10866	Fentanyl (17.97%) + Caffeine (nan%)	0	TN	TN	0	TN	TN
10876	Fentanyl (12.02%) + Caffeine (nan%)	0	TN	TN	0	TN	TN
10880	Fentanyl (9.14%) + Caffeine (nan%)	0	TN	TN	0	TN	TN
10898	Fentanyl (10.96%) + Caffeine (nan%)	0	TN	FP	0	TN	TN
12816	Fluorofentanyl Base (71.48%) + Fentanyl (0.75%)	0	TN	TN	0	TN	FP
12992	Fentanyl (2.21%) + Fluorofentanyl (7.94%) + Caffeine (nan%) + ANPP (1.66%)	0	TN	TN	0	TN	FP
13141	Fluorofentanyl (nan%) + ANPP (1.08%) + Fentanyl (71.68%) + 4-Anilino-boc-piperidine (nan%)	0	TN	TN	0	TN	FP
13163	Fentanyl (0.16%) + Fluorofentanyl (12.13%) + Bromazolam (0.82%) + Xylazine (8.47%) + Caffeine (nan%)	0.82	TP	TP	8.47	TP	TP
13167	Fluorofentanyl (19.7%) + Xylazine (0.1%) + Caffeine (nan%)	0	FP	FP	0.1	FN	TP
13830	Fluorofentanyl (37.02%) + Fentanyl (0.29%) + Caffeine (nan%)	0	TN	FP	0	TN	FP
13833	Fluorofentanyl (5.77%) + Xylazine (5.43%) + Caffeine (nan%)	0	TN	FP	5.43	TP	TP
14414	Fluorofentanyl (19.24%) + Fentanyl (0.16%) + Bromazolam (1.29%) + Xylazine (13.06%) + Caffeine (nan%)	1.29	TP	TP	13.06	TP	TP
14793	Fentanyl (5.57%) + Fluorofentanyl (3.74%) + Bromazolam (0.9%) + Xylazine (1.84%) + Caffeine (nan%)	0.9	TP	TP	1.84	TP	TP
14799	Fentanyl (11.63%) + Fluorofentanyl (4.35%) + Bromazolam (0.96%) + Xylazine (0.79%) + Caffeine (nan%)	0.96	TP	TP	0.79	FN	TP
15080	Fluorofentanyl (nan%)	0	TN	TN	0	TN	FP
15083	Fluorofentanyl (21.16%) + Fentanyl (1.69%) + Bromazolam (10.89%) + Caffeine (nan%)	10.89	FN	TP	0	TN	TN
15095	Fluorofentanyl (3.49%) + Fentanyl (2.46%) + Bromazolam (1.7%) + Caffeine (nan%)	1.7	TP	FN	0	TN	FP
15096	Fentanyl (25.53%) + Fluorofentanyl (1.89%) + Xylazine (0.47%) + Benzodiazepine (undifferentiated) (nan%) + Caffeine (nan%)	0	FP	TN	0.47	TP	TP
15106	Fentanyl (8.94%) + Fluorofentanyl (0.83%) + Bromazolam (3.01%) + ANPP (1.3%) + Chloroisobutyryl fentanyl (0.14%) + Caffeine (nan%)	3.01	TP	FN	0	TN	FP
15107	Fentanyl (4.6%) + Fluorofentanyl (2.29%) + Bromazolam (2.01%) + Caffeine (nan%)	2.01	TP	TP	0	TN	TN
15108	Fentanyl (12.51%) + Chloroisobutyryl fentanyl (1.1%) + ANPP (3.9%) + Acetylfentanyl (0.29%) + Caffeine (nan%)	0	TN	TN	0	TN	FP
15265	Fentanyl (13.24%) + Isobutyryl fentanyl (4.54%) + Heroin (3.16%) + Bromazolam (9.01%) + Caffeine (nan%)	9.01	TP	TP	0	FP	FP
15269	ANPP (0.57%) + Fentanyl (14.89%) + Fluorofentanyl (0.63%) + Bromazolam (0.37%) + Caffeine (nan%)	0.37	FN	TP	0	TN	TN
15481	Fluorofentanyl (21.0%) + Caffeine (nan%)	0	TN	FP	0	TN	FP
15482	Fentanyl (0.4%) + Fluorofentanyl (3.59%) + Bromazolam (1.46%) + Xylazine (4.46%) + Caffeine (nan%)	1.46	TP	TP	4.46	TP	TP
15483	ANPP (0.59%) + Fentanyl (34.6%) + Bromazolam (35.69%) + Flualprazolam (1.06%) + Xylazine (7.48%)	35.69	TP	TP	7.48	TP	TP
15487	Heroin (2.47%) + Fentanyl (11.32%) + Bromazolam (1.68%) + Caffeine (nan%)	1.68	TP	FN	0	TN	FP
15489	Fentanyl (1.76%) + Fluorofentanyl (0.38%) + Bromazolam (0.32%) + Flubromazolam (2.15%) + Xylazine (0.15%) + Caffeine (nan%)	0.32	TP	TP	0.15	TP	TP
15498	ANPP (0.65%) + Fentanyl (12.62%) + Isobutyryl fentanyl (3.73%) + Bromazolam (12.38%) + Caffeine (nan%)	12.38	TP	TP	0	TN	FP
15501	ANPP (0.5%) + Fentanyl (12.92%) + Fluorofentanyl (0.55%) + Bromazolam (0.43%) + Caffeine (nan%)	0.43	FN	TP	0	TN	TN
15503	Fentanyl (1.68%) + Fluorofentanyl (2.0%) + Bromazolam (5.88%) + Caffeine (nan%)	5.88	TP	TP	0	TN	FP
15504	Fentanyl (9.83%) + Fluorofentanyl (15.42%) + Bromazolam (5.57%) + Caffeine (nan%)	5.57	TP	TP	0	TN	TN
15505	Fentanyl (11.5%) + Fluorofentanyl (24.25%) + Bromazolam (3.11%) + Caffeine (nan%)	3.11	TP	TP	0	TN	TN
15550	Fentanyl (18.35%) + Bromazolam (8.23%) + Xylazine (15.3%) + Flubromazolam (1.94%) + Caffeine (nan%)	8.23	TP	TP	15.3	TP	TP
15617	Fentanyl (58.05%) + Bromazolam (nan%) + Xylazine (15.36%) + ANPP (0.97%) + Flualprazolam (1.9%)	≤ 23.72	TP	TP	15.36	TP	TP

Table 3.6: Complete breakdown of validation set ($n = 50$) sample composition as identified by paper spray mass spectrometry and benzodiazepine immunoassay test strips. Concentration values of bromazolam (green) and xylazine (blue) are described next to true positive (TP), true negative (TN), false positive (FP), and false negative (FN) prediction results of the SERS and IR models. Correct predictions by the respective models are shown in white, while incorrect predictions are in grey.

analyze, warranting the continued exploration of tailoring models to real-world samples. It should be noted that the unpredictable nature of illicit street samples can also introduce complications in model development. The variability of compounds and adulterant concentrations in street samples, limits its use in the development of quantitative models. The complex nature of these samples also means the model performance may be influenced by variations in matrix components not encountered in the training phase. Future SERS studies alongside targeted mass spectrometry to develop larger labeled data sets could vastly benefit the drug checking community and expand automated algorithms for the rapid and accurate detection of trace adulterants in complex mixtures.

3.6 Conclusions

The trace detection of two common adulterants, bromazolam and xylazine, was demonstrated in complex illicit opioid samples using a random forest classifier trained on SERS spectral data. In both cases, the performance exceeded that obtained using FTIR spectroscopy which is more commonly employed in point-of-care drug checking applications, even when similar machine learning approaches were used. Visualization of the features determined to have the highest contribution to the predictive success revealed characteristic regions within the SERS spectra of bromazolam and xylazine. This enabled a secondary model to be trained that further improved sensitivity and specificity in the classification of samples in the low confidence range. Overall, this illustrates the capability of SERS when applied to real-world drug samples for the accurate detection and differentiation of potentially harmful sedatives that are increasingly present in the opioid supply.

Chapter 4

Optimized Machine Learning Approaches to SERS–IR Data Fusion Strategies for Trace Detection of Xylazine Adulterated in Opioids

4.1 Introduction

Implementing multiple strategies for compound identification can enhance drug checking capabilities and offset the limitations of individual techniques [40]. Many point-of-care facilities currently employ this approach by combining general sample information from portable spectroscopic instruments, such as Raman or IR spectroscopy, with the specific identification capabilities of immunoassay test strips. This combination ensures both broad-spectrum analysis and trace detection of fentanyl or benzodiazepine classes, reducing the likelihood of false negatives or undetected substances.

Maintaining broad-spectrum analysis while improving the LoD for spectroscopic instruments can be facilitated by using a surface-sensitive technique like SERS for trace detection of adulterants. This can complement currently implemented techniques used in point-of-care drug checking, specifically IR which is widely used for detecting actives and bulk cutting agents. Facilitating a method for fusing the predictive capabilities of two spectral platforms, SERS and IR, can lead to more robust and accurate models for compound identification. Data fusion can significantly increase predictive success in model development and can be implemented by combining spectral features directly or fusing predictive results from individual models [110].

Given the low performance of the RF model coupled with IR spectroscopy for detecting xylazine, combining spectral data from a high-accuracy method like SERS could provide additional feature importance and improve the performance of both platforms. Investigation into different supervised chemometric approaches, such as support vector machine (SVM) and *k*-nearest neighbors (KNN), have demonstrated success alongside RF in data fusion applications [111]. Supervised models like RF, SVM, and KNN can be optimized through hyperparameter tuning and preprocessing adjustments to maximize their effectiveness in handling fused spectral data.

As drug checking becomes more widespread, it is important to recognize that no single analytical method can fully revolutionize and perfect compound identification [40]. Therefore, exploring synergistic combinations of techniques to enhance automated trace compound identification is a valuable area of research.

Two data fusion strategies using SERS and IR spectral data were examined to evaluate the predictive performance of RF, SVM, and KNN optimized models for the detection of xylazine in complex opioid samples. Given the spectral interference that hinders xylazine detection using IR, six advanced supervised machine learning classifiers were developed to assess overall improvements or reductions to performance for early and late data fusion strategies. The findings indicate that the two random forest models built using concatenated SERS–IR spectral data and a high level weighted fusion of SERS and IR models outperform SVM and KNN models. The high sensitivity and specificity afforded to RF classification using either fusion method illustrate the potential for effectively utilizing IR data with SERS for component prediction in point-of-care drug checking.

4.2 Data Fusion Strategies

This study explores two data fusion strategies: a hybrid and high level approach. An overview of the procedure is presented in Figure 4.1, inspired by a similar representation in Ref. 110.

4.2.1 Hybrid Data Fusion

Typical data fusion strategies that collate raw spectra or features before model development are classified as low level and mid level fusion [112]. Both are practical approaches for combining data, preserving the original spectral information, and maintaining a rich dataset for future model development.

4.2.1.1 Low Level Data Fusion

This involves combining raw or preprocessed data directly from different spectral sources before any feature extraction or modeling. This method retains all the original data information, providing a comprehensive dataset for analysis.

4.2.1.2 Mid Level Data Fusion

This fusion approach combines features that have been extracted and individually processed from different spectral data sets before integration. This approach balances detail and simplicity by focusing on essential features while reducing data complexity.

In this study, a hybrid fusion technique is employed, integrating aspects of both low level and mid level fusion.

The hybrid SERS–IR data fusion technique involves concatenating the spectral data from both SERS and IR, integrating the information at the raw data level (low level fusion). This process enhances the richness of the data by leveraging the complementary information provided by both SERS and IR techniques. By concatenating the SERS and IR data before performing outlier detection, variability is analyzed in a unified principal component space. This approach allows for the identification and removal of samples

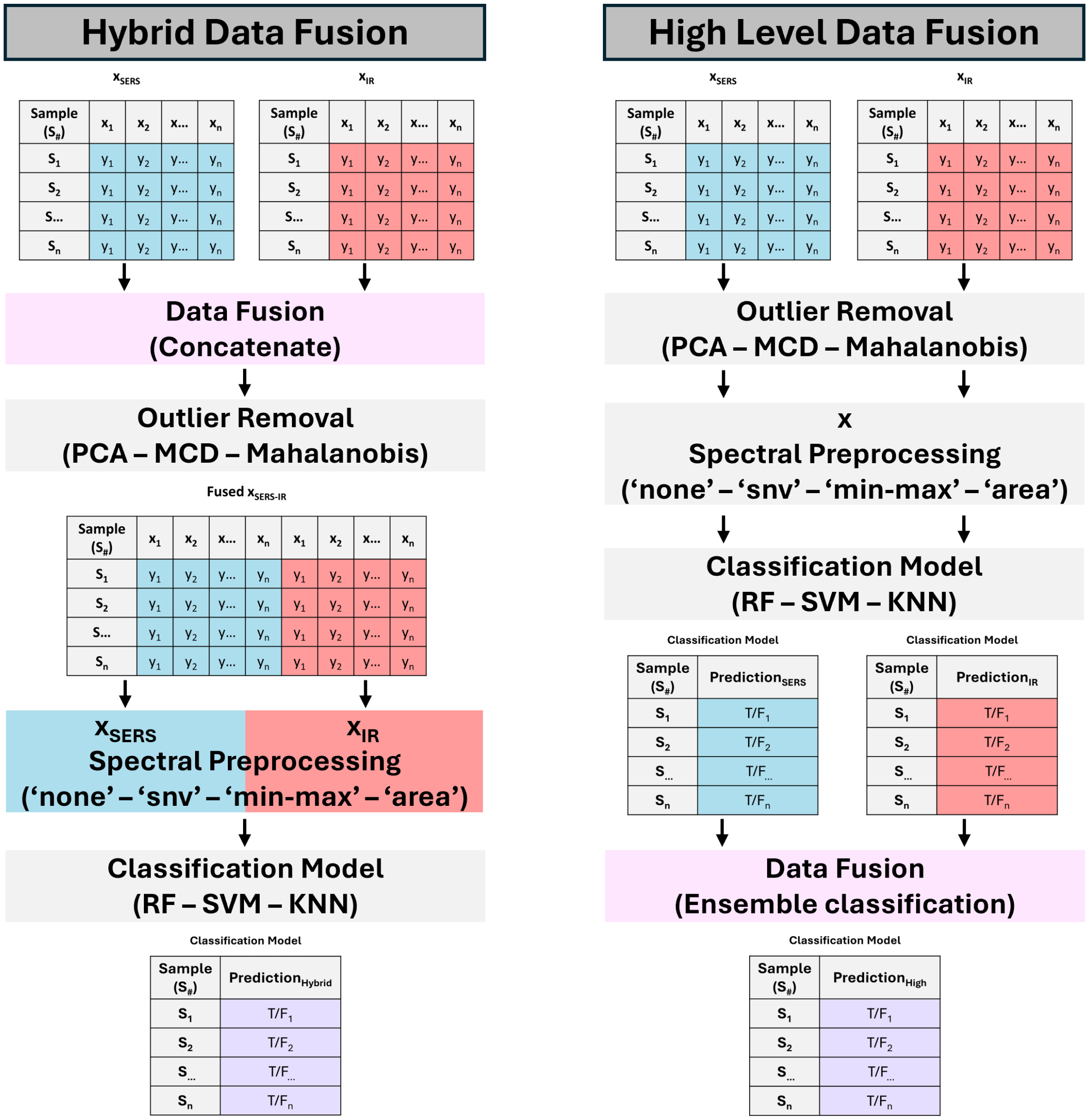


Figure 4.1: Flowchart of hybrid and high level fusion strategies.

with high variance based on the combined spectral data, rather than treating each dataset separately. Consequently, the samples causing variability are reflective of both SERS and IR data, leading to a more consistent and robust outlier detection process before model development.

The variance reduced SERS–IR hybrid training set ($n = 154$) undergoes a hyperparameter tuning stage where different spectral preprocessing techniques, including derivatives and normalization, are tested separately on SERS and IR data within the fused dataset (mid level fusion). This approach is facilitated by a grid search for hyperparameter tuning of the RF, SVM, and KNN cross-validated models. By preprocessing each part of the fused spectrum separately within the hyperparameter search, the specific characteristics and features of SERS and IR data are optimized individually before combining them for model development. Therefore, the optimal combination of SERS and IR preprocessing techniques can be computationally tested along with model parameters to potentially enhance the performance of the classification models in the training phase. Results for the optimized model parameters for the hybrid data fusion method using the concatenated SERS–IR spectral data are shown for RF in Table 4.1, SVM in Table 4.2, and KNN in Table 4.3.

4.2.2 High Level Data Fusion

High level data fusion combines the outputs of models built on different spectral data sets [112, 113]. This approach enhances overall performance and robustness through the merger of these decisions. Utilizing high level fusion through ensemble methods allows for improved accuracy and consistency, and has proven successful in model development across diverse applications [110, 114]. A significant advantage of this approach is that combining decisions from multiple models is often easier to implement into existing systems compared to early fusion techniques. High level data fusion enables the use of various models and techniques for different data sources, allowing the

Random Forest (RF)

Parameter Name	Description	Grid Search Values	Hybrid SERS—IR Data Fusion	
			Best Parameter	F1 Score
n_estimators	The number of trees in RF	5, 10, 20, 30, 40, 50, 100, 150, 200, 300, 400, 500, 800, 1000, 2000	50	0.743
max_depth	The maximum depth of a single tree	None, 1, 5, 10, 20, 30, 50	None	
min_samples_leaf	The minimum number of samples required to be at a leaf node	1, 5, 10, 20, 50	1	
deriv_sers	The order of derivatives of the SERS spectral data	0, 1, 2	0	
deriv_ir	The order of derivatives of the IR spectral data	0, 1, 2	0	
norm_sers	The normalization process of the SERS spectral data	‘min-max', 'snv', 'none', 'area'	‘snv’	
norm_ir	The normalization process of the IR spectral data	‘min-max', 'snv', 'none', 'area'	'snv’	

Table 4.1: Investigation of RF hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search for the detection of xylazine using the hybrid SERS–IR data fusion method. Showing the combinations tested and the optimal set of parameters selected based on F1 score.

Support Vector Machine (SVM)

Parameter Name	Description	Grid Search Values	Hybrid SERS—IR Data Fusion	
			Best Parameter	F1 Score
C	Regularization parameter. The strength of the regularization is inversely proportional to C	1, 10, 100	10	0.881
kernel	Specifies the kernel type to be used in the algorithm	'linear', 'rbf', 'poly', 'sigmoid'	‘rbf’	
gamma	Kernel coefficient for 'rbf', 'poly', and 'sigmoid' kernels. It defines how far the influence of a single training example reaches	'auto’, ‘scale'	‘scale’	
degree	Degree of the polynomial kernel function ('poly')	1, 2, 3, 4, 5	N/A	
coef0	Independent term in the kernel function (Only significant for ‘poly’ and ‘sigmoid’ kernels)	0.0, 0.1, 0.5, 1.0, 2.0	N/A	
class_weight	Weights associated with classes	None, ‘balanced’	‘balanced’	
deriv_sers	The order of derivatives of the SERS spectral data	0, 1, 2	0	
deriv_ir	The order of derivatives of the IR spectral data	0, 1, 2	2	
norm_sers	The normalization process of the SERS spectral data	‘min-max', 'snv', 'none', 'area'	‘area’	
norm_ir	The normalization process of the IR spectral data	‘min-max', 'snv', 'none', 'area'	'none’	

Table 4.2: Investigation of SVM hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search for the detection of xylazine using the hybrid SERS–IR data fusion method. Showing the combinations tested and the optimal set of parameters selected based on F1 score.

k-Nearest Neighbors (KNN)

Parameter Name	Description	Grid Search Values	Hybrid SERS—IR Data Fusion	
			Best Parameter	F1 Score
n_neighbors	The number of neighbors to use for k-neighbors queries	1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 35, 37, 39	1	0.748
weights	Weight function used in prediction	'uniform', 'distance'	‘uniform’	
algorithm	Algorithm used to compute the nearest neighbors	'auto', 'ball_tree', 'kd_tree', 'brute'	‘auto’	
metric	How the algorithm calculates the distance between data points	'euclidean', 'manhattan', 'minkowski'	‘manhattan’	
deriv_sers	The order of derivatives of the SERS spectral data	0, 1, 2	0	
deriv_ir	The order of derivatives of the IR spectral data	0, 1, 2	0	
norm_sers	The normalization process of the SERS spectral data	‘min-max', 'snv', 'none', 'area'	‘area’	
norm_ir	The normalization process of the IR spectral data	‘min-max', 'snv', 'none', 'area'	'area’	

Table 4.3: Investigation of KNN hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search for the detection of xylazine using the hybrid SERS–IR data fusion method. Showing the combinations tested and the optimal set of parameters selected based on F1 score.

integration of predictive results from models built with different training data [115–118] or classifiers [113].

The high level data fusion used in this study merges predictive results from two classification models: one built with SERS spectral data and the other with IR spectral data, similar to the approach described in Chapter 3. RF, SVM, and KNN models for both SERS and IR data are developed using a 5-fold cross-validation grid search. Results for the optimized SERS and IR model parameters for the high level data fusion method are shown for RF in Table 4.4, SVM in Table 4.5, and KNN in Table 4.6. Each pair of SERS and IR models is then combined using an ensemble classification method, resulting in fused RF, SVM, and KNN models that utilize information from both spectral sources.

The ensemble classification tool is a crucial component of the high level data fusion process and can significantly affect prediction results. Several methods for ensemble classification exist, including voting, stacking, and blending. A voting classifier combines the predictions of multiple models by averaging (soft voting) or taking a majority vote (hard voting). Weighted voting assigns different weights to each model's prediction based on performance, directly incorporating model effectiveness into the ensemble. In the case of single component identification of xylazine, if one model significantly outperforms the other, as seen in Section 3.4.3, overall model performance can be improved by leveraging the strengths of the superior model. Therefore, a weighted voting ensemble was selected for the high level fusion process in this study.

A weighted voting classifier (sklearn.ensemble) was constructed to combine the predictions from both SERS and IR models. The weights of the prediction models were optimized using a 5-fold cross-validation grid search on the training data. The voting classifier was initialized in 'soft' voting mode to allow for probability-based weighting. The optimal combination of weights for the RF, SVM, and KNN models was selected based on the F1 score. The best weights identified were used to train the final high level fused models, which were then evaluated on the validation set to assess performance.

Random Forest (RF)

Parameter Name	Grid Search Values	SERS		IR	
		Best Parameter	F1 Score	Best Parameter	F1 Score
n_estimators	5, 10, 20, 30, 40, 50, 100, 150, 200, 300, 400, 500, 800, 1000, 2000	1000	0.802	10	0.530
max_depth	None, 1, 5, 10, 20, 30, 50	None		50	
min_samples_leaf	1, 5, 10, 20, 50	1		5	
deriv	0, 1, 2	0		0	
norm	‘min-max', 'snv', 'none', 'area'	‘area’		‘none’	

Table 4.4: Investigation of RF hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search for the detection of xylazine using SERS and IR data. Showing the combinations tested and the optimal set of parameters selected based on F1 score.

Support Vector Machine (SVM)

Parameter Name	Grid Search Values	SERS		IR	
		Best Parameter	F1 Score	Best Parameter	F1 Score
C	1, 10, 100	100	0.869	100	0.790
kernel	'linear', 'rbf', 'poly', 'sigmoid'	‘poly’		‘poly’	
gamma	'auto’, ‘scale'	‘auto’		‘scale’	
degree	1, 2, 3, 4, 5	2		3	
coef0	0.0, 0.1, 0.5, 1.0	0.5		0.1	
class_weight	None, ‘balanced’	‘balanced’		‘balanced’	
deriv	0, 1, 2	0		1	
norm	‘min-max', 'snv', 'none', 'area'	‘snv’		‘snv’	

Table 4.5: Investigation of SVM hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search for the detection of xylazine using SERS and IR data. Showing the combinations tested and the optimal set of parameters selected based on F1 score.

k-Nearest Neighbors (KNN)

Parameter Name	Grid Search Values	SERS		IR	
		Best Parameter	F1 Score	Best Parameter	F1 Score
n_neighbors	1, 2, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 25, 30	1	0.690	1	0.618
weights	'uniform', 'distance'	‘uniform’		‘uniform’	
algorithm	'auto', 'ball_tree', 'kd_tree', 'brute'	‘auto’		‘auto’	
metric	'euclidean', 'manhattan', 'minkowski'	‘manhattan’		'euclidean'	
deriv	0, 1, 2	0		2	
norm	‘min-max', 'snv', 'none', 'area'	‘area’		‘min-max’	

Table 4.6: Investigation of KNN hyperparameters and spectral preprocessing techniques using a 5-fold cross-validation grid search for the detection of xylazine using SERS and IR data. Showing the combinations tested and the optimal set of parameters selected based on F1 score.

4.3 Optimization of RF, SVM, and KNN Models

4.3.1 Hybrid Data Fusion Model Development

F1 scores of 0.743, 0.881, and 0.748 were obtained from the optimized RF (Table 4.1), SVM (Table 4.2), and KNN (Table 4.3) classifiers, respectively, for predicting xylazine using concatenated SERS–IR spectral data during cross-validation. All hybrid models demonstrated relatively high predictive performance, with the SVM model achieving the highest precision and recall (F1 score of 0.881) within the SERS–IR training set.

The optimal hyperparameters for the hybrid SVM model (Table 4.2) included the ‘rbf’ kernel, which effectively handles non-linear classification by mapping the input space into a higher-dimensional space, and the ‘scale’ kernel coefficient, which adjusts the gamma parameter based on the number of features, thereby controlling the influence of each data point. The chosen preprocessing methods suggest that applying the second derivative to the IR data enhances relevant features, while the raw SERS data, normalized by area, is sufficiently informative. Thus, the optimized hybrid SVM model is configured to manage imbalanced data, fit complex non-linear relationships, and emphasize absolute intensities in the IR spectra while standardizing the SERS spectra.

For the hybrid RF model (Table 4.1), spectral preprocessing selected the same hyperparameters for both SERS and IR regions, with no derivative (0) and SNV normalization, indicating a consistent approach to reducing intensity variations across the entire dataset. Similarly, the hybrid KNN model (Table 4.3) selected area normalization, ensuring consistent total signal intensity across all samples. This uniformity in spectral preprocessing for the RF and KNN models suggests that the concatenated SERS–IR spectral data share underlying characteristics best captured when treated as a whole.

The selection of 50 decision trees for the hybrid RF model, indicates a balanced approach to model complexity and computational efficiency. The use of 50 trees allows the model to capture more diverse patterns related to xylazine while mitigating the risk of overfitting to the training data. In contrast, the hybrid KNN model employed a simple approach by using a single neighbour for predictions, relying on the most immediate data point's characteristics, indicating a potential for high bias to the training set.

4.3.2 High Level Data Fusion Model Development

4.3.2.1 Random Forest Models

F1 scores of 0.802 for SERS and 0.530 for IR indicate varied predictive performance during cross-validation (Table 4.4). RF models built with the two spectroscopic platforms show similar trends to the F1 scores and optimal hyperparameters to Section 3.4. The SERS RF model demonstrates good initial performance with a high number of decision trees and consistent normalization techniques, capturing complex patterns in the SERS data. Conversely, the IR RF model, with only 10 trees and a max depth of 50, likely suffers from underfitting, resulting in poorer performance, and an F1 score close to that of a baseline classifier. This suggests that IR features are less effective for predicting xylazine. However, high precision and recall of the high level SVM IR model elucidate the potential of IR features with different methods.

4.3.2.2 Support Vector Machine Models

SVM models built with SERS and IR data achieved the highest individual precision and recall scores during cross-validation (Table 4.5), with F1 scores of 0.869 for SERS and 0.790 for IR. Early performance metrics on the training set suggest both spectral devices seem compatible with SVM classification. The optimized SVM parameters for both SERS and IR data indicate a strong focus on capturing complex, non-linear relationships in the data. The use of polynomial kernels with different degrees (2 for SERS and 3 for IR) allows the models to capture interactions between features, while 'balanced' class weights help address class imbalance. High values of C suggest that the models are less regularized, aiming to fit the training data closely. Overall, these parameters highlight the tailored approach taken to optimize the SVM models for the specific characteristics of SERS and IR spectral data, resulting in high initial predictive performance as indicated by their respective F1 scores.

4.3.2.3 *k*-Nearest Neighbors Models

KNN models for SERS (F1 score 0.690) and IR (F1 score 0.618) showed no strong preference for either platform in detecting xylazine during cross-validation (Table 4.6). The individual KNN models both demonstrated lower precision and recall scores compared to their hybrid counterpart. The use of a single neighbour and Manhattan distance by the SERS model emphasizes its immediate data points, while the IR model applies the second derivative and Euclidean distance to capture subtle features. However, both utilize simple approaches with a single neighbour which may lead to underfitting and limit performance on the validation set.

4.3.2.4 Weighted Voting Ensemble

Despite varied F1 scores, all individual models consistently favoured SERS as a more robust platform for detecting xylazine compared to IR. SERS models demonstrated a distinct preference for normalization techniques over derivatives, aligning with the optimal preprocessing methods selected for all hybrid models as well.

High level fusion of SERS and IR models evaluated F1 scores of all weight combinations summing to 1 during cross-validation on the training set. The fused RF, SVM, and KNN models, with optimal weights [SERS, IR], achieved F1 scores of 0.744, 0.795, and 0.637, respectively. These values were consistent with the initial predictive performance results for each spectral model.

Optimal weights prioritized SERS information in all the final high level fusion models. A comprehensive breakdown of the optimal parameters for individual and hybrid models, the weights selected for high level data fusion, and F1 scores of the final ensemble models on the training set are illustrated in Table 4.7.

Data Fusion	Model	Training Data	Hyperparameters	F1 Score	Voting Classifier	Weights [SERS, IR]	Ensemble F1 Score
Hybrid	RF	SERS—IR	n_estimators: 50, max_depth: None, min_samples_leaf: 1, deriv_sers: 0, deriv_ir: 0, norm_sers: 'snv', norm_ir: 'snv'	0.743	-	-	-
	SVM	SERS—IR	C: 10, kernel: 'rbf', gamma: 'scale', degree: N/A, coef0: N/A, class_weight: 'balanced', deriv_sers: 0, deriv_ir: 2, norm_sers: 'area', norm_ir: 'area'	0.881	-	-	-
	KNN	SERS—IR	n_neighbors: 1, weights: 'uniform', algorithm: 'auto’, metric: 'manhattan', deriv_sers: 0, deriv_ir: 0, norm_sers: 'area', norm_ir: 'area'	0.748	-	-	-
High Level	RF	SERS	n_estimators: 1000, max_depth: None, min_samples_leaf: 1, deriv: 0, norm: 'area'	0.802	Weighted Voting	[0.9, 0.1]	0.744
		IR	n_estimators: 10, max_depth: 50, min_samples_leaf: 5, deriv: 0, norm: 'none'	0.530			
	SVM	SERS	C: 100, kernel: 'poly', gamma: 'auto', degree: 2, coef0: 0.5, class_weight: 'balanced', deriv: 0, norm: 'snv'	0.869	Weighted Voting	[0.9, 0.1]	0.795
		IR	C: 100, kernel: 'poly', gamma: 'scale', degree: 3, coef0: 0.1, class_weight: 'balanced', deriv: 1, norm: 'snv'	0.790			
	KNN	SERS	n_neighbors: 1, weights: 'uniform', algorithm: 'auto’, metric: 'manhattan', deriv: 0, norm: 'area'	0.690	Weighted Voting	[0.6, 0.4]	0.637
		IR	n_neighbors: 1, weights: 'uniform', algorithm: 'auto’, metric: 'euclidean', deriv: 2, norm: 'min-max'	0.618			

Table 4.7: Optimized parameters and F1 scores for all models built with SERS, IR, and SERS–IR training data for the detection of xylazine. Optimal weights selected for SERS and IR model contribution and their respective F1 scores are detailed for the construction of high level classification models.

4.4 Analysis of Models

While the cross-validated F1 scores are a good metric to determine initial performance on the training set folds, the results of the fused models on unseen data is crucial to assessing their performance. All fused models using the two strategies (hybrid and high level) are tested on the validation set ($n = 50$) to evaluate model performance for the prediction of xylazine. The evaluation metrics include Area Under the Curve (AUC), optimal threshold, accuracy, precision, sensitivity, and specificity for the detection of xylazine in complex opioid samples. The performance of all models is evaluated to identify the most effective data fusion strategy and classification algorithm for this application. ROC analysis was performed on the hybrid and high level models for the validation set as shown in Figure 4.2. AUC values were calculated and optimal thresholds were determined for final prediction results.

AUC values of 0.92, 0.86, and 0.68 were calculated for the hybrid RF, SVM, and KNN models, respectively (Figure 4.2a). Both the RF and SVM hybrid models reflect high discriminatory power across the range of thresholds. The hybrid KNN model demonstrated a weaker performance across the range of predicted probabilities, likely due to the selected number of neighbours, where $k = 1$ results in predictions that are highly sensitive to the nearest neighbour (Table 4.3). In scenarios where the nearest neighbour clearly belong to one class (0 or 1), the predicted probability of a tested sample will reflect the class and the scores for all tested samples will be either 0 or 1. Consequently, setting a threshold becomes arbitrary and no improvements can be made by changing it. Therefore, despite relatively high initial performance within the training group, the hybrid KNN model does not seem well suited to generalize to unseen data or to assign meaningful probabilities.

AUC values calculated for the high level RF (0.93), SVM (0.87), and KNN (0.79) models (Figure 4.2b) show increased performance over the range of predicted probabilities compared to the hybrid AUC values. Most notably, the high level KNN model improved prediction results significantly when combining high level outputs from the independent

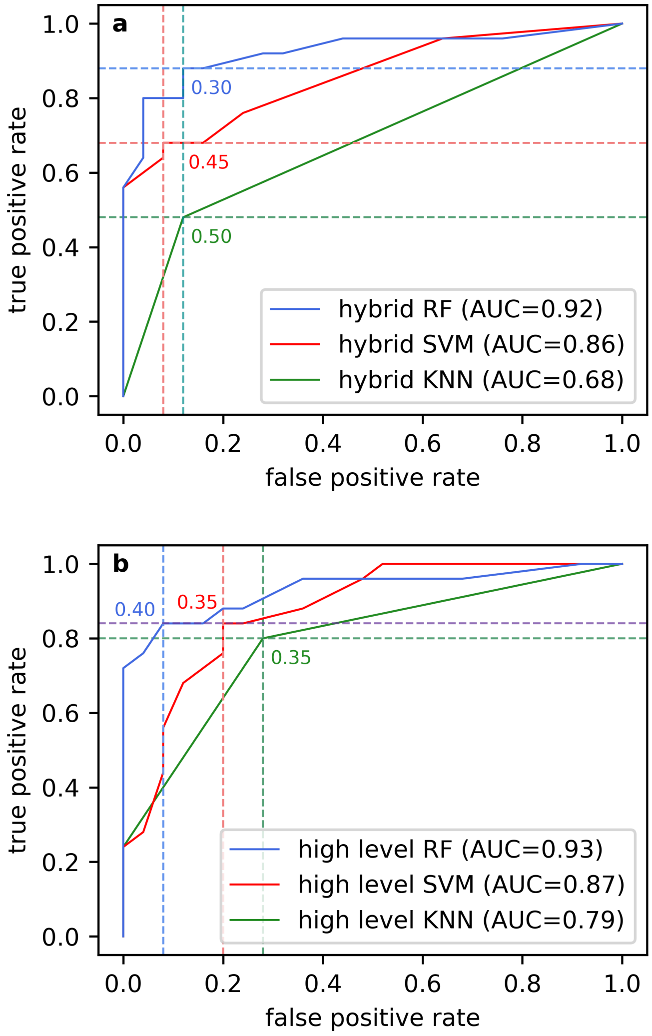


Figure 4.2: A comparison between the RF model, SVM model and KNN model prediction results of xylazine on the test set ($n = 50$) for the (a) hybrid and (b) high level data fusion strategies. The ROC curves illustrate the true positive rate and false positive rate across various classification thresholds for xylazine detection. Optimal threshold selection for balancing sensitivity and specificity and area under the curve (AUC) is highlighted for all models.

SERS and IR models.

Optimal predicted probability score thresholds for all models were determined using Youden's J statistic and were applied to the test set. The xylazine prediction results are shown for the hybrid (Table 4.8a) and high level (Table 4.8b) models and were used to evaluate the performance of the 2 fusion strategies.

a. Hybrid SERS—IR Data Fusion

	RF Model on Test Set ($n = 50$)	SVM Model on Test Set ($n = 50$)	KNN Model on Test Set ($n = 50$)
True positive (n)	22	17	12
True negative (n)	22	23	22
False positive (n)	3	2	3
False negative (n)	3	8	13

b. High Level SERS—IR Data Fusion

	RF Model on Test Set ($n = 50$)	SVM Model on Test Set ($n = 50$)	KNN Model on Test Set ($n = 50$)
True positive (n)	21	21	20
True negative (n)	23	20	18
False positive (n)	2	5	7
False negative (n)	4	4	5

Table 4.8: Xylazine prediction results of the test set ($n = 50$) for the RF, SVM, and KNN models developed using the (a) hybrid SERS–IR data fusion method and (b) high level SERS–IR data fusion method.

Of the hybrid models (Table 4.8a), Random forest correctly identified 22 out of the 25 xylazine-positive samples and 22 out of the 25 xylazine-negative samples. Showing a balanced sensitivity and specificity towards detecting xylazine. Both SVM and KNN hybrid models demonstrated a higher affinity towards identifying negative classes than they did for correctly assigning positive classes. Initial performance in the training group for SVM showed the highest precision and recall scores, but the performance on the validation set is indicative of overfitting of the hybrid SVM model to the training data.

The hybrid KNN model performs poorly in correctly assigning positive classes, likely due to an imbalance in the training dataset, where the nearest neighbour is often from the majority class (xylazine-negative). This leads to underprediction of the minority class (xylazine-positive), as the single nearest neighbour may not accurately reflect the true class distribution. KNN's effectiveness also decreases with high-dimensional data, where the distance between points becomes less informative. Concatenating SERS and IR data significantly increases dimensionality, making the nearest neighbour less representative of the true class. Additionally, the computational intensity required for large data sets [119] further highlights the unsuitability of this model development strategy.

The hybrid data fusion method appears to best suit the RF binary classifier, leveraging spectral information from both platforms to achieve high accuracy and precision compared to other supervised techniques (Table 4.9).

Hybrid SERS—IR Data Fusion

Model	AUC	Optimal Threshold	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)
RF	0.92	0.30	88	88	88	88
SVM	0.86	0.45	80	89	68	92
KNN	0.68	0.50	68	80	48	88

Table 4.9: Performance metrics summary for RF, SVM, and KNN models using hybrid SERS–IR data fusion method for xylazine detection. Summary of the Area Under the Curve (AUC), optimal threshold, accuracy, precision, sensitivity, and specificity for opioid samples ($n = 50$) tested with RF, SVM, and KNN models.

Among the high level models (Table 4.8b), KNN has the worst overall performance; however, it manages to prioritize and improve the classification of positive classes compared to the hybrid KNN model. Although overprediction of the positive class is a concern for any model development, the high level KNN model achieves a relatively balanced true positive and false positive rate. With a moderately high AUC value (0.79) and, with a threshold of 0.35, demonstrates better performance of the high level KNN model across a range of predicted probability scores compared to the hybrid KNN model. This improvement may be due to the combination of high outputs, which reduces dimensionality and the computational strain KNN experiences with large data sets.

The high level SVM model effectively balances the number of positive and negative classes, correctly identifying an additional four xylazine-positive samples compared to the hybrid model. Despite a decline in true negative performance, all high level models show a more meaningful balance of correctly assigned predictions.

The high level RF model yielded the best predictions among the high level models. Correctly identifying 21 true positive samples and 23 true negative samples. With a 0.9 weight assigned to the RF SERS contribution, the prediction results of the high level RF model (Table 4.10) are comparable to the high performance established by the xylazine SERS model in Section 3.4.

High Level SERS—IR Data Fusion

Model	Voting Classifier	Weights [SERS, IR]	AUC	Optimal Threshold	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)
RF	Weighted Voting	[0.9, 0.1]	0.93	0.40	88	91	84	92
SVM		[0.9, 0.1]	0.87	0.35	82	81	84	80
KNN		[0.6, 0.4]	0.79	0.35	76	74	80	72

Table 4.10: Performance metrics summary for RF, SVM, and KNN models using high level SERS–IR data fusion method for xylazine detection. Presenting the selected voting classifier and the weights of the individual SERS and IR models that created the high level fused RF, SVM, and KNN models. Summary of the area under the curve (AUC), optimal threshold, accuracy, precision, sensitivity, and specificity for opioid samples ($n = 50$) tested with high level fused models.

Both the high level SVM and KNN models outperformed their hybrid counterparts. However, the highest performing models were those using random forest classification, both in hybrid and high level strategies, achieving accuracies of 88% and sensitivities of 88% and 84%, respectively, as shown in Tables 4.9 and 4.10.

RF binary classification demonstrated high predictive success for both data fusion strategies, with the high level RF classifier achieving increased precision (91%) and specificity (92%). The compatibility of the concatenated SERS–IR data for RF model development illustrates robust performance for xylazine detection when IR data is directly incorporated into model training. However, given the RF hyperparameters and identical spectral preprocessing values selected for both SERS and IR regions, it is likely that the hybrid model continues to leverage SERS data for its predictions. Similarly, the high level RF model weighted the individual SERS model predictions at a disproportionate ratio of 9:1 over the IR model.

The findings of this study indicate no reduction in performance when implementing the two spectroscopic techniques for fused model development. Considering the known spectral overlap and difficulties in detecting xylazine with IR (Section 3.4.3), the high level data fusion method is the preferred strategy. Leveraging predictions from individual models prioritizes the stronger method, resulting in robust performance, especially when dealing with the complexities of real-world training and testing groups. This approach of combining RF model outputs is a beneficial strategy for expanding drug checking services and targeting compounds identifiable by either platform. The complete breakdown of sample composition and model results of the validation set is illustrated in Table 4.11.

4.5 Conclusions

The trace detection of xylazine in complex illicit opioid samples was demonstrated using random forest classification trained on two data fusion strategies incorporating SERS and IR spectral data. In both strategies, the performance of the random forest models surpassed that of the SVM and KNN models. Given the known interferences and challenges of identifying xylazine using IR alone, implementing a high level data fusion technique with random forest model development maintained high sensitivity and specificity utilizing both platforms. This approach effectively prioritized the sensitivity of SERS for trace xylazine detection while preserving valuable predictive information from IR, without compromising performance. The integration of multiple analytical techniques, supported by advanced data fusion and machine learning approaches, represents a promising strategy for advancing drug checking. This holistic approach ensures that rapid, affordable, and

AERL	Sample Composition	Xylazine						
		Concentration (% w/w)	Hybrid Fused Model Results			High level Fused Model Results		
			RF	SVM	KNN	RF	SVM	KNN
5184	Fentanyl (6.95%) + Etizolam (1.07%) + Flubromazepam (0.73%) + Xylazine (0.5%) + Fluorofentanyl (0.46%) + Caffeine (nan%)	0.5	TP	TP	TP	TP	TP	TP
5404	Fentanyl (4.36%) + Bromazolam (0.99%) + Xylazine (0.64%) + Caffeine (nan%)	0.64	TP	FN	FN	TP	TP	FN
5405	Fentanyl (3.68%) + Bromazolam (0.95%) + Xylazine (0.5%) + Caffeine (nan%)	0.5	FN	TP	FN	FN	FN	TP
5416	Fentanyl (17.62%) + Fluorofentanyl (0.37%) + Xylazine (0.96%) + Caffeine (nan%)	0.96	TP	TP	FN	TP	TP	TP
5417	Fentanyl (6.45%) + ANPP (0.46%) + Fluorofentanyl (0.65%) + Xylazine (1.74%) + Caffeine (nan%)	1.74	TP	TP	FN	TP	TP	TP
5424	Fentanyl (6.21%) + Fluorofentanyl (0.71%) + Xylazine (1.64%) + Benzodiazepine (undifferentiated) (nan%) + Caffeine (nan%)	1.64	TP	TP	TP	TP	TP	TP
5452	Fentanyl (15.48%) + Etizolam (1.06%) + Flubromazepam (2.8%) + Xylazine (4.57%) + Caffeine (nan%)	4.57	TP	FN	TP	TP	TP	TP
5453	Fentanyl (10.53%) + Etizolam (1.06%) + Xylazine (0.88%) + Caffeine (nan%)	0.88	TP	TP	FN	TP	TP	TP
5464	Fentanyl (4.4%) + Fluorofentanyl (0.49%) + Bromazolam (4.32%) + Xylazine (0.98%) + Caffeine (nan%)	0.98	TP	FN	FN	TP	TP	TP
5493	Fentanyl (11.18%) + Fluorofentanyl (0.65%) + Xylazine (4.12%) + Benzodiazepine (undifferentiated) (nan%) + Caffeine (nan%)	4.12	TP	TP	FN	TP	TP	TP
5779	Fentanyl (6.63%) + Xylazine (14.88%) + Caffeine (nan%)	14.88	TP	TP	FN	TP	TP	FN
5934	Fentanyl (6.48%) + Xylazine (1.15%) + Fluorofentanyl (0.29%) + Benzodiazepine (undifferentiated) (nan%) + Caffeine (nan%)	1.15	FN	FN	FN	FN	TP	FN
10073	Fentanyl (16.21%) + Caffeine (nan%)	0	TN	FP	TN	TN	TN	TN
10543	Fentanyl (1.36%) + Fluorofentanyl (7.87%) + Bromazolam (3.15%) + Xylazine (3.17%) + Caffeine (nan%)	3.17	TP	TP	FN	FN	FN	TP
10549	Fentanyl (20.46%) + Fluorofentanyl (0.46%) + + Caffeine (nan%)	0	FP	TN	FP	FP	FP	TN
10866	Fentanyl (17.97%) + Caffeine (nan%)	0	TN	TN	TN	TN	TN	TN
10876	Fentanyl (12.02%) + Caffeine (nan%)	0	TN	TN	TN	TN	TN	TN
10880	Fentanyl (9.14%) + Caffeine (nan%)	0	TN	TN	TN	TN	TN	TN
10898	Fentanyl (10.96%) + Caffeine (nan%)	0	TN	TN	TN	TN	TN	TN
12816	Fluorofentanyl Base (71.48%) + Fentanyl (0.75%)	0	TN	TN	TN	TN	TN	TN
12992	Fentanyl (2.21%) + Fluorofentanyl (7.94%) + Caffeine (nan%) + ANPP (1.66%)	0	TN	TN	TN	TN	TN	TN
13141	Fluorofentanyl (nan%) + ANPP (1.08%) + Fentanyl (71.68%) + 4-Anilino-boc-piperidine (nan%)	0	TN	TN	TN	TN	TN	FP
13163	Fentanyl (0.16%) + Fluorofentanyl (12.13%) + Bromazolam (0.82%) + Xylazine (8.47%) + Caffeine (nan%)	8.47	TP	TP	TP	TP	TP	TP
13167	Fluorofentanyl (19.7%) + Xylazine (0.1%) + Caffeine (nan%)	0.1	TP	TP	TP	TP	TP	TP
13830	Fluorofentanyl (37.02%) + Fentanyl (0.29%) + Caffeine (nan%)	0	TN	TN	TN	TN	TN	FP
13833	Fluorofentanyl (5.77%) + Xylazine (5.43%) + Caffeine (nan%)	5.43	TP	TP	TP	TP	TP	TP
14414	Fluorofentanyl (19.24%) + Fentanyl (0.16%) + Bromazolam (1.29%) + Xylazine (13.06%) + Caffeine (nan%)	13.06	TP	TP	TP	TP	TP	TP
14793	Fentanyl (5.57%) + Fluorofentanyl (3.74%) + Bromazolam (0.9%) + Xylazine (1.84%) + Caffeine (nan%)	1.84	TP	FN	FN	TP	TP	FN
14799	Fentanyl (11.63%) + Fluorofentanyl (4.35%) + Bromazolam (0.96%) + Xylazine (0.79%) + Caffeine (nan%)	0.79	FN	FN	FN	FN	FN	FN
15080	Fluorofentanyl (nan%)	0	TN	TN	TN	TN	TN	TN
15083	Fluorofentanyl (21.16%) + Fentanyl (1.69%) + Bromazolam (10.89%) + Caffeine (nan%)	0	TN	TN	TN	TN	TN	TN
15095	Fluorofentanyl (3.49%) + Fentanyl (2.46%) + Bromazolam (1.7%) + Caffeine (nan%)	0	TN	TN	TN	TN	TN	TN
15096	Fentanyl (25.53%) + Fluorofentanyl (1.89%) + Xylazine (0.47%) + Benzodiazepine (undifferentiated) (nan%) + Caffeine (nan%)	0.47	TP	TP	TP	TP	TP	TP
15106	Fentanyl (8.94%) + Fluorofentanyl (0.83%) + Bromazolam (3.01%) + ANPP (1.3%) + Chloroisobutyryl fentanyl (0.14%) + Caffeine (nan%)	0	FP	TN	TN	TN	FP	FP
15107	Fentanyl (4.6%) + Fluorofentanyl (2.29%) + Bromazolam (2.01%) + Caffeine (nan%)	0	TN	TN	TN	TN	TN	TN
15108	Fentanyl (12.51%) + Chloroisobutyryl fentanyl (1.1%) + ANPP (3.9%) + Acetylfentanyl (0.29%) + Caffeine (nan%)	0	TN	TN	TN	TN	FP	FP
15265	Fentanyl (13.24%) + Isobutyryl fentanyl (4.54%) + Heroin (3.16%) + Bromazolam (9.01%) + Caffeine (nan%)	0	TN	FP	TN	TN	TN	FP
15269	ANPP (0.57%) + Fentanyl (14.89%) + Fluorofentanyl (0.63%) + Bromazolam (0.37%) + Caffeine (nan%)	0	TN	TN	TN	TN	TN	TN
15481	Fluorofentanyl (21.0%) + Caffeine (nan%)	0	TN	TN	TN	TN	TN	TN
15482	Fentanyl (0.4%) + Fluorofentanyl (3.59%) + Bromazolam (1.46%) + Xylazine (4.46%) + Caffeine (nan%)	4.46	TP	TP	TP	TP	TP	TP
15483	ANPP (0.59%) + Fentanyl (34.6%) + Bromazolam (35.69%) + Flualprazolam (1.06%) + Xylazine (7.48%)	7.48	TP	TP	TP	TP	TP	TP
15487	Heroin (2.47%) + Fentanyl (11.32%) + Bromazolam (1.68%) + Caffeine (nan%)	0	TN	TN	TN	TN	TN	TN
15489	Fentanyl (1.76%) + Fluorofentanyl (0.38%) + Bromazolam (0.32%) + Flubromazolam (2.15%) + Xylazine (0.15%) + Caffeine (nan%)	0.15	TP	FN	TP	TP	TP	TP
15498	ANPP (0.65%) + Fentanyl (12.62%) + Isobutyryl fentanyl (3.73%) + Bromazolam (12.38%) + Caffeine (nan%)	0	TN	TN	FP	TN	TN	FP
15501	ANPP (0.5%) + Fentanyl (12.92%) + Fluorofentanyl (0.55%) + Bromazolam (0.43%) + Caffeine (nan%)	0	TN	TN	TN	TN	FP	FP
15503	Fentanyl (1.68%) + Fluorofentanyl (2.0%) + Bromazolam (5.88%) + Caffeine (nan%)	0	TN	TN	FP	TN	TN	TN
15504	Fentanyl (9.83%) + Fluorofentanyl (15.42%) + Bromazolam (5.57%) + Caffeine (nan%)	0	FP	TN	TN	TN	TN	TN
15505	Fentanyl (11.5%) + Fluorofentanyl (24.25%) + Bromazolam (3.11%) + Caffeine (nan%)	0	TN	TN	TN	FP	FP	TN
15550	Fentanyl (18.35%) + Bromazolam (8.23%) + Xylazine (15.3%) + Flubromazolam (1.94%) + Caffeine (nan%)	15.3	TP	TP	TP	TP	FN	TP
15617	Fentanyl (58.05%) + Bromazolam (nan%) + Xylazine (15.36%) + ANPP (0.97%) + Flualprazolam (1.9%)	15.36	TP	FN	FN	TP	TP	TP

Table 4.11: Complete breakdown of validation set ($n = 50$) sample composition as identified by paper spray mass spectrometry and benzodiazepine immunoassay test strips. Concentration values of xylazine (blue) are described next to true positive (TP), true negative (TN), false positive (FP), and false negative (FN) prediction results of the hybrid and high level SERS–IR data fusion models. Correct predictions by the respective RF, SVM, and KNN models are shown in white, while incorrect predictions are in grey.

portable spectroscopic devices are utilized to their full potential, ultimately leading to more reliable and comprehensive compound identification in various settings.

Chapter 5

Conclusions

5.1 Summary of Work

Community drug checking is an essential harm reduction strategy to address the overdose crisis and the challenges of an unregulated drug market. As opioid samples have undergone dramatic changes in primary active ingredients and potent adulterants over the years, analytical techniques must evolve accordingly. Research employing signal amplification techniques, in addition to the traditional IR and Raman spectrometers currently used in drug checking, enhances the detection and differentiation of constantly evolving adulterants.

Opioid samples were collected by the Vancouver Island Drug Checking Project, where sample composition was identified by mass spectrometry and benzodiazepine immunoassay test strips. Among the 218 powdered samples analyzed, fentanyl, fluorofentanyl, or heroin was identified as the primary active, with various cutting agents and adulterants. These samples were measured using SERS and FTIR spectroscopy and labeled according to confirmatory testing.

A subset of 168 complex illicit opioid samples was used for model development, while 50 samples with even representation of bromazolam and xylazine were used as the validation set. Both SERS and IR real world training libraries were used to develop the random forest binary classification models and were optimized through a hyperparameter tuning method using a 5-fold cross-validation grid search for the detection of bromazolam and xylazine. A visualization strategy using RF feature importance bands identified adulterant-specific SERS spectral regions, which improved classification accuracy of low confidence samples.

The comprehensive optimization of RF parameters and spectral preprocessing led to robust SERS models that outperformed IR models, particularly in detecting trace xylazine. The predictive success on the validation set highlights the importance of real-world samples for model training and demonstrates the superior sensitivity of SERS in detecting low-concentration adulterants. The final SERS models showed high sensitivity (88% and 92%) and specificity (88% and 96%) for bromazolam and xylazine detection. This approach provides a valuable, affordable, and accessible tool for standardizing spectral analysis in point-of-care drug checking services, demonstrating a robust prediction range for identifying bromazolam (0.32–36% w/w) and xylazine (0.15–15% w/w) in complex illicit drug matrices.

Given that complex opioid samples tested by Substance found an average of more than 4 compounds present [18], maintaining broad spectrum analysis and incorporating IR spectroscopy remains desirable for drug checking. The current widespread use of IR for identifying bulk cutting agents, actives, and adulterants over the limit of detection, led to an exploration into model development using SERS–IR data fusion strategies.

Data fusion of SERS and IR measurements using an early (hybrid) and late (high level) fusion technique was used to assess overall model prediction results for the detection of xylazine. SERS, IR, and concatenated SERS–IR spectral data of the training library samples were used to construct and evaluate three different classification models. A comprehensive investigation of RF, SVM, and KNN hyperparamters via grid search with spectra specific preprocessing techniques led to the construction of the 3 hybrid models and 3 high level models (fused with optimized SERS and IR weights). Using the same real world samples for training models, data fusion techniques saw increased predictive performance than what was established for IR alone. However, RF continued to demonstrate superior prediction capabilities for either hybrid or high level fusion methods, maintaining sensitivity (88% and 84%) and specificity (88% and 92%). High level data fusion using a random forest classification tool effectively prioritized SERS predictive information at 90%, demonstrating no reduction in model performance when incorporating IR data.

5.2 Future Work

Continued development of models using SERS and the integration of multiple instruments is essential for enhancing adulterant detection and overall compound identification. A simple and effective approach to begin with would be expanding the training group with diverse opioid samples, which can improve model building and standardize sample information to reflect the current drug market. Repeating the study in Chapter 4 with an increased number of xylazine-positive samples would establish a more balanced dataset, potentially reducing the high bias in SVM and KNN models.

Further practical work should include expanding research into supervised models and hyperparameter tuning strategies to investigate more optimized building metrics. Specifically, exploring neural networks, partial least squares discriminant analysis (PLS-DA), and other supervised models for better trace detection and evaluating model compatibility with SERS and IR data. Employing PCA and other dimensionality reduction techniques for inputting specific component features (mid-level fusion) could improve training data sets and reduce computational intensity in fused methods. Therefore, evaluating PCA-KNN or PCA-SVM models to enhance computational efficiency is a promising strategy for future development.

Data fusion strategies should also consider model fusion, incorporating high performing classifiers into a single late fusion strategy for difficult to identify components in complex matrices. A fusion technique with RF SERS and SVM IR outputs, given their individual performance on their data sets, warrants further investigation.

Finally, the success of the SERS technique for identifying trace adulterants heavily prompts future work into a combined fusion approach with bulk Raman measurements. As it would be more feasible to implement into drug checking operations, maximizing the information from a single instrument and reducing the financial burden of additional equipment.

References

[1] BC Coroners Service, “Illicit Drug Toxicity Deaths in BC”, Technical Report, Ministry of Public Safety & Soliciter General, Victoria, BC, 2023.

[2] Larnder, A.; Saatchi, A.; Borden, S. A.; Moa, B.; Gill, C. G.; Wallace, B.; Hore, D. *Drug Alcohol Depend.* **2022,** *235,* 109427.

[3] Bhuiyan, I.; Tobias, S.; Ti, L. *Am. J. Drug Alcohol Abuse* **2023,** *49,* 685–690.

[4] Beaulac, M.; Richardson, L.; Tobias, S.; Lysyshyn, M.; Grant, C.; Ti, L. *Int. J. Drug Policy* **2022,** *105,* 103707.

[5] Misailidi, N.; Papoutsis, I.; Nikalaou, P.; Dona, A.; Spiliopoulou, C.; Athanaselis, S. *Forensic Toxicol.* **2018,** *36,* 12–32.

[6] Borden, S.; Mercer, S. R.; Saatchi, A.; Wong, E.; Stefan, C. M.; Wiebe, H.; Hore, D. K.; Wallace, B.; Gill, C. G. *Drug Test Anal.* **2023,** *15,* 484–494.

[7] Kim, H. K.; Nelson, L. S. *Expert Opin. Drug Saf.* **2015,** *14,* 1137–1146.

[8] So, R.; Hamarneh, Y. A.; Barnes, M.; Beazely, M. A.; Boivin, M.; Laroche, J.; Patel, H.; Sihota, A.; Smith, T.; Tsuyuki, R. T. *Can. Pharm. J* **2020,** *153,* 352–356.

[9] Wermeling, D. P. *Ther. Adv. Drug Saf.* **2015,** *6,* 20–31.

[10] Measham, F. *Br. J. Clin. Pharmacol.* **2020,** *86,* 420–428.

[11] Laing, M. K.; Tupper, K. W.; Fairbairn, N. *Int. J. Drug Policy* **2018,** *62,* 59–66.

[12] Bardwell, G.; Kerr, T. *Subst. Abuse Treat. Prev. Policy* **2018,** *13,* 20.

[13] Wallace, B.; van Roode, T.; Burek, P.; Pauly, B.; Hore, D. *Drugs Education Prevention Policy* **2023,** *30,* 443–452.

[14] Laing, M. K.; Ti, L.; Marmel, A.; Tobias, S.; Shapiro, A. M.; Laing, R.; Lysyshyn, M.; Socias, M. E. *Int. J. Drug Policy* **2021,** *1,* 103169.

[15] Vahedi, H. S. M.; Hajebi, H.; Vahidi, E.; Nejati, A.; Saeedi, M. *World J. Emerg. Med.* **2019,** *10,* 27–32.

[16] Russell, C.; Law, J.; Bonn, M.; Rehm, J.; Ali, F. *Int. J. Drug Policy* **2023,** *111,* 103933.

[17] Johnson, C. F.; Barnsdale, L. R.; McAuley, A. "Investigating the Role of Benzodiazepines in Drug-Related Mortality - A Systematic Review Undertaken on Behalf of The Scottish National Forum on Drug-Related Deaths", Technical Report, NHS Health Scotland, 1 South Gyle Crescent, Edinburgh EH12 9EB, 2016.

[18] "Substance Drug Checking: Annual Review 2023", `https://substance.uvic.ca/files/reports/Substance%20Drug%20Checking%202023%20Annual%20Report.pdf`, Accessed: 2024-07-01.

[19] Edinoff, A. N.; Nix, C. A.; Odisho, A. S.; Babin, C. P.; Derouen, A. G.; Lutfallah, S. C.; Cornett, E. M.; Murnane, K. S.; Kaye, A. M.; Kaye, A. D. *Neurol. Int.* **2022,** *14,* 648–663.

[20] Hikin, L. J.; Coombes, G.; Rice-Davies, K.; Couchman, L.; Smith, P. R.; Morley, S. R. *Forensic Sci. Int.* **2024,** *354,* 111891.

[21] Rodda, L. N. *Addict.* **2024,** *1,* 1487–1490.

[22] O'Neil, J.; Kovach, S. *N. Engl. J. Med.* **2023,** *388,* 2274.

[23] Sue, K. L.; Hawk, K. *Addiction* **2024,** *119,* 606–608.

[24] Kacinko, S. L.; Mohr, A. L. A.; Logan, B. K.; Barbieri, E. J. *J. Anal. Toxicol.* **2022,** *46,* 911–917.

[25] Tobias, S.; Shapiro, A. M.; Wu, H.; Ti, L. *Can. J. Addiction* **2020,** *11,* 28–32.

[26] Bowles, J. M.; McDonald, K.; Maghsoudi, N.; Thompson, H.; Stefan, C.; Beriault, D. R.; Delaney, S.; Wong, E.; Werb, D. *Harm Reduct. J.* **2021,** *18,* 104.

[27] Torruella, R. A. *Subst. Abuse Treat. Prev. Policy* **2011,** *6,* 7.

[28] Shapiro, A.; Sim, D.; Wu, H.; Mogg, M.; Tobias, S.; Patel, P.; Ti, L. "Detection of Etizolam, Flualprazolam, and Flubromazolam by Benzodiazepine-Specific Lateral Flow Immunoassay Test Strips", Technical Report, BC Centre on Substance Use, 2020.

[29] Green, T. C.; Park, J. N.; Gilbert, M.; McKenzie, M.; Struth, E.; Lucas, R.; Clarke, W.; Sherman, S. G. *Int. J. Drug Policy* **2020,** *77,* 102661.

[30] Gozdzialski, L.; Aasen, J.; Larnder, A.; Ramsay, M.; Borden, S. A.; Saatchi, A.; Gill, C. G.; Wallace, B.; Hore, D. K. *Int. J. Drug Policy* **2021,** *97,* 103409.

[31] Miskulin, A.; Wallace, B.; Gill, C.; Hore, D. *Drug Test. Anal.* **2024,** *online ahead of print,* DOI 10.1002/dta.3630.

[32] Vandergrift, G. W.; Gill, C. G. *J. Mass Spectrom.* **2019,** *54,* 729–737.

[33] BTNX, "Harm Reduction", 2024.

[34] Tsai, J. S.-C.; Lin, G. L. Drug-Testing Technologies and Applications. In *Drugs of Abuse: Body Fluid Testing*; Wong, R. C.; Tse, H. Y., Eds.; Humana Press: Totowa, NJ, 2005.

[35] Thompson, H.; McDonald, K. *Int. J. Environ. Res. Public Health* **2023,** *20,* 6486.

[36] Moon, K. J.; Whitehead, H. D.; Trinh, A.; Hasenstab, K. A.; Hayes, K. L.; Stanley, D.; Carter, B.; Barclay, R.; Lieberman, M.; Nawaz, S. *Harm Reduction J.* **2024,** *21,* 11.

[37] Carver, H.; Falzon, D.; Masterton, W.; Wallace, B.; Aston, E. V.; Measham, F.; Hunter, C.; Sumnall, H.; Gittins, R.; Raeburn, F.; Craik, V.; Priyadarshi, S.; Rothney, L.; Weir, K.; Parkes, T. *Harm Reduct. J* **2023,** *20,* 94.

[38] Gozdzialski, L.; Hutchison, A.; Wallace, B.; Gill, C.; Hore, D. *Drug Test. Anal.* **2024,** *16,* 83–92.

[39] Gozdzialski, L.; Wallace, B.; Hore, D. *Harm Reduction J.* **2023,** *20,* 39.

[40] Wallace, B. *et al. Drug Test. Anal.* **2021,** *13,* 734–746.

[41] Bunaciu, A. A.; Aboul-Enein, H. Y.; Fleschin, S. *Appl. Spectrosc. Rev.* **2010,** *45,* 206–219.

[42] Tobias, S.; Shapiro, A. M.; Grant, C. J.; Patel, P.; Lysychyn, M.; Ti, L. *Drug Alcohol Depend.* **2021,** *218,* 108300.

[43] Johnson, C. S.; Stansfield, C. R.; Hassan, V. R. *Forensic Sci. Int.* **2020,** *313,* 110367.

[44] Gozdzialski, L.; Ramsay, M.; Larnder, A.; Wallace, B.; Hore, D. K. *J. Raman Spectrosc.* **2021,** *52,* 1308–1316.

[45] Gerace, E.; Seganti, F.; Luciano, C.; Lombardo, T.; Di Corcia, D.; Teifel, H.; Vincenti, M.; Salomone, A. *Drug Alcohol Rev.* **2019,** *38,* 50–56.

[46] Mullin, A.; Scott, M.; Vaccaro, G.; Gittins, R.; Ferla, S.; Schifano, F.; Guirguis, A. *Int. J. Environ. Res. Public Health* **2023,** *20,* 4793.

[47] Vankeirsbilck, T.; Vercauteren, A.; Baeyens, W.; der Weken, G. V.; Verpoort, F.; Vergote, G.; Remon, J. P. *Trends Anal. Chem.* **2002,** *21,* 869–877.

[48] Lanzarotta, A.; Witkowski, M.; Batson, J. *J. Forensic Sci.* **2020,** *65,* 421–427.

[49] Mabbott, S.; Alharbi, O.; Groves, K.; Goodacre, R. *Analyst* **2015,** *140,* 4399–4406.

[50] Burr, D. S.; Fatigante, W. L.; Lartey, J. A.; Jang, W.; Stelmack, A. R.; McClurg, N. W.; Standard, J. M.; Wieland, J. R.; Kim, J.-H.; Mulligan, C. C.; Driskell, J. D. *Anal. Chem.* **2020,** *92,* 6676–6683.

[51] Masterson, A. N.; Hati, S.; Ren, G.; Liyanage, T.; Manicke, N. E.; Goodpaster, J. V.; Sardar, R. *Anal. Chem.* **2021,** *93,* 2578–2588.

[52] Sha, X.; Fang, G.; Cao, G.; Li, S.; Hasi, W.; Han, S. *Analyst* **2022,** *147,* 5785–5795.

[53] Segawa, H.; Fukuoka, T.; Itoh, T.; Imai, Y.; Iwata, Y. T.; Yamamuro, T.; Kuwayama, K.; Tsijikawa, K.; Kanamori, T.; Inoue, H. *Analyst* **2019,** *144,* 2158–2165.

[54] Shende, C.; Brouillette, C.; Farquharson, S. *Analyst* **2019,** *144,* 5449–5454.

[55] Haddad, A.; Comanescu, M. A.; Green, O.; Kubic, T. A.; Lombardi, J. R. *Anal. Chem.* **2018,** *90,* 12678–12685.

[56] Wang, H.; Xue, Z.; Wu, Y.; Gilmore, J.; Wang, L.; Fabris, L. *Anal. Chem.* **2021,** *93,* 9373–9382.

[57] Kimani, M. M.; Lanzarotta, A.; Batson, J. S. *J. Forensic Sci.* **2020,** *66,* 491–504.

[58] Leonard, J.; Haddad, A.; Green, O.; Birke, R. L.; fand A. Kocka, T. K.; Lombardi, J. R. *J. Raman Spectrosc.* **2017,** *48,* 1323–1329.

[59] Gozdzialski, L.; Rowley, A.; Borden, S.; Saatchi, A.; Gill, C. G.; Wallace, B.; Hore, D. K. *Int. J. Drug Policy* **2022,** *102,* 103611.

[60] Pilot, R.; Signorini, R.; Durante, C.; Orian, L.; Bhamidipati, M.; Fabris, L. *Biosensors* **2019,** *9,* 57.

[61] Yaffe, N. R.; Ingram, A.; Graham, D.; Blanch, E. W. *J. Raman Spectrosc.* **2010,** *41,* 618–623.

[62] Fularz, A.; Almohammed, S.; Rice, J. H. *ACS Sustainable Chem. Eng..* **2021,** *9,* 16808–16819.

[63] Karawdeniya, B. I.; Bandara, Y. M. N. D. Y.; Whelan, J. C.; Dwyer, J. R. *ACS Appl. Nano Mater.* **2018,** *1,* 960–968.

[64] Han, Y.; Wu, S.-R.; Tian, X.-D.; Zhang, Y. *ACS Appl. Mater. Interfaces* **2020,** *12,* 28965–28974.

[65] Zhang, C.; Chen, S.; Jiang, Z.; Shi, Z.; Wang, J.; Du, L. *ACS Appl. Mater. Interfaces* **2021,** *13,* 29222–29229.

[66] Brulé, T.; Bouhelier, A.; Dereux, A.; Finot, E. *ACS Sens.* **2016,** *1,* 676–680.

[67] Pavan Kumar, G. V.; Ashok Reddy, B. A.; Arif, M.; Kundu, T. K.; Narayana, C. *J. Phys. Chem. B* **2006,** *110,* 16787–16792.

[68] Lai, H.; Z.Yu,; Li, G.; Zhang, Z. *J. Chromatogr. A* **2022,** *1675,* 463181.

[69] Hu, W.; Ye, S.; Zhang, Y.; Li, T.; Zhang, G.; Luo, Y.; Mukamel, S.; Jiang, J. *J. Phys. Chem. Lett.* **2019,** *10,* 6026–6031.

[70] Lee, S.; Oh, J.; Lee, K.; Cho, M.; Paulson, B.; Kim, J. K. *Anal. Chem.* **2022,** *94,* 17477–17484.

[71] Luo, S.-H.; Wang, W.-L.; Zhou, Z.-F.; Xie, Y.; Ren, B.; Liu, G.-K.; Tian, Z.-Q. *Anal. Chem.* **2022,** *94,* 10151–10158.

[72] Mirsafavi, R.; Moskovits, M.; Meinhart, C. *Analyst* **2020,** *145,* 3440–3446.

[73] Wang, L.; Vendrell-Dones, M. O.; Deriu, C.; Dogruuer, S.; Harrington, P. B.; McCord, B. *Appl. Spectrosc.* **2021,** *75,* 1225–1236.

[74] dos Santos, D. P.; Sena, M. M.; Almeida, M. R.; Mazali, I. O.; Olivieri, A. C.; Villa, J. E. L. *Anal. Bioanal. Chem.* **2023,** *415,* 3945–3966.

[75] Guo, M.; Li, M.; Fu, H.; Zhang, Y.; Chen, T.; Tang, H.; Zhang, T.; Li, H. *Spectrochim. Acta, Part A* **2023,** *287,* 122057.

[76] Seifert, S. *Sci. Rep.* **2020,** *10,* 5436.

[77] Weng, S.; Qiu, M.; Dong, R.; Wang, F.; Huang, L.; Zhang, D.; Zhao, J. *Spectrochim. Acta, Part A* **2018,** *200,* 20–25.

[78] Breiman, L. *Mach. Learn.* **2001,** *45,* 5–32.

[79] Chauhan, V. K.; Dahiya, K.; Sharma, A. *Artif. Intell. Rev.* **2019,** *52,* 803–855.

[80] Sha'abani, M. N. A. H.; Fuad, N.; Jamal, N.; Ismail, M. F. KNN and SVM Classification for EEG: A Review. In ; Nasir, A. N. K.; Ahmad, M. A.; Najib, M. S.; Wahab, Y. A.; Othman, N. A.; Ghani, N. M. A.; Irawan, A.; Khatun, S.; Ismail, R. M. T. R.; Saari, M. M.; Daud, M. R.; Faudzi, A. A. M., Eds.; Springer Nature: Singapore, 2020.

[81] Xing, W.; Bei, Y. *IEEE access* **2020,** *8,* 28808–28819.

[82] Cooman, T.; Trejos, T.; Romero, A.; Arroyo, L. *Chem. Phys. Lett.* **2022,** *787,* 139283.

[83] Wallace, B.; Gozdzialski, L.; Qbaich, A.; Azam, S.; Burek, P.; Hutchison, A.; Teal, T.; Louw, R.; Kielty, C.; Robinson, D.; Moa, B.; Storey, M.-A.; Gill, C.; Hore, D. *Drugs Habits and Social Policy* **2022,** *23,* 220–231.

[84] Smith, M.; Logan, M.; Bazley, M.; Blanchfield, J.; Stokes, R.; Blanco, A.; McGee, R. *J. Forensic Sci.* **2020,** *66,* 505–519.

[85] Borden, S. A.; Saatchi, A.; Vandergrift, G. W.; Palaty, J.; Lyshyshyn, M.; Gill, C. G. *Drug Alcohol Rev.* **2022,** *41,* 410–418.

[86] Borden, S. A.; Saatchi, A.; Krogh, E. T.; Gill, C. G. *Anal. Sci. Adv.* **2020,** *1,* 97–108.

[87] Pedregosa, F. *et al. J. Mach. Learn. Res.* **2011,** *12,* 2825–2830.

[88] Weng, S.; Dong, R.; Zhu, Z.; Zhang, D.; Zhao, J.; L.Huang,; Liang, D. *Spectrochim. Acta A: Mol. Biomol. Spectrosc.* **2018,** *189,* 1–7.

[89] Li, B.; Wu, Y.; Wang, Z.; Xing, M.; Xu, W.; Zhu, Y.; Du, P.; Wang, X.; Yang, H. *Anal. Methods* **2021,** *13,* 5264–5273.

[90] Huang, M.-W.; Chen, C.-W.; Lin, W.-C.; Ke, S.-W.; Tsai, C.-F. *PLoS One* **2017,** *12,* 161501.

[91] Kazemzadeh, M.; Martinez-Calderon, M.; Xu, W.; Chamley, L. W.; Hisey, C. L.; Broderick, N. G. R. *Anal. Chem.* **2022,** *94,* 12907–12918.

[92] Syriopoulos, P. K.; Kalampalikis, N. G.; Kotsiantis, S. B.; Vrahatis, M. N. *Ann. Math. Artif. Intell.* **2023,** *1,* 1.

[93] Zhang, M.-L.; Zhou, Z.-H. *Pattern Recognit.* **2007,** *40,* 2038–2048.

[94] Othman, N. H.; Lee, K. Y.; Radzol, A. R. M.; Mansor, W.; Wong, P. S.; Looi, I. *Int. J. Intell. Inf. Database Syst.* **2018,** *10752,* 335–346.

[95] Adade, S. Y.-S. S.; Lin, H.; Johnson, N. A. N.; Qianqian, S.; Nunekpeku, X.; Ahmad, W.; Kwadzokpui, B. A.; Ekumah, J.-N.; Chen, Q. *Food Chem.* **2024,** *454,* 139836.

[96] Raiaan, M. A. K.; Sakib, S.; Fahad, N. M.; Mamun, A. A.; Rahman, A.; Shatabda, S.; Mukta, S. H. *Decis. Anal. J.* **2024,** *11,* 100470.

[97] Karthika, M. S.; Rajaguru, H.; Nair, A. R. *Bioeng. (Basel)* **2024,** *11,* 314.

[98] Zhang, X.; Chang, Z.; Li, Y.; Wu, H.; Tan, J.; Li, F.; Cui, B. *Proc. VLDB Endow.* **2022,** *15,* 1808–1821.

[99] Li, Y.; Shen, Y.; Jiang, H.; Zhang, W.; Li, J.; Liu, J.; Zhang, C.; Cui, B. *Proc. VLDB Endow.* **2022,** *15,* 1256–1265.

[100] Xionga, Z.; Cuia, Y.; Liua, Z.; Zhaoa, Y.; Hub, M.; Hu, J. *Comput. Mater. Sci.* **2020,** *171,* 109203.

[101] Wagmann, L.; Manier, S. K.; Felske, C.; Gampfer, T. M.; Richter, M. J.; Eckstein, N.; Meyer, M. R. *J. Anal. Toxicol.* **2021,** *45,* 1014–1027.

[102] Harkins, W. D.; Bowers, H. E. *Phys. Rev.* **1931,** *38,* 1845–1857.

[103] Cleaver, K. D.; Davies, J. E. D.; Wood, W. J. *J. Mol. Struct.* **1975,** *25,* 222–224.

[104] Kimani, M. M.; Smith, S. W.; Lanzarotta, A.; Brueggemeyer, J. L.; Batson, J. S. *Forensic Sci. Int.* **2022,** *338,* 111390.

[105] Doctor, E. L. *Development of a Surface-Enhanced Raman Spectroscopy Method for the Detection of Benzodiazepines in Urine,* Thesis, Florida International University, 2014.

[106] Neville, G. A.; Shurvell, H. F. *J. Raman Spectrosc.* **1990,** *21,* 9–19.

[107] Neville, G. A.; Beckstead, H. D.; Shurvell, H. F. *Vib. Spectrosc.* **1991,** *1,* 287–297.

[108] Anban, J. D.; James, C.; Kumar, J. S.; Pradhan, S. *SN Appl. Sci.* **2020,** *2,* 1685.

[109] Ramsay, M.; Gozdzialski, L.; Larnder, A.; Wallace, B.; Hore, D. K. *Vib. Spectrosc.* **2021,** *114,* 103243.

[110] Ahmmed, F.; Fuller, I. D.; Killeen, D. P.; Fraser-Miller, S. J.; Gordon, K. C. *ACS Food Sci. Technol.* **2021,** *1,* 570–578.

[111] Wang, H.; Barone, G.; Smith, A. *Struct. Infrastruct. Eng.* **2023,** *online ahead of print*,.

[112] Borràs, E.; Ferré, J.; Boqué, R.; Mestres, M.; Aceña, L.; Busto, O. *Anal. Chim. Acta* **2015,** *891,* 1–14.

[113] Stahlschmidt, S. R.; Ulfenborg, B.; Synnergren, J. *Briefings Bioinf.* **2022,** *23,* 1–15.

[114] Li, Y.; Huang, Y.; Xia, J.; Xiong, Y.; Min, S. *Vib. Spectrosc.* **2020,** *108,* 103060.

[115] Gao, W.; Han, Y.; Chen, L.; Tan, X.; Liu, J.; Xie, J.; Li, B.; Zhao, H.; Yu, S.; Tu, H.; Feng, B.; Yang, F. *Analyst* **2023,** *148,* 5650–5657.

[116] Du, J.; Huang, Z.; Li, C.; Jiang, L. *RSC Adv.* **2024,** *14,* 12428–12437.

[117] Li, Q.; Huang, Y.; Zhang, J.; Min, S. *Spectrochim. Acta, Part A* **2021,** *247,* 119119.

[118] Li, M.; Xue, J.; Du, Y.; Zhang, T.; Li, H. *Energy Fuels* **2019,** *33,* 12286–12294.

[119] Maillo, J.; Ramìrez, S.; Triguero, I.; Herrera, F. *Knowledge Based Syst.* **2017,** *117,* 3–15.