Validation of Fast Spectrochemical Screening Methods for the Identification of Counterfeit Pharmaceutical Packaging

Emily Ann Haase
West Virginia University, eah0033@mix.wvu.edu

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Part of the Forensic Science and Technology Commons

Recommended Citation
https://researchrepository.wvu.edu/etd/7549

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.
Validation of Fast Spectrochemical Screening Methods for the Identification of Counterfeit Pharmaceutical Packaging

Emily Ann Haase

Thesis submitted
to the Eberly College of Arts and Science
at West Virginia University

in partial fulfillment of the requirements for the degree of

Master of Science in
Forensic and Investigative Science

Luis Arroyo, Ph.D., Chair
Tatiana Trejos, Ph.D.
Marina Galvez, PharmD, PhD

Department of Forensic and Investigative Science

Morgantown, West Virginia
2020

Keywords:
Counterfeit pharmaceuticals, LIBS, ATR-FTIR, pharmaceutical packaging, forensics

Copyright 2018 Emily Haase
ABSTRACT

Validation of Fast Spectrochemical Screening Methods for the Identification of Counterfeit Pharmaceutical Packaging

Emily Ann Haase

Counterfeit pharmaceuticals are an actively developing health and economic threat worldwide. Particularly prevalent are counterfeit pharmaceuticals distributed in emerging nations and through internet pharmacies or e-pharmacies. Although technology has been developed that discourages anti-counterfeiting practices (such as optically variable devices, invisible ink, and track-and-trace technology), it remains somewhat novel and expensive to implement on a widespread scale.

In this study, Laser Induced Breakdown Spectroscopy (LIBS) and Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy (ATR-FTIR) were proposed as fast and non-invasive tools for the identification of counterfeit pharmaceutical packages. The main objective of this research was to develop and evaluate the capabilities of LIBS and ATR-FTIR to determine chemical differences between counterfeit and authentic pharmaceutical packaging samples. LIBS and ATR-FTIR possess several characteristics that render them suitable for rapid on-site detection. They produce analytical results in less than one minute per sample, with high sensitivity and selectivity, limited sample preparation, and minimal destructivity.

The methods were evaluated through the analysis of a dataset of 166 packages (112 counterfeits and 54 authentic sources). The dataset was divided into two main subsets. The first subset was evaluated to identify the informative value of LIBS for fast screening of black barcodes and the carton substrate (100 counterfeit and 35 authentic). The multi-color inks and paper of the second subset was investigated for variation of chemical profiles within and between sources, and the method’s capabilities to distinguish between counterfeits (112) and authentic samples (12).

One hundred and twelve counterfeit pharmaceutical cartons were printed from five different sources, mimicking six authentic counterparts. The authentic subset consisted of twelve
secondary packages of six common medical products, including packages from the same and different manufacturing lots. The selected products consisted of vasodilators, antivirals, steroids, and other commonly counterfeited pharmaceuticals.

Intra-source variation of the counterfeit subset was investigated; it was determined to be sufficiently lower than inter-source variation. False exclusion rates were calculated to be less than 20% for samples originating from the same source (e.g., same package, intra-lots, replicate printouts).

Using LIBS, a two-class classification system was used for the combined black barcode ink and paperboard carton spectra (n = 135, 100 counterfeit, 35 authentic packages). As black barcode ink is very common on pharmaceutical packaging, this system was used as a general screening technique to quickly identify a sample as authentic or counterfeit, regardless of counterfeit printing source. In general, the correct classification rates for this set were over 92%.

The classification models were established using six machine learning methods: Random Forest, Naïve Bayes, Neural Networks, k-Nearest Neighbors, Quadratic Discriminant Analysis, and Linear Discriminant Analysis. A random split of 60% and 40% of the dataset was applied for training and testing of the classifier algorithms. Principal Component Analysis (PCA) was utilized on the LIBS and ATR-FTIR data for variable reduction purposes. The principal components for each ink type were combined prior to classification.

Also, a six-class system was also used to classify the dataset using LIBS, ATR-FTIR, and combined data from both techniques (n = 124, 112 counterfeit, 12 authentic packages). The machine learning methods classified the samples as belonging to one of five counterfeit printing sources or their corresponding authentic counterpart. Seven ink colors (red, blue, yellow, green, brown, pink, black) were analyzed; additionally, in ATR-FTIR, the paperboard substrate was also analyzed. In most comparisons, LIBS had a successful classification rate of over 70% and ATR-FTIR had a successful classification rate of over 85%. When the data from both techniques were combined, the discrimination power of the system increased to 93% correct classification. Although LIBS and ATR-FTIR had a low misclassification rate when used in isolation, the misclassification rate could be reduced even further through data combination.
The results of this study are encouraging for the inclusion of LIBS and ATR-FTIR as a screening method for the detection of counterfeit pharmaceutical packaging. The utilization of combined data to discover chemical signatures addresses an urgent need in the investigation of counterfeit pharmaceuticals. Also, the classification of counterfeit samples into their specific counterfeit source may benefit investigators as they make determinations in the counterfeit pharmaceutical packaging supply chain. This study is anticipated to offer relevant tools to both government and pharmaceutical industry in the detection and fight against counterfeit pharmaceuticals.
Acknowledgements

All the thanks in the world to both the Trejos and Arroyo lab groups for all their assistance and support on this project! Additional thanks to Dr. Rafael Ortiz and Maira Kerpel Dos Santos from the Porto Alegre Federal Police, Dr. Renata Pereira-Limberger from the Federal University of Rio Grande do Sul (UFRGS), and Dr. Galvez from WVU for project assistance and sample acquisition.

Thank you to Dr. Trejos and Dr. Arroyo for all of your help with the ups and downs of this project, as well as my undergraduate student Mandy for all her work with lab analysis!
# Table of Contents

ABSTRACT .......................................................................................................................... ii

Acknowledgements .............................................................................................................. v

1. Introduction ...................................................................................................................... 1

   1.1. The Impact of Counterfeit Pharmaceuticals to Human Health and Economy .......... 1

   1.3. Ink Composition ........................................................................................................ 10

   1.4. Instrumental Techniques ......................................................................................... 16

       1.4.1. Laser Induced Breakdown Spectroscopy (LIBS) .............................................. 16

       1.4.2. ATR-FTIR ........................................................................................................ 21

   1.5. Machine Learning Techniques ................................................................................ 26

       1.5.2. Random Forest (RF) ....................................................................................... 27

       1.5.3. Naïve Bayes (NB) .......................................................................................... 28

       1.5.4. Neural Network (NN) ..................................................................................... 30

       1.5.5. k-Nearest Neighbors (KNN) .......................................................................... 31

       1.5.6. Linear Discriminant Analysis (LDA) .............................................................. 33

       1.5.7. Quadratic Discriminant Analysis (QDA) ......................................................... 35

2. Objectives of the Project ............................................................................................... 38

   2.1. Motivation of the Study and Objective Summary .................................................... 38

   2.2. Task 1. Create and visualize a dataset of pharmaceutical packages to assess the utility of LIBS and FTIR for identification of counterfeits ......................................................... 40

   2.3. Task 2. Optimize and validate Laser-Induced Breakdown Spectroscopy for the analysis and classification of authentic and counterfeit pharmaceutical packaging ......................... 42

   2.4. Task 3. Validate Attenuated Total Reflectance – Fourier Transform Infrared Spectroscopy for the analysis and classification of authentic and counterfeit pharmaceutical packaging. .... 44

   2.5. Task 4. Combine LIBS and ATR-FTIR data to classify authentic and counterfeit pharmaceutical packaging using a more comprehensive chemical profile ..................................... 45
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.</td>
<td>Dataset Creation and Visualization</td>
<td>46</td>
</tr>
<tr>
<td>3.1.</td>
<td>Methodology</td>
<td>46</td>
</tr>
<tr>
<td>3.2.</td>
<td>Results and Discussion</td>
<td>50</td>
</tr>
<tr>
<td>3.3.</td>
<td>Conclusion</td>
<td>57</td>
</tr>
<tr>
<td>4.</td>
<td>Method Optimization and Validation for Laser-Induced Breakdown Spectroscopy</td>
<td>58</td>
</tr>
<tr>
<td>4.1.</td>
<td>Methodology</td>
<td>58</td>
</tr>
<tr>
<td>4.2.</td>
<td>Results and Discussion</td>
<td>63</td>
</tr>
<tr>
<td>4.3.</td>
<td>Conclusion</td>
<td>83</td>
</tr>
<tr>
<td>5.</td>
<td>Method Validation of Attenuated Total Reflectance – Fourier Transform Infrared Spectroscopy</td>
<td>84</td>
</tr>
<tr>
<td>5.1.</td>
<td>Methodology</td>
<td>84</td>
</tr>
<tr>
<td>5.2.</td>
<td>Results and Discussion</td>
<td>85</td>
</tr>
<tr>
<td>5.3.</td>
<td>Conclusion</td>
<td>91</td>
</tr>
<tr>
<td>6.</td>
<td>Combination of Laser-Induced Breakdown Spectroscopy and Attenuated Total Reflectance – Fourier Transform Infrared Spectroscopy</td>
<td>92</td>
</tr>
<tr>
<td>6.1.</td>
<td>Methodology</td>
<td>92</td>
</tr>
<tr>
<td>6.2.</td>
<td>Results and Discussion</td>
<td>93</td>
</tr>
<tr>
<td>6.3.</td>
<td>Conclusion</td>
<td>97</td>
</tr>
<tr>
<td>7.</td>
<td>Conclusion Summary</td>
<td>98</td>
</tr>
<tr>
<td>8.</td>
<td>References</td>
<td>101</td>
</tr>
<tr>
<td>8.</td>
<td>Appendix</td>
<td>117</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table 1. Example composition of inkjet inks [93, 95] ................................................................. 11
Table 2. Discrimination power of inkjet inks using different analytical techniques and elements identified through LA-ICP-MS [96, 97, 98] ............................................................................. 11
Table 3. Example composition of toner inks [92,99] ...................................................................... 12
Table 4. Discrimination power of toner inks using different analytical techniques and elements identified through LA-ICP-MS [96,97,98] ............................................................................. 13
Table 5. Example composition of offset inks [100, 101, 102, 103] ............................................. 14
Table 6. Discrimination power of offset inks using different analytical techniques and elements identified through LA-ICP-MS [96, 97, 98] ............................................................................. 14
Table 7. Common components in document paper [104, 105] .................................................... 15
Table 8. Common elements identified in document paper [105, 106] ........................................ 15
Table 9. Summary of the machine learning methods used in the study ....................................... 37
Table 10. Full list of authentic dataset information ...................................................................... 47
Table 11. Full description of counterfeit printing sources. Replicate printouts grouped as a range.............................................................................................................................. 49
Table 12. Medical purposes for authentic pharmaceutical dataset .............................................. 51
Table 13. Medical purposes for counterfeit pharmaceutical samples .......................................... 53
Table 14. Tested parameters for preliminary LIBS optimization .................................................. 59
Table 15. 27 Box Behnken experiments with tested parameters .................................................. 60
Table 16. Tested parameters for Box Behnken optimization of LIBS ......................................... 61
Table 17. Informative elements for analyzed ink colors and substrate ........................................ 69
Table 18. False exclusion rates within same-lot authentic samples ............................................. 72
Table 19. Misclassification rates for a two-class system using combined black ink and paperboard substrate LIBS data (RF: Random Forest, NB: Naïve Bayes, NN: Neural Network, KNN: k-Nearest Neighbor, QDA: Quadratic Discriminant Analysis, LDA: Linear Discriminant Analysis). ................................................................................................................................. 79
Table 20. Six-class system: misclassification rates of LIBS data using six machine learning methods .................................................................................................................. 81
Table 21. ATR-FTIR Parameters ................................................................................................. 84
Table 22. Six-class system: misclassification rates of ATR-FTIR data using six machine learning methods. .................................................................................................................................................. 89

Table 23. Misclassification rates of dataset using combined LIBS/ATR-FTIR data. ................. 95
LIST OF FIGURES

Figure 1. An example of two anti-counterfeit technologies on a pharmaceutical package: tamper evident packaging and an optically variable device (hologram) [29]................................................. 4
Figure 2. Diagram of a Q-switched Nd:YAG laser. Pumping source is a flashback lamp [109]. 18
Figure 3. Diagram of LIBS apparatus [107]. .................................................................................. 20
Figure 4. Diagram of FTIR instrument [121] ................................................................................. 22
Figure 5. Diagram of Michelson Interferometer [121] .................................................................. 23
Figure 6. Diagram of ATR accessory apparatus [118] .................................................................... 24
Figure 7. An example of a basic decision tree. .............................................................................. 27
Figure 8. An example of a basic Naive Bayes classifier ................................................................. 29
Figure 9. An example of an artificial neural network, with three hidden layers ......................... 30
Figure 10. An example of a basic KNN classifier ......................................................................... 32
Figure 11. An example of a basic LDA classifier .......................................................................... 34
Figure 12. An example of a basic QDA classifier ........................................................................ 35
Figure 13. Experimental design of study ...................................................................................... 39
Figure 14. Examples of sample identification system. ................................................................. 46
Figure 15. Description for authentic pharmaceutical dataset. Colors analyzed for intra-source variability are shown (gray = paperboard substrate). ......................................................... 53
Figure 16. Visual display of authentic reference A4 and printouts from C1-C5. Analyzed colors were blue, yellow, black (barcode), and paperboard substrate ............................................. 54
Figure 17. Description for counterfeit pharmaceutical dataset. Colors analyzed for counterfeit/authentic comparison are shown (gray = paperboard substrate). .............................. 56
Figure 18. An authentic barcode sample after eight analytical runs ............................................ 64
Figure 19. Box Behnken graph showing signal to noise of the calcium peak on Pharmaceutical B and how it varies with frequency and laser energy .................................................. 65
Figure 20. LIBS spectra displaying the black ink for authentic reference samples A1 and A2. .. 66
Figure 21. LIBS spectra indicating black ink on authentic reference samples A3 and A4. ......... 67
Figure 22. LIBS spectrum indicating black ink on authentic reference A5 .................................... 68
Figure 23. Box plots depicting intra-source variability in same-lot pharmaceutical samples: titanium in red ink (Pharmaceutical A), cadmium in blue ink (Pharmaceutical B), and calcium in paper substrate (Pharmaceutical C), ........................................................................................................... 71
Figure 24. Top: Boxplot of variance of cadmium content in the blue ink of counterfeit counterpart 1 within counterfeit printing sources. Bottom: ANOVA plot of variance of cadmium content in blue ink of counterfeit counterpart 1, C2.

Figure 25. False exclusion rates for two authentic templates (A1 and A2) from five counterfeit printing sources.

Figure 26. False exclusion rates for two authentic templates (A3 and A4) from five counterfeit printing sources.

Figure 27. False exclusion rates for two authentic templates from five (A5) or three (A6) counterfeit printing sources.

Figure 28. LDA Canonical plot for 2-class system using LIBS black ink and paperboard substrate.

Figure 29. PCA plot displaying maximized variance between the six classes, utilizing the second principal component of the green ink on A2.

Figure 30. QDA canonical plot displaying class separation in A2 for a six-class system using LIBS.

Figure 31. ATR-FTIR spectrum of black ink on A1 (bottom), as well as magnified portions of the fingerprint region of A1 - A5.

Figure 32. ATR-FTIR spectra of authentic template A1 (black), as well as a replicate printout from C4 (gray).

Figure 33. PCA plot of yellow ink for A1, depicting class separation with one feature PC1 component.

Figure 34. QDA canonical plot, displaying class separation, of A1 using ATR-FTIR data.

Figure 35. Two emission lines (Al at 396 nm and Ca at 397 nm) and magnified ATR-FTIR Spectra of blue ink on A1 and 5 corresponding counterfeit counterparts (1 (authentic class: red, counterfeit printing source 1: blue, counterfeit printing source 2: green, counterfeit printing source 3: orange, counterfeit printing source 4: purple, counterfeit printing source 5: navy).

Figure 36. LDA canonical plot, displaying class separation, of A1 using combined LIBS and ATR-FTIR data.
1. Introduction

1.1. The Impact of Counterfeit Pharmaceuticals to Human Health and Economy

The World Health Organization [WHO] estimated in 2006 that 10% of the entire global pharmaceutical trade includes the transfer of counterfeit pharmaceuticals [1]. A more accurate estimate would vary widely with location, pharmaceutical accessibility, and reporting standards. In emerging countries, up to 25% of pharmaceutical trade may be counterfeit whereas in developed countries, it may affect only 1% of the pharmaceutical trade [2]. However, the growing emergence of online ‘e-pharmacies’ presents an alternative route for the sale of counterfeit pharmaceuticals. Due to lax security screening procedures and accreditation processes, the vast majority of e-pharmacies are unlicensed and, as a result, WHO estimated that 50% [2] of pharmaceuticals sold over the Internet are counterfeit.

Although counterfeit pharmaceuticals present a looming concern for global industry, more systematic, collaborative investigations into the issue are necessary to determine more precise estimates about the scale of their impact. The problem of counterfeit pharmaceuticals has become more prevalent within the last five years, with global reports of pharmaceutical crime increasing 102.3% from 2014 (2107) to 2018 (4405), according to the Pharmaceutical Security Institute [3].

Definitions of counterfeit pharmaceuticals can differ between countries and investigating bodies. The most common definition is used by WHO, where a counterfeit pharmaceutical is described as a pharmaceutical “that has been deliberately or fraudulently mislabeled with respect to its identity and/or source.” [4] In order to encompass all definitions of counterfeit pharmaceuticals, WHO also developed an acronym to assist its investigation into fake drugs: SSFFC, or “Substandard, Spurious, Falsely Labeled, Falsified, and Counterfeit” pharmaceuticals [5]. Of these, counterfeit pharmaceuticals may include the packaging materials, the drugs (incorrect/insufficient API or additional impurities), or both entities.

The estimated global impact of counterfeit pharmaceuticals was US $ 75 billion in 2010. It is important to focus on the indirect ramifications associated with the presence of counterfeit
pharmaceuticals, and not only the direct cost [6,7]. For instance, in 2005, the economic effect of counterfeit pharmaceuticals in the US was $6 billion, representing a 92% increase in just five years [8]. This incorporates the broader consequences of this criminal enterprise, such as cost to the health care system and fewer legitimate pharmaceutical sales [9].

With no standard global reporting body taking action on this growing situation, the public health toll of SSFFC pharmaceuticals influence is more difficult to establish. For example, negative health impacts from these fake drugs are generally only reported in the news when they are, in some way, exceptional. In 2008, it was determined that a counterfeit version of heparin (blood thinner), which contained the incorrect active pharmaceutical ingredient (over sulfated chondroitin sulfate), was distributed to patients and resulted in hundreds of reports of severe consequences, including death [10]. In 1995-1996, cough syrup tainted with diethylene glycol (antifreeze) caused death by renal failure of over 80 Haitian children [11]. Anti-malarial drugs are another serious concern, particularly in emerging countries that lack a robust pharmaceutical security supply chain. Fake anti-malarial medications are estimated to attribute to 450,000 deaths annually worldwide [12]. Other estimations indicate that more than 33 - 53% of artesunate in mainland southeastern Asia is counterfeit [13]. Other prominent cases of SSFFC pharmaceuticals are Avastin (anticancer, no active ingredient, US), Viagra/Cialis (erectile dysfunction, unreported active ingredient, UK), Zidolam-N (HIV/AIDS, falsified, Kenya) [14].

Counterfeit pharmaceuticals have become a danger to public health and global economy, but few deterrents exist for counterfeiters. Several factors exist that encourage the rapid growth of fake pharmaceutical manufacturing: expensive authentic drugs, little regulatory legislation, limited legal punishments, drug agency underfunding, consumer ignorance of dangers, and little communication between drug agency bodies [11, 15].

A compounding threat is the lack of regulation on online pharmacies, or e-pharmacies. According to the Food and Drug Administration, almost 25% of all Internet users will make a purchase from e-pharmacies, resulting in a yearly sale of US $11 billion [9]. Additionally, a separate study from the National Association of Boards of Pharmacy [16] indicated that 97% of e-pharmacies are non-compliant with federal standards. Other than occasional shut-down operations made by federal or
global entities [17,18], e-pharmacies operate mostly without repercussion. Legal consequences for counterfeit pharmaceutical manufacturers remain much less drastic than illegal drug manufacturers. For example, smuggling a kilogram of cocaine may result in a decades-long prison sentence, whereas a kilogram of sildenafil (Viagra©) may carry a sentence of only three years [19]. All of these factors contribute to the evolving counterfeit pharmaceutical industry. However, in 2013, the Drug Supply Chain Security Act (Title II of the Drug Quality and Security Act) was made into law, mandating strict requirements for pharmaceutical manufacturers to prevent weaknesses in the drug supply chain [20]. Full track-and-trace technology is expected to be implemented into the drug supply chain by 2023, which will assist in preventing counterfeit manufacturers from infiltrating the drug supply chain [21].

An important aspect of current anti-counterfeit technologies in the pharmaceutical industry is the utilization of pharmaceutical packaging, which can consist of three components: primary, secondary, and tertiary packaging [22]. Primary packaging is the portion that comes into physical contact with the pharmaceutical (e.g., glass/plastic containers, blister packages). Secondary packaging encloses the primary packaging (typically paperboard cartons). Finally, tertiary packaging facilitates shipment of the pharmaceuticals, such as pallets or barrels, and is not typically seen by the end user.

Preventative measures for the manufacture of counterfeit pharmaceuticals exist at the industry level and can be incorporated on all packaging components. Anti-counterfeit technologies (ACTs) broadly fall into three different categories: tamper evident, product authentication, and track and trace technology [23]. Tamper evident packaging intends to indicate whether a pharmaceutical packaging has been prematurely opened prior to reaching the consumer, and includes shrink wrap [24], blister packs [25], and breakable ampules [26].

Product authentication involves the inclusion of extra details or features on the packaging to indicate its authenticity. This category can be further divided into covert and overt technology [27]. Pharmaceuticals with overt ACTs can have their authenticity verified by simple visual observation. These can include optical variable devices [28], color-changing [30] or complex printing patterns. An example of an optically variable device is shown in Figure 1. In contrast,
pharmaceuticals with covert ACTs require additional methods (such as an alternative light source or chemical analysis) to verify their authenticity. Covert ACTs can include digital watermarks [31], invisible ink [32], or chemical taggants [33].

![Figure 1](image)

Figure 1. An example of two anti-counterfeit technologies on a pharmaceutical package: tamper evident packaging and an optically variable device (hologram) [29].

The last category of ACTs includes “track and trace” technology, which can include traditional assigned serial numbers physically printed on the package (barcodes, drug pedigrees, mass serialization) or, more novel, radio frequency identification tags [34]. This category intends to secure the drug supply chain and prevent theft or fabrication of pharmaceuticals.

1.1.1. Common Types of Counterfeit Pharmaceuticals

It should be emphasized that all pharmaceutical products have the potential to be counterfeited, but certain factors increase the likelihood for illegal manufacture, such as brands with public familiarity or high cost. The prevalence of counterfeited pharmaceuticals, as well as trends in reporting, largely differs with geographic region and time period. For instance, in the 1980s, 78% of counterfeit drug reports submitted to the World Health Organization originated from developing countries; but this number has shifted to 60% from developing, and 40% from developed countries in 2000 [35]. These changes in reporting data does not reflect a decrease in the number of cases originating from emerging countries, but reflect the growth of counterfeit pharmaceuticals sold online in both areas, affecting the total number of reports.
Reports logged from 2009-2011 in the Counterfeit Incident System from the Pharmaceutical Security Institute (a not-for-profit organization that focuses on monitoring the prevalence of counterfeit pharmaceuticals and holds many major pharmaceutical companies as members) determined that most reports came from government health agencies (58%). The top five reporting countries were China (27.6% of reports), Peru (11.6%), Uzbekistan (10.9%), Russia (8.4%), and Ukraine (7.2%) [36].

Other studies have attempted to determine which pharmaceuticals are more likely to be counterfeited, but excluding geographic location. A review conducted by Koczwara and Dressman [2] indicated that the top 6 therapeutic drug classes collected during academic research were antimalarials, antibiotics, phosphodiesterase inhibitors (PDE-V), anthelmintic, and growth factors. The World Health Organization largely agrees with this finding based on reports received from 1999 – 2002, determining that the most commonly counterfeited drugs were: antibiotics, hormones, anti-histamines, anti-malarial, and analgesics [35]. Due to widespread access to the Internet, developed countries typically have greater instances of counterfeit medicines distributed online, such as vasodilators, opioids, and benzodiazepines. Pfizer, a prominent pharmaceutical company, found that 96% of 250 Xanax© units purchased online were counterfeit [37].

In developing countries, particularly those struggling with malaria epidemics, counterfeit antimicrobials and antibiotics are common. Kaur et al. [38] examined 10,000 artemisinin samples from six developing countries: Equatorial Guinea, Cambodia, Ghana, Nigeria, Rwanda, and Tanzania. They determined that the percentage of falsified anti-malarial ranged from 0.0% - 7.4% of the collected samples, and the percentage of substandard anti-malarial ranged from 1.6% - 37.0% of the data set studied. Furthermore, the World Health Organization estimated a staggering number of 200,000 preventable deaths associated with the use of fake antimalarial medication. Both the WHO and Interpol estimating that 30-40% of antimalarials are falsified or substandard within Africa and Southeast Asia [39].
1.1.2. Analytical Investigations into Counterfeit Pharmaceuticals

1.1.2.1. Literature Reviews Concerning the Analysis of Counterfeit Pharmaceuticals

Many studies have focused on analytical techniques to identify counterfeit pharmaceuticals. Given the complex nature of medicinal composition, there are several ways to characterize and identify counterfeit products. A review by Görög [40] studied analytical techniques effective for the identification of bulk pharmaceuticals and the API: traditional color tests/colorimetry [41, 42, 43], spectroscopy (ATR-FTIR [44, 45, 46, 47], UV [48, 49, 50], NMR [51, 52, 53], optical rotation [54]), chromatography (HPLC [55, 56, 57, 58], TLC [59, 60, 61, 62]), and electromigration (CE [63, 64]). They additionally advocate for mass spectrometry [64, 65, 66, 67] for impurity detection within the drug. One review [69] recommends the coupling of orthogonal NMR-HPLC to identify active pharmaceutical ingredients.

Almuzaini et al. [70] specifically evaluated studies concerning to the identification of counterfeit pharmaceuticals. In this review, the authors discussed major findings from fifteen articles, all of which were related to fake anti-microbials and provided proper details for replication. Most of the analytical results originated from liquid chromatography systems.

Antibacterial medicines are one of the most commonly counterfeit therapeutic classes and have been the subject of numerous review manuscripts. Almuzaini et al. [70] examined scientific articles referring to SSFFC antibacterials and other medicines. Tshilumba et al. [71] identified 41 papers focusing on this topic: 30 papers utilized liquid chromatography, 9 papers utilized UV/Vis, and 2 papers utilized both methods. They additionally found that most research focused on beta-lactams, quinolines, and macrolide compounds. Another review [72] described 24 sufficiently rigorous studies relating to the study of anti-infectives, indicating that HPLC, TLC, and colorimetry were common and effective analysis techniques.

Instead of a specific therapeutic class, other reviews have focused on a specific geographic location. Nayyar et al. [73] reviewed studies pertaining to the chemical analysis of anti-microbial drugs in Southeast Asia and sub-Saharan Africa. The authors reported the use of HPLC, TLC, LC-MS, XRD, Raman, and physical methods to authenticate anti-malarial samples. Of these, 43% of
the Southeast Asia (n=894) and 35% of the sub-Saharan Africa (n=2297) samples were determined to be SSFFCC pharmaceuticals due to the results of a chemical analysis.

Since the analysis of SSFFC products often involves more than one examination and results in multivariate data, a variety of chemometric tools have also been used to analyze counterfeit pharmaceuticals. Kumar et al. [74] reviews common chemometric tools used in forensic science, breaking them into two distinct multivariate classes: supervised (LDA, PLSDA, kNN, SIMCA, ANN, SVM) and unsupervised (PCA and cluster analysis). A portion of the review is also dedicated to analyzing counterfeit medicine studies. Two of the studies (one for various pharmaceutical tablets, other for Viagra) utilized PCA, SVM, KNN, LDA, and SIMCA.

Krakowska et al. [75] focused on the combination of analytical and chemometric techniques to identify counterfeit pharmaceuticals. Of the analytical techniques, the authors noted colorimetric, chromatographic (TLC, GC, HPLC), spectroscopic (NIR [76, 77, 78], FT-IR, ATR-FTIR, Raman [79, 80], NMR), and spectrometric (HPLC-MS) methods were used. In examining the chemometric methods, the authors divided them into unsupervised techniques (projection, clustering) and supervised techniques (discrimination, classification). The discrimination techniques discussed included classification and regression trees (CART), LDA, partial least squares-discriminant analysis (PLS-DA), k-NN (k-Nearest Neighbors), multi-layer feedforward neural networks (MLFNN) and support vector machines (SVM). The classification techniques discussed are soft independent modeling of class analogies (SIMCA), absolute centered residuals (ACR), classification and influence matrix analysis method (CAIMAN). All chemometric techniques were identified in studies pertaining to the discrimination or classification of counterfeit pharmaceuticals.

1.1.2.2. Screening Field Technologies for Counterfeit Pharmaceuticals

The availability and efficiency of screening technologies are particularly important for the identification of counterfeit pharmaceuticals. If questioned pharmaceuticals are found in bulk, or need to be analyzed on-site, screening technologies are essential to quickly make decisions about their authenticity.
One of the most prominent instruments available for the field testing of questioned pharmaceuticals is the German Pharma Health Fund (GPHF) – MiniLab. The GPHF-MiniLab utilizes four steps for the authentication of pharmaceuticals: visual inspection (drugs and packaging), drug-disintegration test, color tests (drug only), and TLC (drug only). It reports that it can perform up to 3000 color tests (identifying 80 separate drug compounds) and 1000 TLC analyses per suitcase kit [81]. Although it is somewhat cheaper than most handheld screening instruments (approximately 4000 USD), the sensitivity is reported to be much lower than, for example, a handheld Raman. While a Raman could reportedly detect whether a pharmaceutical had 95% of its specified API or less, a GPHF MiniLab could only detect whether a pharmaceutical had 80% of the same metric [35].

De Araujo et al. [82] investigated portable analytical technologies for general uses in forensic chemistry, including questioned document and drug identification. For questioned documents, they identified a study [83] which used a paperspray ionization source to date questioned documents. They also identified numerous portable screening technologies available for the detection of drugs on-site, including traditional color tests and mass spectrometry methods (such as DART-MS).

Another review [84] focused on the accessibility of screening technologies for the identification of drugs, listing the intended usage of the technology and relative cost. For screening technologies meant for the identification of major components of the pharmaceutical, they noted the capital costs for either the portable technique or instrumentation as follows: IR ($25,000 - $70,000), NIR ($7000 - $60,000), Raman ($18,000 - $60,000), XRF ($30,000 - $60,000), ion mobility spectrometry ($25,000 - $60,000) GHF minilab chromatography and disintegration ($4500 - $10,000). The review emphasized the importance of testing and validating screening technologies to assist in the global counterfeit pharmaceutical problem. Vickers et al. [85] identified multiple technologies, focusing on Raman, NIR, MIR, TLC, paper-based, capillary electrophoresis, laser absorption/fluorescence, and others. They also identified several large gaps in the literature about screening technologies for pharmaceutical identification concerning training needs, cost-effectiveness studies, and API range of the technologies.
Finally, another review focused [86] on the usage of screening technologies for antimalarials in lower middle-income countries. They identified 39 manuscripts that used either visual inspection, disintegration, colorimetry, or TLC. Due to wide variability in the drug quality reporting of the studies identified (out of 4621 records found, only 39 studies were eligible to be included in the study), the authors suggested the incorporation of standard guidelines to improve the studies’ generalizability.

1.2. Counterfeit Pharmaceutical Packaging

As compared to the analysis of the pharmaceuticals themselves, the amount of studies focused on pharmaceutical packaging materials is relatively modest. For instance, Kwok et al. [87] used Raman spectroscopy to identify chemical differences between authentic and counterfeit Cialis© cardboard packages with two different ink colors (yellow and white). They hypothesized that variations in the white ink and yellow ink composition could distinguish between a counterfeit as compared to an authentic package. Additionally, Rodomonte et al. [88] utilized a colorimeter to analyze authentic Viagra© (blue ink), Cialis© (green ink), and Levitra© (violet ink) paperboard packages. Due to differences in providers’ manufacturing process, the authors recommended a frequently updated spectral library of packaging inks in order to reduce the false negative (authentic determined to be counterfeit) rate.

Dégardin et al. published two studies [89, 90] on the analysis of counterfeit pharmaceuticals. The first analyzed the differences between seven authentic pharmaceutical packaging samples (same brand and product) from two different manufacturing sites and five counterfeit pharmaceutical packaging samples [89, 90]. The identity of the product was unspecified in the study, but the authors analyzed seized paperboard packaging, glass vials with an aluminum cap, and paper leaflets. Six techniques were used to discriminate between the authentic and counterfeit packaging: visual observation, X-Ray Fluorescence, Raman microscopy, ATR-FTIR, Scanning Electron Microscopy (SEM-EDS), and microcomputed tomography. Although there was variant success in identifying counterfeits when only one technique was used, the authors were able to determine all
counterfeits and chemical links between counterfeit types when the techniques’ data were combined with one another.

The second study from Degardin et al. focused on the authentication of glass vials [90]. Five components of glass vials were investigated: flip-off caps, aluminum crimping caps, stoppers, the vial body, and vial labels. To discriminate them from their authentic counterparts, Raman spectroscopy/microscopy, X-Ray fluorescence, ATR-IR, and optical microscopy were successfully used and were able to note eight linked counterfeits through chemical differences and similarities.

1.3. Ink Composition

1.3.1. General Ink Analysis

Ink is composed of both organic and inorganic components, varying due to the general class of ink and specific formulation from the manufacturer. Broadly, ink is separated into three major components: colorant, vehicle, and additives. The colorant imparts the color to the ink. The vehicle acts as transport to place the ink on the substrate. Additives affix various qualities to the ink (surfactants, extenders, biocides) [91]. The colorant is typically composed of either a dye, a pigment, or a combination. A dye is typically soluble within its vehicle, whereas a pigment consists of small insoluble particles suspended within the vehicle. The ink class usually determines the colorant type; some liquid inks such as inkjet are commonly dye-based to prevent clogging of the printing system, other inks such as toner are commonly pigment-based [92]. However, pigments have recently been used in inkjet inks to increase photo reproduction quality [93] Offset inks are generally used for the printing of pharmaceutical packaging [94], but many other types of inks can be used in the production of counterfeit packaging. The two most common types of home and office printer inks (inkjet and toner), as well as offset ink, will therefore be discussed.

1.3.2. Inkjet Ink

Inkjet inks are liquid inks that are deposited onto an appropriate substrate from a nozzle through the use of a piezoelectric force. This is done through either a continuous inkjet system (whereupon
the ink is deposited continuously through the nozzle) or a drop-on-demand inkjet system, where
the ink is only deposited when required. The ink deposition is controlled through a pressure pulse
within the system [93].

The 4 CMYK (Cyan, Magenta, Yellow, Black) pigment set is becoming increasingly common in
inkjet printing [93]. Common composition of inkjet inks is found in Table 1.

Table 1. Example composition of inkjet inks [93, 95].

<table>
<thead>
<tr>
<th>Ink Component</th>
<th>Chemical Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicle</td>
<td>Water, alcohols</td>
</tr>
<tr>
<td>Colorant</td>
<td>Carbon black (black ink), copper phthalocyanine (cyan), quinacridone (magenta), azo compounds (yellow)</td>
</tr>
<tr>
<td>Inorganic Additives</td>
<td>Ti, Zn, Mg, Al, Ca, P, Zr, Cr, Fe</td>
</tr>
</tbody>
</table>

Several studies (Williamson et al., Corzo et al., Trejos et al.) have used various analytical
techniques to successfully discriminate printing inks: LA-ICP-MS, LIBS, DART-MS, SEM-EDS, and
ATR-FTIR [96, 97, 98]. The discrimination power of these techniques for inkjet inks, along
with the list of inorganic elements from a 78 inkjet set detected by LA-ICP-MS, are shown in
Table 2.

Table 2. Discrimination power of inkjet inks using different analytical techniques and elements
identified through LA-ICP-MS [96, 97, 98].

<table>
<thead>
<tr>
<th>Discrimination Power (%)</th>
<th>Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>99.8%</td>
<td>LA-ICP-MS</td>
</tr>
<tr>
<td>94.3%</td>
<td>LIBS</td>
</tr>
<tr>
<td>NA</td>
<td>DART-MS</td>
</tr>
<tr>
<td>49.2%</td>
<td>SEM-EDS</td>
</tr>
<tr>
<td>NA</td>
<td>ATR-FTIR</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Elements Detected</th>
<th>Technique</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al, B, Ba, Cu, Hf, K, Li, Mg, Na, S, Sn, Zr</td>
<td>LA-ICP-MS</td>
</tr>
</tbody>
</table>
Of the examined techniques, LA-ICP-MS, LIBS, and DART-MS had excellent discrimination power (>90%) while SEM-EDS and ATR-FTIR had poor-to-passable discrimination power (78.0% and 50% respectively). ATR-FTIR spectra were dominated by cellulose or solvent peaks found within the paper substrate or liquid ink respectively, complicating discrimination. While SEM-EDS had relatively high discrimination capability for non-black ink, the technique was not sensitive enough to discriminate between the trace inorganic materials found within the carbon black pigment.

1.3.3. Toner Ink

As opposed to inkjets, toner inks typically come in a solid powder consisting of pigments suspended in a resin. A rotating drum within the printer applies the toner to the substrate, and a heated fuser melts the toner onto it [92]. Examples of common resins, pigments, and inorganic additives is shown in Table 3.

<table>
<thead>
<tr>
<th>Ink Component</th>
<th>Chemical Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resin</td>
<td>Polyesters, acrylate copolymers, styrene</td>
</tr>
<tr>
<td>Pigments</td>
<td>Carbon black, nigrosine, quinacrodone</td>
</tr>
<tr>
<td>Inorganic Additives</td>
<td>Si, B, Sn, Zn, Ti, Al, Ba, K, Ca, Cr, Fe, Mg,</td>
</tr>
<tr>
<td></td>
<td>Fe, Ni, Al, Ag, Au, Cu</td>
</tr>
</tbody>
</table>

Several studies have also successfully discriminated different brands of toner inks. The discrimination capabilities of the studied analytical techniques, as well as the inorganic elements from a 76 toner set determined by LA-ICP-MS, are shown in Table 4.
Table 4. Discrimination power of toner inks using different analytical techniques and elements identified through LA-ICP-MS [96,97,98].

<table>
<thead>
<tr>
<th>Technique</th>
<th>Black</th>
<th>Magenta</th>
<th>Cyan</th>
<th>Yellow</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA-ICP-MS</td>
<td>97.6%</td>
<td>100%</td>
<td>98.5%</td>
<td>97.4%</td>
<td>99.0%</td>
</tr>
<tr>
<td>SEM-EDS</td>
<td>94.2%</td>
<td>95.5%</td>
<td>95.5%</td>
<td>94.9%</td>
<td>97.1%</td>
</tr>
<tr>
<td>LIBS</td>
<td>88.7%</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>88.7%</td>
</tr>
<tr>
<td>ATR-FTIR</td>
<td>75.7%</td>
<td>71.2%</td>
<td>71.2%</td>
<td>79.5%</td>
<td>86.0%</td>
</tr>
<tr>
<td>DART-MS</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>82.0%</td>
</tr>
</tbody>
</table>

There was much less variation in the discrimination power of the techniques as compared to inkjet. In general, they all performed acceptably for a screening method (>80%). DART-MS experienced lower discrimination capabilities due to some interferences from uneven melting of the toner.

1.3.4. Offset Ink

Offset ink is applied to a substrate through the use of a rubber plate. The ink is first applied onto it, and then the rubber plate presses against the substrate to form the image. Unlike inkjet, where ink becomes embedded onto the substrate fibers, and toner, where micron-sized particles fuse to the substrate surface, offset ink forms a thin film partially embedded onto the surface [93]. Like toner inks, offset inks are typically composed of a resin and a pigment. Common offset ink compositions are shown on Table 5.
Table 5. Example composition of offset inks [100, 101, 102, 103].

<table>
<thead>
<tr>
<th>Ink Component</th>
<th>Chemical Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resin</td>
<td>Phenols, olefins, paraffins</td>
</tr>
<tr>
<td>Pigment</td>
<td>Azo, calcium carbonate, titanium dioxide</td>
</tr>
<tr>
<td>Inorganic Additives</td>
<td>Bi, Ti, Fe, Sn, Cr, Zn, Ni, Cu</td>
</tr>
</tbody>
</table>

As in the previous sections, the discrimination power of various analytical techniques on black offset ink brands are shown in Table 6 as well as the inorganic elements from a 79 offset ink set detected by LA-ICP-MS.

Table 6. Discrimination power of offset inks using different analytical techniques and elements identified through LA-ICP-MS [96, 97, 98].

<table>
<thead>
<tr>
<th>Technique</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA-ICP-MS</td>
<td>99.8%</td>
</tr>
<tr>
<td>ATR-FTIR</td>
<td>92.6%</td>
</tr>
<tr>
<td>SEM-EDS</td>
<td>80.0%</td>
</tr>
<tr>
<td>DART-MS</td>
<td>56.1%</td>
</tr>
</tbody>
</table>

Technique | Element Identification
-----------|------------------------
LA-ICP-MS  | Al, Ba, Bi, Ca, Co, Cu, Fe, Hf, K, Mg, Mn, Mo, Na, Nb, Pb, Sb, Si, Sn, Sr, Ti, W, Zn, Zr

Like toner inks, the analytical techniques appreciated considerable discrimination power when discriminating offset inks (generally > 80%). The analytical technique with the lowest discrimination power was DART-MS (56.1%); there was difficulty in interpreting the DART-MS spectra due to the partially embedded ink on the substrate.

1.3.5. Paperboard Substrate

Along with ink analysis, the paperboard substrate of the sample can also reveal important information about the sample. The main constituents of paper are cellulose and hemicellulose,
which originates from wood pulp [104]. However, many additives can be added to the paper to impart certain properties. Examples are shown in Table 7.

Table 7. Common components in document paper [104, 105].

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Chemical Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main structural component</td>
<td>Cellulose, hemicellulose</td>
</tr>
<tr>
<td>Fillers</td>
<td>Silicates, carbonates, oxides, sulfides</td>
</tr>
<tr>
<td>Filler retention compounds</td>
<td>Polyacrylamide, starch</td>
</tr>
<tr>
<td>Coating pigments</td>
<td>Calcium carbonate, talc, titanium dioxide</td>
</tr>
</tbody>
</table>

Several techniques have been employed in order to analyze paper in a questioned document context. The most common techniques are inorganic in nature, particularly varieties of XRF, ICP-MS (as well as LA-ICP-MS), and LIBS. A study by Trejos et al., analyzed a sample set of 26 different-source document papers and found the discrimination capabilities for laser ablation methods to be quite high for paper (99.4% for LA-ICP-MS and 97.7% for LIBS) [106]. Although most document papers share a joint base, the use of inorganic additives adds identifying characteristics to each paper batch. A list of elements detected by LA-ICP-MS, LIBS, and EDXRF are listed in Table 8.

Table 8. Common elements identified in document paper [105, 106].

<table>
<thead>
<tr>
<th>Technique</th>
<th>Elements Identified</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA-ICP-MS</td>
<td>Na, Al, Zn, Mg, Sr, Fe, Mn, Cu, Ti, Ba, Zr</td>
</tr>
<tr>
<td>LIBS</td>
<td>Na, Al, Ca, Sr, Mg</td>
</tr>
<tr>
<td>EDXRF</td>
<td>Pb, Rb, Sr, Y, Zr, K, Ca, Ti, V, Cr, Mn, Fe, Ni, Cu, Zn, Co</td>
</tr>
</tbody>
</table>
1.3.6. Summary of Potential Use of Ink Analysis for Source Classification

When choosing instrumental techniques for the discrimination between authentic and counterfeit pharmaceutical packaging samples, several qualifications were considered. Characteristics inherent to the technique were considered for their applicability to remote analysis. Many questioned pharmaceutical analyses happen in the field, so the technique had to possess a relatively quick analysis time, minimal destruction to the sample, limited sample preparation, sensitivity, and portable instruments available.

Additionally, techniques that had previously shown success in ink discrimination in the literature were prioritized. As explained in 1.3.1 - 1.3.5., ink and paper possess both organic and inorganic constituents. It was hypothesized that more accurate discrimination between sources could occur if the techniques yielded complementary information, such as organic and inorganic data, that could be analyzed in combination with one another.

As explained in 1.4.1.1. and 1.4.2.1., LIBS and ATR-FTIR both possess the technical advantages described previously. LIBS was chosen as the primary technique to acquire an inorganic profile of the sample. Meanwhile, ATR-FTIR was chosen as the primary technique to acquire an organic profile of the sample. The information from these two spectroscopic techniques could easily be combined to provide a robust account of the sample composition, with enough informative features to adequately discriminate between sources.

1.4. Instrumental Techniques
   1.4.1. Laser Induced Breakdown Spectroscopy (LIBS)
      1.4.1.1. LIBS Overview

LIBS is an atomic, ionic, and molecular emission technique that measures the wavelengths of light given off by electronic transitions. It is able to identify and quantitate most inorganic elements on the periodic table. There are several advantages intrinsic to the LIBS technique, including low sample acquisition time (<1 min), sensitivity, quasi-non-destruction of the sample, and portability [107].
This spectroscopic technique utilizes a temporal micro-plasma to induce electronic excitation on the surface of the sample (with only a penetration depth of only a few nanograms). The subsequent relaxation into the electronic ground state releases light, with wavelengths characteristic to each element. These wavelengths can be collected, separated, and translated into electrical signal for analysis. Overall, a LIBS spectrum is collected with intensities at each wavelength stretching from the ultra-violet to infrared region allowing for both qualitative and quantitative analysis.

LIBS is especially suitable for printing ink analysis. The penetration depth of the laser can be modified by a series of adjustable parameters (such as gate delay, spot size, laser energy %). When these parameters are optimized, sufficient ablation can occur so that ink signal is collected while paper signal is minimized. The sensitivity for LIBS also allows for the identification of trace constituents of the ink formulation, which is critical for the purposes of discrimination.

1.4.1.2. LIBS Theory
1.4.1.2.1. Lasers

Lasers are an integral component of the LIBS technique. There are several types of lasers, however, for purposes of this research we will focus the discussion on the Nd:YAG, utilized in this study. LASERs (Light Amplification by Stimulated Emission of Radiation) broadly consists of three major components: a pump source, a gain medium (e.g., Nd:YAG) crystal), and an optical resonator [109]. As the gain medium was a solid crystal, this laser formation would be termed as solid-state laser.

The pump source is typically a flashlamp or diode stack. The purpose of the pump source is to pump light into the lasing material. The physical crystal is an yttrium aluminum garnet (YAG, Y₃Al₅O₁₂,) that has been doped with a neodymium ³⁺ ion (Nd) up to 1.4% of the original crystal composition [108]. The neodymium ions absorb the light, exciting their electrons to the next energy level. Soon, the electrons relax (radiation-less) into a metastable state (above the ground state, but less unstable than the excited state). Neodymium ions (and rare earth elements in general, such as Eu) are advantageous for lasing because of the long lifetime state in their metastable level
(230 µs, as opposed to ground level 30 ns) [108]. Due to this, a population inversion quickly occurs, wherein most of the electrons in the medium are found in the excited energy level rather than the ground level.

Some instruments, including the one used in the study, utilize a Q-Switch. A Q-Switch increases the power of the laser. Before total population inversion occurs, the Q-switch is closed and prevents photons from inducing energy emission on the excited electrons. Once the energy within the lasing cavity crosses a certain threshold, the Q-switch opens, and light is allowed to reflect back into the chamber between the optical resonators. The optical resonators consist of several mirrors placed around the lasing material, meant to reflect emitted light back into the lasing cavity. This light strikes the excited ions and stimulated emission occurs, allowing the ions to decay into the ground state and excite the photons (or, allow them to have laser gain). Allowing the light to reflect back into the chamber all at once increases the irradiance of the laser pulse as it leaves the chamber [110]. A general laser schematic is shown in Figure 2.

![Figure 2. Diagram of a Q-switched Nd:YAG laser. Pumping source is a flashback lamp [109].](image)

### 1.4.1.2.2. Plasma Formation and Light Emission

Once the laser pulse is fired, it is focused through a quartz lens onto the sample. Each pulse is typically 10-25 mJ for a 266-nm Nd:YAG laser, approximating $2.5 \times 10^{17}$ protons per pulse [107]. The interaction of the laser with the sample ablates away a dry aerosol of nanoparticles, consuming
few nanograms to micrograms of the material, depending on the ablation conditions. As the laser beam impacts the sample, the energy from the pulse creates a micro-plasma that reaches 8,000 – 10,000K, consisting of the ionized vapor.

This plasma produces excited atoms, ions, and molecules, and will only last on the orders of nanoseconds to microseconds before the excited species within it start to decay back to the ground state. In doing so, they release light in wavelengths characteristic to their element. The resulting light is transferred via fiber optic cable to the spectrograph.

The spectrograph consists of a series of fine gratings and mirrors and for the purpose of this research the spectrograph used was a Czerny-Turner. When the light enters the spectrograph, it is highly polychromatic in nature. The polychromatic light enters through the exit slit, where it strikes a mirror. This mirror focuses the light on a fine grating, which separates the light into discrete wavelengths. It reflects off the grating onto a second mirror, whereupon the monochromatic light exits through the exit slit. From there, the photons are focused on an array detector. This array detector will allow monochromatic light to move to the detector [111].

A CCD (charge-coupled device) is used to convert the received light into a digital signal. The light strikes a number of pixels on an array. This induces a charge on the pixels. Dimensions for this array can differ with design, with common dimension sizes being 512 x 1024 or 1024 x 1024 pixels. For each wavelength, the charge on each row of pixels is transferred into the bottom horizontal register, which converts this charge into electrical signal. This electrical signal is read by computer software. The process is repeated for all rows in the array, allowing for construction of a complete spectrum [107]. A rough schematic of the LIBS instrument in shown in Figure 3.
One of the advantages of the LIBS technique is that the instrument parameters can be optimized and modified to suit specific sample types. Due to the transient nature of the signal, a critical component of parameter optimization is related to the proper selection of signal acquisition time. For example, gate delay refers to the portion of time that the spectrometer waits before collecting data. When the electrons first start to decay, a series of bound-unbound and unbound-unbound energy collisions occur, resulting in intensity readings at non-characteristic wavelengths and causing high levels of background “noise” in the spectrum. Increasing the gate delay allows for the decay of this background noise but may also reduce the signal at wavelengths of interest (characteristic emission lines of ions and atoms). However, the decay in background occurs at a faster rate than the decay in signal. Other parameters that affect the ablation process are laser frequency, laser energy, and spot size. Frequency refers to the rate of shots being fired per seconds and is expressed in Hz. Laser energy does not explicitly refer to the percentage of utilized laser energy but is the software unit for the laser power. This correlates to the fluence (received energy per surface area), and irradiance (received work per surface area; or received energy per surface area per time). While a higher laser energy does correlate to more energy impacting the surface of the sample, it will also increase background noise and penetration depth. Spot size refers to the diameter of spot made by the laser.
1.4.1.3. LIBS Application

Since its creation in the 1960s, LIBS has found analytical use in many different fields, from astronomy \[112\] to zoology \[113\]. Due to the camera incorporated on the instrument and the low destructivity, LIBS has found extensive utility in trace evidence analysis, including ink and paper analysis \[114, 115, 116, 117\].

Using LIBS, Rzecki et al. \[114\] discriminated thirty classes of pen ink and paper using seven computational intelligence methods, eventually achieving accuracies of 97% - 98% and discrimination powers of 96 – 99%. Lennard et al. \[115\] examined thirty-three different brand/batch paper samples and a hundred and thirty one ink samples (inkjet, ballpoint, and toner) and found discrimination capabilities of 99% (paper), 99% (black ballpoint), and 93-97% (printing ink) respectively.

Subedi et al. \[116\] utilized both LIBS and LA-ICP-MS to examine different categories of printing ink: inkjet (ten samples), toner (nine samples), offset (twelve samples), and intaglio (twelve samples). They achieved discrimination rates of 98%, 100%, 92%, and 88% respectively. Metzinger et al. \[117\] focused on differentiating between both paper types (six sources) and black/color printers (eight printers). They achieved an accuracy of 96% for discriminating paper types and an accuracy of 83% in discriminating printers.

1.4.2. ATR-FTIR

1.4.2.1. ATR-FTIR Overview

Attenuated Total Reflectance Fourier Transformed Infrared (ATR-FTIR) is a molecular spectroscopic technique that evaluates primarily organic functional groups through their vibrational and rotational frequencies. It can be used to identify the functional groups of a molecule qualitatively or to quantitate. ATR-FTIR has several advantages, including fast analysis time (<1 min), sensitivity, no sample preparation, no sample destruction, and the option for portability.
ATR-FTIR focuses on the absorption and transmission of infrared light with a sample. In order for a molecule to be considered IR-active, it must have a dipole moment (i.e. the molecule must be overall polar). If the frequency of light matches the vibrational frequency of the bond, the bond will absorb a portion of the evanescent wave. The evanescent wave (now attenuated) will return to the crystal and reflect out of it before being transported to the detector. The detector then translates this information into an electrical signal.

Due in part to its lack of sample destruction, ATR-FTIR is predicted to be a satisfactory technique for analysis. The ATR attachment enables the samples to be analyzed with no sample preparation. Unlike the LIBS, the light source is not intense enough to destroy any portion of the sample. ATR-FTIR only has a penetration depth of a few microns, so the paper is not anticipated to have a large contribution to the ink spectra.

1.4.2.2. ATR-FTIR Theory

Like LIBS, ATR-FTIR utilizes light to determine the chemical properties of samples. However, there are many theoretical and structural differences between the two techniques. A general schematic of an FTIR instrument is shown in Figure 4.

![Figure 4. Diagram of FTIR instrument [121].](image)

First, the light source (in the study, a Globar was used, though other varieties such as Nernst and tungsten sources exist) is heated and acts as a black body. Once heated to roughly 1000 K – 2000 K, it emits polychromatic light radiation in the infrared region. This light directly enters the Michelson interferometer.
A specific schematic of a Michelson interferometer is shown in Figure 5.

![Diagram of Michelson Interferometer](image)

Figure 5. Diagram of Michelson Interferometer [121].

Once the light enters the Michelson interferometer, it strikes a beam splitter. The beam splitter used was a partially reflective KBr crystal that is capable of measuring from 400 cm\(^{-1}\) to 4000 cm\(^{-1}\). A portion of the light traverses directly through the beam splitter to hit a totally reflective mirror, whereas the other portion is reflected to hit another mirror. The light reflects off the mirrors and encounters the beam splitter again, where another portion of light continues to the sample.

One of these mirrors remains at a fixed point. The other is movable and adjusts portions of millimeters. This shifts the phase of the light waves that strikes the movable mirror. The consequence of this is, when the light recombines at the beam splitter for the second time, they experience a change in amplitude. This recombined wave is permitted to travel, through a series of reflecting mirrors, on to the ATR accessory.
Attenuated Total Reflectance (ATR) simplifies sample preparation by allowing the bulk sample to be placed directly on the optical window, allowing for direct solid and liquid sample analysis. A simplified design of an ATR accessory is shown in Figure 6. The sample is pressed against the crystal, and the light from the interferometer is passed into a crystal (diamond, in the case of the Perkin Elmer Spectrum Two used in this study). It’s imperative that the crystal has a higher refractive index than the sample. If this is not the case, some energy will be lost to the sample at every reflection. While the light reflects off the interior of the crystal, it strikes the sample and penetrates a few microns [118].

![Diagram of ATR accessory apparatus](image)

Figure 6. Diagram of ATR accessory apparatus [118].

It is possible for certain molecules to absorb a portion of the light, according to the fundamentals of vibrational spectroscopy. Molecules vibrate at specific frequencies, according to the strength of the bond and the atoms associated with it. However, molecules are also able to vibrate in different ways, termed vibrational modes: symmetrical stretching, asymmetrical stretching, bending, rocking, wagging, and twisting. The amount of vibrational modes per molecule can be determined by the equation $3N - 6$, where $N$ is the number of atoms in the molecule. If the molecule is linear, the equation shifts to $3N - 5$ due to the symmetry on one plane. Vibrational frequencies of different modes of the same bond are typically distinct from one another [121]. Additionally, a molecule is only considered IR active if it possesses a dipole moment, or a non-zero polarity across the atoms in the bond.

If an extant photon approaches the system, and the frequency of the light matches the vibrational frequency of the bond, then the photon can be absorbed into the system. This causes the bond to gain energy and rise into an excited vibrational state. If the extant photon’s frequency does not
match the vibrational frequency, then it will be reflected back into the diamond. Therefore, the presence of certain bonds at certain vibrational modes can be indicated by the amount of light returned to the system (or, in other words, how well the light transmits through the sample system, where a higher transmittance indicates less photon frequency/bond frequency “matches”).

From there, the light exits the ATR accessory and approaches the detector (lithium tantalate, LiTaO₃, MW 236 g/mol). The detector transforms the light into an electrical signal, creating an interferogram. As the signal was measured over a wide variance of optical path lengths (and therefore interference patterns), this interferogram is initially created in the time domain (with the y-axis indicating intensity, and x-axis indicating time). The use of the Fourier Transform equation, can be used to transform the pattern into the more readily interpretable frequency domain. This process has three main advantages over monochromatic spectroscopy: resolution, multiplexing, and throughput all increase [119]. From there, major vibrational bands can be observed, indicating specific functional groups within the molecule. The spectrum is divided into three major regions: far (10 – 400 cm⁻¹), mid (400 – 4000 cm⁻¹), and near (4000 – 14000 cm⁻¹). In this study, the mid-IR region was investigated.

1.4.2.3. ATR-FTIR Application

ATR-FTIR’s versatility facilitates its use in many fields, including ink analysis. Williamson et al. [101] used ATR-FTIR to distinguish inkjet (n=78), toner (n=76), offset (n=79), and intaglio (n=86) inks from one another. The reported discrimination capabilities were then 49.5%, 86.3%, 92.6%, and 33.6% respectively. Due to the ink deposition process of inkjet and intaglio inks, the authors indicated that the ATR-FTIR spectra of inkjet and intaglio samples were very similar and showed prominent peaks of the paper substrate. Further, the authors identified the discrimination capabilities of the different colors of toner ink separately (black, magenta, cyan, and yellow) and determined them to be 75.7%, 71.2%, 71.2%, and 79.5%.

The technique has also been used to identify counterfeit drugs and discriminate them from their authentic counterparts. Scaffi and Pasquini used the near-IR portions of the spectrum to provide compositional information of authentic and counterfeit tablets, solutions, and capsules. They
determined that it was possible to use ATR-FTIR for this purpose with a variety of active pharmaceutical ingredients (aspirin, ampicillin, levonorgestrel) [121].

As Viagra and Cialis are a popular choice for counterfeiters, multiple studies have focused on the utilization of ATR-FTIR to discriminate between these authentic and counterfeit vasodilators. Anzanello et al. [122] specifically investigated on which portions of the ATR-FTIR spectra were more informative for this purpose. It was concluded that the lactose bands in the spectra were the most discriminating, specifically the bands occurring in the 1091 – 1046 cm⁻¹ region.

1.5. Machine Learning Techniques

1.5.1. Machine Learning Techniques Overview

Machine learning techniques aim to ‘teach’ a computer how to perform a task without being explicitly programmed to do so. For the purposes of classification, a machine learning technique creates a mathematical model. This model utilizes inputted features of sample classes in order to best highlight the disparate and similar features between them, and to determine a class for an otherwise unassigned, or “questioned”, sample.

Machine learning can also be broken down into supervised and unsupervised techniques. Unsupervised techniques focus on examining the distribution of the data and do not require labeling of the sample class for training the algorithm. Clustering techniques are commonly unsupervised, but because a known dataset is not implemented, classification is difficult without the formal construction of a model. Conversely, supervised techniques utilize a known dataset (where both features and class label are inputted into the system) in order to train a model. This model is then applied to a dataset of known features and unknown classes in order to assign them [124].
1.5.2. Random Forest (RF)

Random Forest is a widely-used classification method that involves decision trees, wherein a sample is classified according to the outcome of a series of binary nodes, or tests. A simple example of a decision tree is located in Figure 7. They have the advantage of ruggedness to different data types and a high number of feature variables, at the cost of overfitting to the training data [125].

![Decision Tree Diagram](image)

**Figure 7.** An example of a basic decision tree.

Random Forest models try to minimize the overfitting problem of decision trees by creating an ensemble classifier that incorporates many independent decision trees and allow them to vote for the most likely class. These independent decision trees utilize both a random subset of predictors (feature bagging, to grow the tree) and a random subset of samples (sample bagging, to classify).
in order to maintain the independence of the trees [126]. Random forests in particular have been shown to maintain high accuracy even with differently-scaled or non-linear data. Still, due to this randomness, they can be sensitive to datasets with high levels of noise.

The random forest classifier has been used with success with LIBS data in previous studies, including studies classifying polypropylene contaminants, ceramics, and iron ore. Using random forest, Rzecki et al. were able to classify 30 pen ink and paper classes, obtaining an accuracy and discrimination power of 99.1% [116]. Random forest classifiers were also successful in studies classifying crop type (>90% correctly classified) [127], brand recognition using potentiometry (99.07% correctly classified) [128], and andiroba oil (100% correctly classified) [129].

1.5.3. Naïve Bayes (NB)

The simplest form of a Bayesian network is a Naïve Bayes classifier, in which one class variable is the parent for a series of nodes (features) which in turn explain the class, such as the example seen in Figure 8. These descendant nodes are assumed to be entirely independent from one another, an assumption which isn’t often met in case data, therefore its term “Naïve”. However, Naïve Bayes has the advantage of being particularly rugged even when several non-explanatory variables are present. It is also able to classify systems with many features and fewer training samples because the features are presumed to be independent [130].
Figure 8. An example of a basic Naive Bayes classifier.

Like Bayesian networks, the Naïve Bayes classifier works according to Bayes theorem, as seen in Equation 1.

\[
p(C|x) = \frac{p(x|C) p(C)}{p(x)}
\]

Equation 1. Bayes' Theorem [131].

In the equation, the posterior probability (or, the conditional probability of witnessing class \( C \) given predictor \( x \)) is set equal to the product of the prior probability of the class and the probability of the predictor given a class, divided by the prior probability of the predictor. The class that yields the highest posterior probability is then determined to be the most probable class, and the class the sample is assigned to.
Naïve Bayes has been successfully used in many applications, from geology to language analysis. A Naïve Bayes classifier was used to predict DNA-binding proteins and was determined to have 80% accuracy [133]. In a study that analyzed the sentiment behind social media statuses, a Naïve Bayes classifier was found to correctly identify the statuses as positive 76% of the time and to correctly identify the statuses as negative 65% of the time [134]. Additionally, Naïve Bayes worked as a classification technique for agricultural land soils, resulting in 100% correct classifications [132].

1.5.4. Neural Network (NN)

Neural Networks are composed of node layers to assist with classification and require an extensive training set to prepare the algorithm to test samples. To simplify, predictors are entered into the model to input nodes and operates simulating the process of neurons in human brain. The input nodes calculate and assign a certain weight to the predictors before passing them along to the activation function. A computation occurs; if the computed weight is larger than a certain threshold, the activation function ‘fires’ and the data is transmitted to the next layer. Eventually, the output layer is reached. The firing of certain nodes on the output layer will indicate which class the questioned sample is assigned to. During the training session, the model records its misclassified samples and adjusts its models accordingly to reduce that error. An example of a simple neural network with three hidden layers is shown in Figure 9. Based on predictor variables 1, 2, and 3, the sample was assigned to class A [135]

Figure 9. An example of an artificial neural network, with three hidden layers.
Neural networks are typically used to mapping complex interacting data, and have no assumptions of independence, normality, or linearity. The extensive corrections made to the model during training allow the technique to adapt to the unique dataset. However, because the model adjusts during the training set cycles, limited adjustments can be made by the operator during that time. Due to this, training the model can be quite computationally expensive and has been criticized for being a ‘black box’ of machine learning [136].

As neural networks are capable of handling even highly correlated data, it has grown in popularity as a classifier, particularly among fields where predictor variables are rarely independent of one another. A neural network classifier was used to correctly classify 900 wine samples based on a sensor array [137]. Using data retrieved from an electrocardiogram (ECG), NN was used to correctly identify the results as stemming from one of four cardiac diseases with 85% accuracy [138]. Similarly, NN yielded 95% accuracy when used to classify a seizure as normal or epileptic [139]. LIBS was used to determine the presence of gunshot residue (GSR) and samples were classified as either originating from a shooter or non-shooter. Using both NB and NN (as well as a critical threshold and logistic regression), accuracies from 87 – 100% were demonstrated [140].

1.5.5. k-Nearest Neighbors (KNN)

k-Nearest Neighbor is a classification technique that involves examining the samples with the most similar attributes (the “neighbors”) to the questioned sample and assigning the questioned a class based on the similarities with neighbor classes. The samples of known classes are displayed and the distances between the known and questioned samples features are calculated. The $k$ smallest distances between the known and questioned samples are noted, and majority rule decides the class of the questioned sample. $k$ is usually chosen to be a small integer (generally <5) and odd, to prevent stalemates [141].

KNN is referred to as a “lazy” learning algorithm, in the sense that there is no “model creation” phase when the training set is introduced. Instead, when a questioned sample is introduced to the system, the “memorized” training set is analyzed for the nearest neighbors and a classification is
made. While this prevents input data from affecting the overall accuracy, it does increase the system’s sensitivity to extreme values [142].

An example of the $k$-Nearest Neighbor technique can be shown in Figure 10. In this case, two out of three closest ‘neighbors’ to the questioned sample are in class A. Therefore, the questioned sample is assigned to class A.

![Figure 10. An example of a basic KNN classifier.](image)

$k$-Nearest Neighbors has several advantages for classification problems. It makes no \textit{a priori} assumptions about the data. However, KNN has decreased accuracy with large datasets and feature scale is recommended [143].

Due partially to the lack of assumptions made about the data, KNN has encountered wide versatility for classification problems. In a study evaluating classification of gasolines, KNN received a 0% (1 ± 1%) error rate when classifying samples [144] and was deemed one of the most effective methods, along with Support Vector Machines and Probabilistic Neural Networks. KNN has also yielded relatively high accuracy for the classification of visual images (71.8%) [145], as well as a low percent error for prediction patterns (<10%) [146]. In the pharmaceutical sector, KNN has proved promising for identifying lead compounds in high throughput screening.
searches, as well as discriminating between counterfeit and authentic spectroscopic spectra of medicines. [147, 148]

1.5.6. Linear Discriminant Analysis (LDA)

A broad description of discriminant analysis is necessary before discussing the details of the quadratic subsection. Discriminant analysis presumes the usage of a continuous predictor variable (independent variable) to determine a categorical dependent variable. Conceptually, discriminant analysis aims to find a combination of variables that maximizes the between-class variance and minimizes the within-class variance in order to increase separation between groups. The model creation of LDA consists of the construction of a discriminant function, or a weighted combination of predictor variables that best separates the classes [149]. For LDA, this discriminant function is given in Equation 2.

\[ \delta_c(x) = -\frac{1}{2} \log |S_c| - \frac{1}{2} (x - \mu_c)^T S_c^{-1} (x - \mu_c) \]

Equation 2. LDA discriminant function (C = class, S = Covariance Matrix, x = sample value) [150].

The class \( k \) that maximizes the value \( \delta \) is the assigned class for the questioned sample. LDA specifically utilizes a pooled class covariance (unlike QDA, where each class’ covariance is calculated separately). Due to this, the probability contours that are determined as a result of determining \( \delta \) are linear in nature, as seen in Figure 12.
Figure 11. An example of a basic LDA classifier.

As to be expected, LDA functions when the covariance between classes is similar. When covariances between classes are highly disparate, QDA may be the better classification technique. However, because LDA utilizes a pooled covariance, it is more rugged to datasets with low sample size.

Discriminant analysis in general and LDA have several assumptions that must be met for maximum efficacy. DA presumes normality, homoscedasticity, and independence in the data. When these assumptions are met, DA can be a sensitive and accurate classification technique. If the assumptions are not meant, accuracy suffers.

When the covariances can be assumed to be roughly equal, LDA is an effective classifier and has been historically used in research for decades. Oravec et al. used LDA as a method for classifying black inkjet ink [154]. LDA had superior classification accuracy over QDA for both classifying carbon black and unspecified black colorant. Another study found that LDA was able to identify common materials found in shipwrecks using LIBS with an 87% accuracy [155] and to identify lymphomata using LIBS with an 99.7% accuracy [156].
1.5.7. Quadratic Discriminant Analysis (QDA)

Like LDA, QDA is a subset of DA and follows the same basic principles and assumptions described in 1.5.6.1. An example of a QDA system is given in Figure 12 where the questioned sample, because it is on one side of the quadratic demarcation line, is assigned to Class A.

![Figure 12. An example of a basic QDA classifier.](image)

QDA specifically is adept at handling datasets with variable class covariances as the covariance matrix is separated for each class separately. On the other hand, this causes noise to dominate the algorithm if intra-class sample sizes are small [149].

Although DA has a number of assumptions about the data, it has nevertheless found a wide application to many fields of science. Mahmodi et al. utilized QDA and other machine learning techniques to classify various diesel-biodiesel blends, resulting in a precision, sensitivity, and specificity of 94.1%, 94.1%, and 97.5% respectively [151].

In microbiology, QDA was used to classify antiviral/HIV, antifungal, anticancer, and antibacterial antimicrobial peptides with over 89% accuracy [152]. QDA was able to classify Venezuelan firewater samples with over a 70% success rate [153].
1.5.8. Machine Learning Summary

Both LIBS and ATR-FTIR produce an output of a spectrum: for LIBS, the spectrum represents emission wavelength v. intensity, for ATR-FTIR, the spectrum represents wavenumber v. %T. Each spectrum can represent thousands of data points. Although specific emission lines are often chosen for LIBS and only a portion of the spectrum is used for ATR-FTIR, each spectrum represents a large packet of data. During the course of this study, nearly 4000 spectra were collected for each technique.

For that reason, traditional statistical methods (such as ANOVA analysis) were infeasible. Instead, machine learning methods were utilized as a tool to create an algorithm that could process a large amount of data at once. As very few studies have been conducted on pharmaceutical packaging, more than one machine learning method was chosen in order to examine which, if any, had an advantage over the others in terms of accurately processing LIBS and ATR-FTIR data.

Three of the chosen machine learning methods (Random Forest, k-Nearest Neighbors, and Neural Networks) make no formal distribution assumptions about the data. These were chosen under the prediction that the LIBS and/or IR data may not be parametric. However, both Random Forest and k-Nearest Neighbors can be sensitive to overfitting and/or outliers. The remaining three methods (Naïve Bayes, Quadratic Discriminant Analysis, and Linear Discriminant Analysis) do make assumptions about data distribution but have shown to be excellent classifiers when these assumptions are meant. All six machine learning methods differ in assumptions, advantages, are limitations, but all have shown utility in either LIBS or ink analysis or both.

A summary of the major foundations, assumptions, advantages and limitations to each of the machine learning methods are shown in Table 9.
Table 9. Summary of the machine learning methods used in the study

<table>
<thead>
<tr>
<th>ML Algorithms</th>
<th>Foundations</th>
<th>Assumptions</th>
<th>Advantages</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Forest</td>
<td>Using outcome of binary tests to determine sample class</td>
<td>Non-parametric, no formal distribution assumptions</td>
<td>Good at many predictors, estimates variable importance, rugged to missing, skewed, multi-modal, continuous, categorical data</td>
<td>Prone to overfitting and sensitive to noisy data</td>
</tr>
<tr>
<td>Naïve Bayes</td>
<td>Maximizing posterior probability to determine sample class</td>
<td>Feature independence</td>
<td>Fast and interpretable, able to classify classes with many features with minimal training, rugged to non-explanatory variables</td>
<td>Independence assumption rarely met</td>
</tr>
<tr>
<td>Neural Network</td>
<td>Using outcome of n node layers to determine sample class</td>
<td>None</td>
<td>Rugged to complex, large datasets, no assumptions, detects all possible interactions between predictor variables</td>
<td>Limited operator input, “black box” makes interpretation difficult, better with larger training datasets</td>
</tr>
<tr>
<td>K-Nearest Neighbors</td>
<td>Using majority rules to determine class of a sample</td>
<td>Similarity can be explained by proximity of data</td>
<td>Easy implementation, new data does not impact accuracy of algorithm</td>
<td>Lower accuracy with large datasets/dimensions, necessary to feature scale to improve accuracy, sensitive to unbalanced data or outliers</td>
</tr>
<tr>
<td>Quadratic Discriminant Analysis</td>
<td>Maximizing between class variability and minimizing within class variability – class covariance</td>
<td>Normality, independence, homoscedasticity</td>
<td>Visually display maximum separation between classes</td>
<td>Accuracy of the predictions dependent on meeting all assumptions</td>
</tr>
<tr>
<td>Linear Discriminant Analysis</td>
<td>Maximizing between class variability and minimizing within class variability – pooled covariance</td>
<td>Normality, independence, homoscedasticity, roughly equal class covariances</td>
<td>Visually display maximum separation between classes</td>
<td>Accuracy of the predictions dependent on meeting all assumptions</td>
</tr>
</tbody>
</table>
2. Objectives of the Project.

2.1. Motivation of the Study and Objective Summary

Visual examination has been the preferred method for the identification of counterfeit packaging due to its simplicity, low cost and non-invasive nature. Nonetheless, this procedure is anticipated to become less effective in the near future as counterfeit packaging becomes more difficult to distinguish due to improvements in the faking technology processes. Therefore, an alternative technique is needed to quickly and accurately discriminate between counterfeit and authentic packaging. Not only should the technique be fast and sensitive, but it should also keep the sample intact and ideally have portable capabilities for on-site characterization.

As a result, this thesis investigates the feasibility of utilizing Laser-Induced Breakdown Spectroscopy (LIBS) and Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy (ATR-FTIR) for the identification of counterfeit pharmaceutical packaging. Previous work (Dégardin et al [87], Subedi et al [111]) has indicated the efficacy of spectroscopic methods to identify pharmaceutical packaging and to discriminate printer inks using LIBS respectively. Therefore, we hypothesize that the combined chemical LIBS and IR profiles of ink and paper packaging will provide accurate discrimination between authentic and counterfeit samples as well as identify features to provide links between counterfeit sources.

This project’s overall objective is to develop and evaluate rapid LIBS and FTIR-ATR methods for the chemical characterization and classification of known authentic and counterfeit pharmaceutical packaging samples, using machine learning classifiers.

Simulated “counterfeit” pharmaceutical packaging samples were created and authentic reference cartons were collected. LIBS and ATR-FTIR were used to identify the inorganic and organic chemical profiles of the samples. Finally, six supervised machine learning methods (Random Forest, Naïve Bayes, Neural Network, k-Nearest Neighbor, Quadratic Discriminant Analysis, and Linear Discriminant Analysis) were used to classify the samples as authentic or counterfeit sources. A graphical representation of the project objectives is shown in Figure 13.
Figure 13. Experimental design of study.

Task 1. Create and visualize a dataset of pharmaceutical packages to assess the utility of LIBS and FTIR for identification of counterfeits.
- \( N_{\text{authentic}} \): 54 paperboard packages
- \( N_{\text{counterfeit}} \): 112 replicate printouts

Task 2. Optimize and validate Laser-Induced Breakdown Spectroscopy for the analysis and classification of authentic and counterfeit pharmaceutical packaging.
- IV: Pharmaceutical packaging source.
- DV: Ink-to-Paper ratio of informative elements.
- \( N_{\text{authentic}} \): 35 black barcodes, 12 multi-color packages
- \( N_{\text{counterfeit}} \): 100 black barcodes, 112 multi-color packages

Task 3. Optimize and validate Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy for the analysis and classification of authentic and counterfeit pharmaceutical packaging.
- IV: Pharmaceutical packaging source.
- DV: Percent transmission.
- \( N_{\text{authentic}} \): 12 multi-color packages
- \( N_{\text{counterfeit}} \): 112 multi-color packages

Task 3. Combine LIBS and ATR-FTIR data to classify authentic and counterfeit pharmaceutical packaging using a more comprehensive chemical profile.
- \( N_{\text{authentic}} \): 12 multi-color packages
- \( N_{\text{counterfeit}} \): 112 multi-color packages
2.2. Task 1. Create and visualize a dataset of pharmaceutical packages to assess the utility of LIBS and FTIR for identification of counterfeits.

2.2.1. Task 1.1. Create a dataset of pharmaceutical packages to assess the utility of LIBS and FTIR for identification of counterfeits.

First, a literature review was conducted to determine common types of pharmaceutical packaging and types of frequently counterfeited pharmaceuticals. It was decided that the study would focus on paperboard packaging due to both the utility of black barcode analysis and multi-color ink analysis. Black barcode analysis was particularly important for this study, as most pharmaceutical packages implement barcode as some form of track-and-trace functionality. If a spectrochemical technique was successful at discriminating between counterfeit and authentic barcodes, it would be favorable for a wider selection of packages. A set of fifty four authentic packages were collected representing the types of commonly counterfeit pharmaceuticals. Of these 54 packages, 41 shared different brands and 46 shared different lots. Thirty five of these packages were used for black barcode analysis and 12 would serve as authentic references for multi-color ink analysis.

Prior to creating the counterfeit dataset, we carried out a review of specific manuscripts focusing in analyzing common printing processes and ink types. Counterfeit printing sources were chosen to mimic how a potential pharmaceutical packaging counterfeiter would recreate their packages. Five sources were ultimately chosen: three different professional printing companies in the Morgantown, WV area (to represent high-quality counterfeits) and two home office printers (to present low-quality counterfeits). Six authentic template references (A1 – A6) were chosen in order to analyze a variety of ink colors. Five of these templates (A1 – A5) were acquired from official secondary packaging schematics from the Drugs@FDA database. These templates were chosen because their authentic reference packages were in the authentic dataset; templates were chosen rather than scans of authentic reference packages to increase the counterfeit package quality. The sixth authentic template (Viagra, A6) was taken from a scan of an authentic reference package; this was done to the frequency of fake Viagra in the counterfeit pharmaceutical industry.
For each authentic template, each counterfeit printing source created four replicate printouts in order to assess intra-source variability. This occurred for authentic templates A1 – A5. The authentic template for A6 was not received until after printing occurred at counterfeit printing sources 1 and 2. In order to ensure consistency between the printouts (as each printout was printed consecutively), counterfeit printing sources 1 and 2 did not print the replicate printouts of A6.

When analyzing the samples, 8 replicates were conducted per ink color/substrate, per sample, per technique. This generated a large number of spectra for analysis, with the prediction that it would allow for a comprehensive chemical characterization of the dataset. It would also allow for a robust training and test set for the classification algorithms.

Overall, the dataset consists of 54 authentic samples, 112 counterfeit samples, 8 colors/substrate. This ultimately generated 3864 LIBS and ATR-FTIR spectral files that were used for analysis.

2.2.2. Task 1.2. Examine the authentic dataset for frequency of anti-counterfeit technologies.

Although several research groups have addressed the composition of anti-counterfeit technologies, none of them have discussed the implication and frequency of anti-counterfeit technologies in authentic pharmaceutical packaging. Therefore, in this thesis a portion of the authentic dataset was examined (18 samples) in order to determine the frequency of “covert”, or unable to be seen by simple visual inspection, anti-counterfeit technologies on the sample.

To this end, a visual spectral comparator (VSC) was used to examine the presence, frequency, and/or characteristics of anti-counterfeit technologies on the paperboard packaging. The samples were analyzed both under flood and transmitted lighting, as well as under UV lighting to examine the presence of any “invisible” ink printed onto the package.
2.3. Task 2. Optimize and validate Laser-Induced Breakdown Spectroscopy for the analysis and classification of authentic and counterfeit pharmaceutical packaging.

2.3.1. Task 2.1. Optimize Laser-Induced Breakdown Spectroscopy for the analysis and classification of authentic and counterfeit pharmaceutical packaging.

LIBS is a very versatile technology, and it has demonstrated to offer broad applicability to a variety of different sample types and materials. However, the technique has several parameters that need to be adjusted before analysis. With this in mind, two procedures were followed: a preliminary rough optimization and a Box-Behnken response surface optimization. The goal during the rough process was to maximize the signal to noise ratio while minimizing the interaction of the laser beam into the paperboard matrix. For optimization, authentic reference template A4 was used. In the preliminary optimization, gate delay, carrier gas identity, laser frequency, laser spot size, and laser % energy were examined. Wide ranges of parameter levels were established based on analysis of ink using LIBS in previous studies.

After the rough preliminary optimization was conducted, a Box-Behnken response surface method was undertaken. This consisted of 27 experiments and tested four parameters: gate delay, carrier gas flowrate, laser frequency, and laser % energy. As it was determined during preliminary optimization that argon yielded superior signal-to-noise to air, carrier gas flowrate was analyzed instead of carrier gas identity. The Box-Behnken response surface method was chosen due to the assumption inherent to the analysis that the optimum parameters are not at the extremes of the tested ranges, and the incorporation of factor-factor effects.

The Box-Behnken response surface plots were analyzed to determine what combination of factor levels yielded the highest maximum signal-to-noise and visual inspection of the ablation occurred to determine what combination of factor levels yielded the lowest ablation depth into the paperboard substrate. After the combination of optimized factor levels was determined, those parameters were used for the validation analysis.
2.3.2. Task 2.2. Validate Laser-Induced Breakdown Spectroscopy for the analysis and classification of authentic and counterfeit pharmaceutical packaging.

The optimized parameters were used from Task 2.1 to validate LIBS’ utility for the analysis of pharmaceutical packaging. This validation consisted of three main subsets: the comparison of within and between source variability of elemental profiles, a 2-class system of black barcodes, and a 6-class system of multi-color inks.

The comparison of within and between source variability of elemental profiles was conducted on the counterfeit dataset. Within source variability was expected to be low for this dataset, as it was known that the replicate printouts were printed consecutively from one another on the same printer for each counterfeit printing sources. However, establishing a false exclusion rate was vital for further classification models.

Prior to the construction of a classification model, the data had to be processed. First, the LIBS spectra were inspected to determine prominent emission lines. The elements responsible for these emission lines were identified through the use of the NIST LIBS database. Finally, the emission lines were integrated and the area for each emission line of interest was calculated. Principal component analysis was performed on each ink color/substrate for each authentic template. Three principal components, accounting for over 85% of the variance in the original dataset, were taken. The principal components were combined for each authentic template before classification occurred. Six machine learning methods were used: Random Forest, Naïve Bayes, Neural Network, k-Nearest Neighbors, Quadratic Discriminant Analysis, and Linear Discriminant Analysis. Misclassification rates were received per each authentic template per each machine learning method.

Further, a 2-class system was analyzed using LIBS. Each sample was classified as either ‘counterfeit’ or ‘authentic’, without taking into account the different counterfeit printing sources. Only the black barcode and paperboard substrate was used for this classification; as explained previously, black barcodes are printed on most pharmaceutical packaging. It was also expected that black barcodes would not reveal much inorganic informative data using LIBS due to organic
nature of the pigment, so black ink was predicted to be the ‘worst case scenario’ for LIBS analysis. In this way, the 2-class system worked as a quick screening method showcasing the utility of LIBS for this type of analysis.

Lastly, a 6-class system was analyzed using LIBS. In this model, the six classes consisted of the authentic class and one of five counterfeit printing classes. Here, multiple ink colors on each authentic template were used for classification. The samples were classified as either ‘authentic’ or one of five counterfeit printing sources; misclassification rates were received for each authentic template per every machine learning method used. This model showed not only the utility of LIBS for the purpose of authentic/counterfeit pharmaceutical packaging analysis, but also the ability of LIBS to discriminate between different counterfeit printing sources. During casework analysis, this would be vital for investigators in order to determine a ‘link’ or common source of origin between counterfeit pharmaceutical packaging samples.


While LIBS was used to reveal inorganic information about the tested specimen, ATR-FTIR was used to display the presence of organic functional groups. Similar to Task 2, a six-class system was constructed. Each sample was classified as either authentic or belonging to one of five counterfeit printing sources using six machine learning methods: Random Forest, Naïve Bayes, Neural Network, k-Nearest Neighbors, Quadratic Discriminant Analysis, and Linear Discriminant analysis.

Prior to classification, the spectra were visually inspected for regions of high variability between sources. It was expected that these regions would provide more discriminating information than the regions of low variability between sources. For this reason, a portion of the IR “fingerprint region”, from 400 – 1400 cm⁻¹, was chosen for analysis. A second derivative transformation was performed on the spectra. Principal component analysis was performed, and one principal
component was chosen for each ink color/paper substrate. All principal components for each authentic template was combined before classification.

After classification, the misclassification rates were received per each authentic template per every machine learning method. They were examined to see if (a) any machine learning method outperformed exceptionally and (b) in general, ATR-FTIR showed an ability to accurately classify between authentic samples and counterfeit pharmaceutical packaging samples.

2.5. Task 4. Combine LIBS and ATR-FTIR data to classify authentic and counterfeit pharmaceutical packaging using a more comprehensive chemical profile.

LIBS and ATR-FTIR were chosen to be complementary techniques to one another. Therefore, the final task in the project was to combine the data from both LIBS and ATR-FTIR to examine whether misclassification rates could be reduced further. The principal components from both LIBS and ATR-FTIR were combined on one spreadsheet. Not only does this fusion the complementary organic and inorganic information from both techniques, but also provides the classification models with more feature variables to analyze.

Like in previous tasks, a six-class system was used to validate the ability of combined LIBS and ATR-FTIR data. Each sample was classified as either authentic or belonging to one of five counterfeit printing sources using six machine learning methods: Random Forest, Naïve Bayes, Neural Network, k-Nearest Neighbors, Quadratic Discriminant Analysis, and Linear Discriminant Analysis.

When the misclassification rates were received, they were compared to the misclassification data from Tasks 2 and 3 to see whether combined, complementary data yielded more accuracy with the classification models. If so, it was determined that combined LIBS/IR analysis would be the superior technique for discriminating between authentic and counterfeit samples.
3. Dataset Creation and Visualization

3.1. Methodology

3.1.1. Unique Identification Tracking Method.

As each sample was collected, they were assigned a unique identification for easy reference in an inventory spreadsheet. An example of ID number allocation is shown in Figure 14. In the inventory, the unique ID, active pharmaceutical ingredient, brand name (if any), manufacturer, and lot number were noted on any received specimen. The samples were placed in a sealed envelope labeled with their ID number and kept in a dark location for storage for room temperature.

![Figure 14. Examples of sample identification system.](image)

The full list of authentic sample IDs, lot numbers, pharmaceuticals, and pharmaceutical companies are listed in Table 10. Gaps between unique identification numbers indicate that the received pharmaceutical package specimen with that ID was not a paperboard package (i.e. the focus of the study). For clarity, authentic templates used for counterfeit production were assigned a secondary ID. A1 is equivalent to A061, A2 is equivalent to A060, A3 is equivalent to A062, A4 is equivalent to OS1 – OS3, A5 is equivalent to A033, and A6 is equivalent to A049 – A053.
Table 10. Full list of authentic dataset information.

<table>
<thead>
<tr>
<th>ID</th>
<th>Lot #</th>
<th>Pharmaceutical</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS1</td>
<td>490048A</td>
<td>Olseletamivir Phosphate</td>
<td>Alvogen</td>
</tr>
<tr>
<td>OS2</td>
<td>490048A</td>
<td>Olseletamivir Phosphate</td>
<td>Alvogen</td>
</tr>
<tr>
<td>OS3</td>
<td>490044A</td>
<td>Olseletamivir Phosphate</td>
<td>Alvogen</td>
</tr>
<tr>
<td>A004</td>
<td>313001A</td>
<td>Desogestrel/Ethinylestradiol</td>
<td>Schering</td>
</tr>
<tr>
<td>A005</td>
<td>6031056</td>
<td>Diazepam</td>
<td>Sanofi</td>
</tr>
<tr>
<td>A006</td>
<td>503137</td>
<td>Flunarizine</td>
<td>Ache</td>
</tr>
<tr>
<td>A007</td>
<td>339701</td>
<td>Amoxicilina</td>
<td>Eurofarma</td>
</tr>
<tr>
<td>A008</td>
<td>13C0043</td>
<td>Levonorgestrel</td>
<td>Neoquimica</td>
</tr>
<tr>
<td>A009</td>
<td>R739425</td>
<td>Fluticasone Furoate</td>
<td>GlaxoSmithKline</td>
</tr>
<tr>
<td>A010</td>
<td>504615</td>
<td>Escitalopram</td>
<td>EMS</td>
</tr>
<tr>
<td>A012</td>
<td>366609</td>
<td>Simeticone</td>
<td>Novaquimica</td>
</tr>
<tr>
<td>A013</td>
<td>BR24327</td>
<td>Metformin</td>
<td>Merck</td>
</tr>
<tr>
<td>A014</td>
<td>317873</td>
<td>Oxymetazoline</td>
<td>EMS</td>
</tr>
<tr>
<td>A015</td>
<td>CT1343</td>
<td>Betamethasone</td>
<td>GSK</td>
</tr>
<tr>
<td>A033</td>
<td>18P0255</td>
<td>Methylprednisolone</td>
<td>Jubilant Cadista Pharmaceuticals Inc.</td>
</tr>
<tr>
<td>A034</td>
<td>GC099T</td>
<td>Allegra</td>
<td>Sanofi</td>
</tr>
<tr>
<td>A035</td>
<td>8GE1927</td>
<td>Ibuprofen</td>
<td>CVS</td>
</tr>
<tr>
<td>A036</td>
<td>602427</td>
<td>Isoprinosine Metisoprinol</td>
<td>Newport Pharmaceuticals International</td>
</tr>
<tr>
<td>A038</td>
<td>PA009RH1</td>
<td>Panadol Ultra</td>
<td>GSK</td>
</tr>
<tr>
<td>A039</td>
<td>5D059B</td>
<td>Kamillosan</td>
<td>MEDA</td>
</tr>
<tr>
<td>A040</td>
<td>7EE1739</td>
<td>Migraine</td>
<td>Novartis</td>
</tr>
<tr>
<td>A041</td>
<td>1504242E</td>
<td>Zithromax</td>
<td>Pfizer</td>
</tr>
<tr>
<td>A042</td>
<td>UFDDA8072A</td>
<td>Cefdinir</td>
<td>Aurobindo Pharma</td>
</tr>
<tr>
<td>A043</td>
<td>33815656A</td>
<td>Junel</td>
<td>TEVA Pharm</td>
</tr>
<tr>
<td>A044</td>
<td>11501771</td>
<td>Koflet</td>
<td>Himalaya Herbal Healthcare</td>
</tr>
<tr>
<td>A045</td>
<td>11602599</td>
<td>Koflet</td>
<td>Himalaya Herbal Healthcare</td>
</tr>
<tr>
<td>A046</td>
<td>1170383</td>
<td>Koflet</td>
<td>Himalaya Herbal Healthcare</td>
</tr>
<tr>
<td>A047</td>
<td>AAS082A</td>
<td>Fluticasone Propionate</td>
<td>West-Ward Hikma Company</td>
</tr>
<tr>
<td>A048</td>
<td>UL80112</td>
<td>Oseltamivir Phosphate</td>
<td>Zydus</td>
</tr>
<tr>
<td>A049</td>
<td>108099</td>
<td>Viagra</td>
<td>Wyeth</td>
</tr>
<tr>
<td>A050</td>
<td>108099</td>
<td>Viagra</td>
<td>Wyeth</td>
</tr>
<tr>
<td>A051</td>
<td>108099</td>
<td>Viagra</td>
<td>Wyeth</td>
</tr>
<tr>
<td>A052</td>
<td>108099</td>
<td>Viagra</td>
<td>Wyeth</td>
</tr>
<tr>
<td>A053</td>
<td>108099</td>
<td>Viagra</td>
<td>Wyeth</td>
</tr>
<tr>
<td>A054</td>
<td>C966853</td>
<td>Cialis</td>
<td>Eli Lilly</td>
</tr>
<tr>
<td>A055</td>
<td>C966853</td>
<td>Cialis</td>
<td>Eli Lilly</td>
</tr>
<tr>
<td>A056</td>
<td>C988839</td>
<td>Cialis</td>
<td>Eli Lilly</td>
</tr>
<tr>
<td>A057</td>
<td>C988839</td>
<td>Cialis</td>
<td>Eli Lilly</td>
</tr>
<tr>
<td>A058</td>
<td>C988839</td>
<td>Cialis</td>
<td>Eli Lilly</td>
</tr>
<tr>
<td>A059</td>
<td>UFDDA8126A</td>
<td>Cefdinir for Oral Suspension</td>
<td>Aurobindo Pharma</td>
</tr>
<tr>
<td>A060</td>
<td>G806703</td>
<td>Azithromycin</td>
<td>Lupin</td>
</tr>
<tr>
<td>A061</td>
<td>MFC006</td>
<td>Sudafed Day+Night</td>
<td>Johnson/Jonson</td>
</tr>
<tr>
<td>A062</td>
<td>C01787</td>
<td>Unisom</td>
<td>Chattem</td>
</tr>
<tr>
<td>A175</td>
<td>362615</td>
<td>Espidifen</td>
<td>Zambon</td>
</tr>
<tr>
<td>A176</td>
<td>6CT0491</td>
<td>Equaline</td>
<td>MPS</td>
</tr>
<tr>
<td>A177</td>
<td>5990</td>
<td>Little Remedies</td>
<td>Medtech</td>
</tr>
<tr>
<td>A178</td>
<td>768527599</td>
<td>Clotrimazole</td>
<td>Taro</td>
</tr>
<tr>
<td>A179</td>
<td>16006</td>
<td>Clotrimazole</td>
<td>Walgreen's</td>
</tr>
<tr>
<td>A180</td>
<td>EH1344</td>
<td>triamcinolone acetonide</td>
<td>Fougera</td>
</tr>
<tr>
<td>A181</td>
<td>L6MZ1</td>
<td>Naproxen</td>
<td>Normon</td>
</tr>
<tr>
<td>A182</td>
<td>YCTA9407</td>
<td>Polaramine</td>
<td>Merck</td>
</tr>
<tr>
<td>A183</td>
<td>048740B</td>
<td>Protopic</td>
<td>Leo</td>
</tr>
<tr>
<td>A184</td>
<td>DAB44a</td>
<td>ProAir HFA</td>
<td>Teva</td>
</tr>
<tr>
<td>A185</td>
<td>L011</td>
<td>Peitel</td>
<td>Ferrer</td>
</tr>
</tbody>
</table>
Accordingly, the full description of the counterfeit dataset per printing source is listed in Table 11. Again, for clarity, printing sources were given a secondary ID. Morgantown Blueprint is equivalent to Counterfeit Printing Source 1 (C1), Morgantown Printing and Binding is equivalent to Counterfeit Printing Source 2 (C2), UPS is equivalent to Counterfeit Printing Source 3 (C3), Downtown Library is equivalent to Counterfeit Printing Source 4 (C4), and Bell Lab is equivalent to Counterfeit Printing Source 5 (C5). Of these, it is important to note that C1 – C3 were classified as ‘professional printing companies’ and C4 – C5 were classified as ‘personal office printers’.

Table 11. Full description of counterfeit printing sources. Replicate printouts grouped as a range.

<table>
<thead>
<tr>
<th>Printing Source</th>
<th>Replicate Printouts</th>
<th>Authentic Reference</th>
<th>Total printouts:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morgantown Blueprint</td>
<td>Q063 - Q066</td>
<td>A1</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Q067 - Q070</td>
<td>A3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q071 - Q074</td>
<td>A2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q075 - Q078</td>
<td>A4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q079 - Q082</td>
<td>A5</td>
<td></td>
</tr>
<tr>
<td>Morgantown Printing and Binding</td>
<td>Q083 - Q086</td>
<td>A1</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Q087 - Q090</td>
<td>A3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q091 - Q094</td>
<td>A2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q095 - Q098</td>
<td>A4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q099 - Q102</td>
<td>A5</td>
<td></td>
</tr>
<tr>
<td>UPS</td>
<td>Q103 - Q106</td>
<td>A1</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Q107 - Q110</td>
<td>A3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q111 - Q114</td>
<td>A2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q115 - Q118</td>
<td>A4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q119 - Q122</td>
<td>A5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q123 - Q126</td>
<td>A6</td>
<td></td>
</tr>
<tr>
<td>Downtown Library</td>
<td>Q127 - Q130</td>
<td>A1</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Q131 - Q134</td>
<td>A3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q135 - Q138</td>
<td>A2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Q139 - Q142</td>
<td>A4</td>
<td></td>
</tr>
</tbody>
</table>
3.1.2. Video Spectral Comparator (VSC)

A Foster & Freeman VSC 6000 (Foster and Freeman, VA, USA) docucenter was used to visually inspect the authentic sample dataset. Images of the packages were taken using the camera incorporated into the instrument using white flood lighting, before the instrument was switched into UV light (312 nm and 365 nm). Any anti-counterfeit technologies, such as “invisible” ink or holograms, visible under the UV light were noted and photographed.

3.2. Results and Discussion

3.2.1. Authentic Dataset

Fifty-four authentic pharmaceutical cartons were collected. The medical purpose of these pharmaceuticals can be seen in Table 12 and represent many classes of commonly counterfeit pharmaceuticals.
Table 12. Medical purposes for authentic pharmaceutical dataset.

<table>
<thead>
<tr>
<th>Medical Purpose</th>
<th>Count</th>
<th>Medical Purpose</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vasodilator</td>
<td>10</td>
<td>Immunosuppressant</td>
<td>1</td>
</tr>
<tr>
<td>Antibiotic</td>
<td>8</td>
<td>Anti-Anxiety</td>
<td>1</td>
</tr>
<tr>
<td>Steroid</td>
<td>7</td>
<td>Anti-Depressant</td>
<td>1</td>
</tr>
<tr>
<td>Anti-Viral</td>
<td>6</td>
<td>Anti-Foaming</td>
<td>1</td>
</tr>
<tr>
<td>Anti-Migraine/Anti-Pain</td>
<td>4</td>
<td>Anti-Fungal</td>
<td>1</td>
</tr>
<tr>
<td>Hormone</td>
<td>3</td>
<td>Anti-Inflammatory</td>
<td>1</td>
</tr>
<tr>
<td>Antihistamine</td>
<td>3</td>
<td>Bronchodilator</td>
<td>1</td>
</tr>
<tr>
<td>NSAID</td>
<td>3</td>
<td>Disinfectant</td>
<td>1</td>
</tr>
<tr>
<td>Decongestant</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Of these fifty-four samples, there were smaller sub-sets of 5, 3, 2, and 2 boxes. Each sub-set contained the same lot number and were used to assess intra-lot and inter-lot variability. The remaining samples differed in manufacturer, product name, or lot number. Each sample was visually inspected to see which/how many ink colors were present in abundance for analysis. If a sample had a black barcode, it was included in the analysis of a barcode authentic subset. This analytical scheme can be seen represented visually in
Figure 15.
3.2.2. Simulated Counterfeit Dataset.

From the authentic dataset, six authentic reference templates (schematics) were obtained to create a simulated counterfeit dataset. The medical purposes of these schematics can be seen in Table 13. One hundred and twelve simulated counterfeit pharmaceutical samples were created using these six authentic reference templates. They originated from five printing sources. Three of these printing sources were professional printing companies and two were office-use printers, all of which utilized toner ink.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Medical Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Decongestant</td>
</tr>
<tr>
<td>A2</td>
<td>Antibiotic</td>
</tr>
<tr>
<td>A3</td>
<td>Antihistamine</td>
</tr>
<tr>
<td>A4</td>
<td>Antiviral</td>
</tr>
<tr>
<td>A5</td>
<td>Steroid</td>
</tr>
<tr>
<td>A6</td>
<td>Vasodilator</td>
</tr>
</tbody>
</table>

Table 13. Medical purposes for counterfeit pharmaceutical samples.

Each schematic had four replicate copies from each printing source, with two exceptions. Authentic A6 was only printed from three printing sources (one professional company, two office
printers). Each color on each replicate copy was randomly analyzed in eight replicate analytical runs for both LIBS and ATR-FTIR. In LIBS, the two-class system utilized 2160 authentic and counterfeit spectra from black barcodes of 35 authentic sources and 100 counterfeit sources. The six-class system for LIBS utilized 2872 spectra (2656 counterfeit, 216 authentic) and 3864 for ATR-FTIR (3552 counterfeit, 312 authentic) from 6 authentic sources (12 sources considering the intra-lot sub-sets) and 112 counterfeit printouts.

Additionally, an example of an authentic reference sample (A4) and a printout from counterfeit printing sources C1-C5 are shown in Figure 16. Visual display of authentic reference A4 and printouts from C1-C5. Analyzed colors were blue, yellow, black (barcode), and paperboard substrate. The analytical scheme for the counterfeit samples is shown in Figure 17.

Figure 16. Visual display of authentic reference A4 and printouts from C1-C5. Analyzed colors were blue, yellow, black (barcode), and paperboard substrate.
Figure 17. Description for counterfeit pharmaceutical dataset. Colors analyzed for counterfeit/authentic comparison are shown (gray = paperboard substrate).
3.2.3. VSC Analysis of Authentic Samples.

A portion of the dataset was investigated to determine the relative presence of anti-counterfeit technologies on the packaging material. Out of the sixteen pharmaceutical brands examined under UV light, fourteen exhibited “invisible” ink anti-counterfeit technology printed on the packaging. These generally took the form of either the manufacture logo or a ‘quality/security’ design on a specified part of the package. Examples of the anti-counterfeit technologies on two packages are located in the appendix.

3.3. Conclusion

It was determined that the authentic and counterfeit pharmaceutical package dataset was suitable for the project with regards to both sample size and type of pharmaceutical collected. After the dataset was created, a chemical analysis was first conducted using the LIBS method on black barcodes to assess within and between samples variability of elemental profiles and to optimize the method. Analysis was then expanded to multi-colored areas of the package by both LIBS and FTIR-ATR. In terms of VSC analysis, it was shown that the majority of the authentic samples (87.5%) tested exhibited some form of invisible ink on the packaging, but the design of the security ink was not particularly complex on most of the examined samples.
4. Method Optimization and Validation for Laser-Induced Breakdown Spectroscopy

4.1. Methodology

4.1.1. Instrument Specifications

The LIBS used in this project was a J200 Applied Spectra (Sacramento, CA, P/N J200T – 0266-EC6-IEPZBC) with a Q-switched 10ns-266 nm Nd:YAG laser source and a six channel Czerny-Turner spectrometer and Charge Coupled Device (CCD) detector. The fixed instrument specifications were 0.1 nm resolution, 1.05 ms gate width, and detection spectral range from 190 nm to 1040 nm.

4.1.2. Technique Optimization

4.1.2.1. Preliminary Optimization

The LIBS instrument has several adjustable parameters that affect the efficiency of laser-to-sample interaction and the detection of ionic and atomic emission lines. Before a detailed optimization was conducted via the Box Behnken method, a preliminary rough optimization was conducted with wider parameter setting to determine, in general, the optimal range settings for the particular application at hand. Three criteria were sought in this optimization: high signal to noise, low percent relative standard deviation within replicate measurements, and minimal paperboard ablation so the chemical signature originates primarily from the ink layer. For all optimization tests, the yellow ink of an authentic sample of Pharmaceutical B (
Figure 15) was utilized, as optimization settings were not expected to change significantly based on ink color.

Five factors were tested during preliminary optimization: gate delay, carrier gas identity, frequency, laser energy, and spot size. After the analytical tests were conducted, the preliminary optimized combination of factors was noted. The tested and optimized parameters are listed in Table 14.

Table 14. Tested parameters for preliminary LIBS optimization.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Level tested</th>
<th>Optimized Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gate Delay</td>
<td>0.5, 1, 1.5 μs</td>
<td>1 μs</td>
</tr>
<tr>
<td>Carrier Gas Identity</td>
<td>Argon, Air</td>
<td>Argon</td>
</tr>
<tr>
<td>Frequency</td>
<td>2, 5, 8, 10 Hz</td>
<td>5 Hz</td>
</tr>
</tbody>
</table>
4.1.2.2. Surface response optimization using Box-Behnken design.

In this study, a Box-Behnken design was applied for the fine-tune-optimization of the LIBS method. Box-Behnken used is a quadratic design considering several treatment combinations in the model to find the optimal regions. The Box Behnken design is efficient in systems where parameter adjustments do not affect the desired response (e.g., signal-to-noise values) independently from one another. One assumption in Box Behnken is that the optimal combination of parameters is not at parameter extremes. For this project, this assumption was considered by selecting ranges for the factor levels in which the midpoint was set as the optimal parameter determined during the preliminary exploratory optimization.

This design requires three levels for each factor of interest. Therefore, each tested parameter, called factors, was assigned three levels at low, medium, and high values. JMP Pro (vers. 14) was used to construct a method of 27 analytical trials consisting of different combinations of factor levels. These trials are shown in Table 15.

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Frequency (Hz)</th>
<th>Energy (%)</th>
<th>Ar Flowrate (L/min)</th>
<th>Gate Delay (us)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0+0−</td>
<td>5</td>
<td>25</td>
<td>0.8</td>
<td>0.5</td>
</tr>
<tr>
<td>00++</td>
<td>5</td>
<td>20</td>
<td>0.5</td>
<td>1.5</td>
</tr>
<tr>
<td>−0−0</td>
<td>2</td>
<td>20</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>0++0</td>
<td>5</td>
<td>25</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>0000</td>
<td>5</td>
<td>20</td>
<td>0.8</td>
<td>1.0</td>
</tr>
<tr>
<td>+0+0</td>
<td>10</td>
<td>20</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>+0−0</td>
<td>10</td>
<td>10</td>
<td>0.8</td>
<td>1.0</td>
</tr>
<tr>
<td>+00+</td>
<td>10</td>
<td>20</td>
<td>0.8</td>
<td>1.5</td>
</tr>
</tbody>
</table>
Although the resultant Box Behnken 3D surface graphs display the combination of factor levels that yielded the highest signal-to-noise responses, there was another objective during the optimization process: minimal ablation into the paperboard substrate. This was determined through visual inspection of the sample after the ablation process.

After the preliminary optimization, critical factors such as argon as the carrier gas and spot size were fixed and the remaining four factors were tested: gate delay, carrier gas flowrate, frequency, and % laser energy. The tested and optimized parameters are shown in Table 16.

<table>
<thead>
<tr>
<th>Parameter (factor)</th>
<th>Level (low, medium, high)</th>
<th>Optimized Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gate Delay</td>
<td>0.5, 1, 1.5 μs</td>
<td>0.9 μs</td>
</tr>
<tr>
<td>Carrier Gas Flowrate</td>
<td>0.5, 0.8, 1 L/min</td>
<td>0.8 L/min</td>
</tr>
<tr>
<td>Frequency</td>
<td>2, 5, 10 Hz</td>
<td>8 Hz</td>
</tr>
</tbody>
</table>

Table 16. Tested parameters for Box Behnken optimization of LIBS.
### 4.1.3. Data Processing and Machine Learning Analysis

Once the spectra were collected through the Axiom software in the LIBS, it was visually examined with the Aurora (v 18.0) data analysis software. Background noise was subtracted from the resultant spectra. Main emission lines were examined, and the occurrence of specific elements was determined when at least 2 major lines per element were present. After this process was completed, the peaks were integrated to receive a signal for that emission line. Noise was determined using the same integrated area for the paperboard spectra of that sample, resulting in the calculated signal-to-noise. This was done as the desired information came from the ink only, so all extraneous data was treated as noise. For this reason, the paperboard substrate was not included in the LIBS analysis except for the binary classification system. Only elements present at signal-to-noise greater than 3 were used for further exploratory, comparative, and classification analysis.

After the elemental integrated data was calculated, principal component analysis was performed. Three principal components (encompassing over 80% of original data variance) were taken for each ink type. All of the principal components for each authentic template were combined for analysis.

In this study, two datasets were analyzed and compared, one consisting of black barcodes and a second one considering multiple colored inks from the packages. The first comparison system was binary in nature. All authentic samples that contained a black ink barcode and all corresponding counterfeit samples were compared, resulting in two classes: counterfeit or authentic. The principal components for the paperboard substrate and black ink was combined for classification of the samples. The majority of examined pharmaceutical packages possessed, at a minimum, a black barcode and paperboard substrate; this classification specifically investigated the utility of LIBS as a quick screening tool for differentiating between authentic and counterfeit samples.
The second comparison model consisted of six classes. The five counterfeit printing sources (C1 to C5) were compared against their corresponding authentic sample (A). More colors (red, blue, yellow, green, pink, brown, and black) were investigated in this model, but analysis of the paperboard substrate alone was only considered during the estimation of ink to paper ratios. The colors analyzed in each dataset is shown in Figure 17. This was conducted not only to examine the efficacy of LIBS at distinguishing counterfeits and authentic samples, but also different counterfeit sources from one another. This could have relevance in real casework, when investigators are attempting to determine a link between the source of counterfeit pharmaceutical manufacture for intelligence purposes.

After integrated data was received and the ink-to-paper ratio calculated, the information was input directly into JMP Pro 14 analysis software where PCA was performed. Six supervised machine learning methods were used: Random Forest, Naïve Bayes, Neural Network, k-Nearest Neighbor, Quadratic Discriminant Analysis, and Linear Discriminant Analysis. For all of these models, a random 60/40 train/test split ratio was used.

4.2. Results and Discussion

4.2.1. Optimization

Two main criteria were used to select the optimal parameters: sampling of the ink with minimal penetration into the substrate and maximum signal to noise. Also, the repeatability of the measurements was monitored as percent relative standard deviation to verify the optimal parameters won’t exceed 20% RSD. The first criterium was determined through visual inspection of the paper. An example of the ablation is shown in Figure 18 after eight analytical runs. The ink is clearly ablated away from the sample, but there is minimal destruction of the paperboard substrate. The effectiveness of the ablation parameters was confirmed by the different elemental profiles observed for ink and paper substrate.
The second criterium of optimal signal response was analyzed through the use of response surface graphs. This is particularly advantageous in displaying the interaction between factors, as it cannot be assumed that the factors are necessarily independent from one another.

Response surface methods (RSM) are useful to find improved or optimal parameter settings in an experiment. A response surface design is often preferred over methods that change a variable at a time because it is more cost-effective as the number of runs needed to find the optimal values are reduced while considering the effects of interactions between variables. An example of such a graph is shown in Figure 19 displaying the signal-to-noise values when considering laser frequency and laser energy for an emission line of calcium. The maximum signal-to-noise value of the graph yields is represented in the darker red regions of the surface response.
4.2.2. Spectral Analysis and Element Identification

The LIBS spectra yielded several emission lines characteristic of inorganic elements, such as the emission lines shown in Figure 20. LIBS spectra displaying the black ink for authentic reference samples A1 and A2, Figure 21, and Figure 22, highlighting emission lines at 285 nm (magnesium), 288 nm (silicon), 396 nm (aluminum), and 397 nm (calcium).
Figure 20. LIBS spectra displaying the black ink for authentic reference samples A1 and A2.
Figure 21. LIBS spectra indicating black ink on authentic reference samples A3 and A4.
Figure 22. LIBS spectrum indicating black ink on authentic reference A5.
The informative elements found within the ink, as determined by examination of the LIBS spectra, is shown in Table 17. Other research studies have confirmed that a majority of these elements have been identified in ink formulations [95, 111].

Table 17. Informative elements for analyzed ink colors and substrate.

<table>
<thead>
<tr>
<th>Ink Color</th>
<th>Informative Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper Carton</td>
<td>Al, Ca, Mg, Sr</td>
</tr>
<tr>
<td>Black</td>
<td>Al, Ca, Cu, K, Mg, Si, Ti</td>
</tr>
<tr>
<td>Blue</td>
<td>Al, Ba, Ca, Cd, Cr, Cu, Fe, K, Mg, Mn, Ni, Si, Sr, Ti</td>
</tr>
<tr>
<td>Yellow</td>
<td>Al, Ba, Ca, Cd, Cr, Cu, Fe, K, Mg, Mn, Ni, Si, Sr, Ti</td>
</tr>
<tr>
<td>Green</td>
<td>Al, Ba, Ca, Cd, Cr, Cu, Fe, K, Mn, Ni, Sr, Ti</td>
</tr>
<tr>
<td>Pink</td>
<td>Al, Ba, Ca, Cd, Cr, Cu, Fe, K, Mg, Mn, Ni, Si, Sr, Ti</td>
</tr>
<tr>
<td>Brown</td>
<td>Al, Ba, Ca, Cd, Cr, Cu, Fe, K, Mn, Si, Ti</td>
</tr>
<tr>
<td>Red</td>
<td>Ba, Ca, Cd, Cr, Cu, Fe, K, Mn, Ni, Si, Sr, Ti</td>
</tr>
</tbody>
</table>

4.2.2. Intra-Sample Variance

4.2.2.1. Intra-Lot Authentic Variance

Pharmaceuticals A, B, and C (
Figure 15) were analyzed for intra-lot and inter-lot variability. Figure 23 depicts boxplots of intra-lot pharmaceutical samples. The boxplots have high overlap, which preliminarily indicate that there is little difference in composition between samples that share a same lot for the monitored elements.
Figure 23. Box plots depicting intra-source variability in same-lot pharmaceutical samples: titanium in red ink (Pharmaceutical A), cadmium in blue ink (Pharmaceutical B), and calcium in paper substrate (Pharmaceutical C),
ANOVA (Kruskal-Wallis with Steel-Dwass post-hoc test, $\alpha = 0.01$) was also used to perform a pairwise comparison, in order to acquire a statistical significance about the variability between the samples. Due to the low sample sizes for some lot subsets, this was only calculated for Lot 1 of Pharmaceutical C ($n=5$). The relevant results and resulting false exclusion rate are shown in Table 18. Many of these differences in variance result from signal-to-noise values that are roughly around the detection limit (SNR ~ 3), where more fluctuation in values is expected. However, false exclusion rates for intra-lot samples is still relatively low. The number of pairwise comparisons conducted was calculated using Equation 3 and multiplied by the number of ink colors/substrate analyzed. The false exclusion rate was then determined by taking the ratio of statistically significant instances of variability and total number of comparisons.

Equation 3. Mathematical calculation for number of pairwise comparisons, where $n$ is the number of samples in the subset.

$$\# \text{Pairwise Comparisons} = \frac{n(n-1)}{2}$$

Table 18. False exclusion rates within same-lot authentic samples.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Lot #</th>
<th>Instances of Intra-Lot Variability</th>
<th>Total Comparisons</th>
<th>False Exclusion Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmaceutical C</td>
<td>Lot 1</td>
<td>5</td>
<td>40</td>
<td>12.5%</td>
</tr>
</tbody>
</table>

4.2.2.2. Inter-Lot Authentic Sample Variance

For the pharmaceuticals that had subsets with different lots), inter-lot variability was also evaluated. Lot 2 of Pharmaceutical B only had one sample in the subset; accordingly, Pharmaceutical B was excluded from inter-lot sample variance analysis. Pharmaceutical A had five instances of intra-lot sample variability: aluminum and calcium in black ink, and nickel, manganese, and calcium in yellow ink. Every instance excepting nickel and calcium in yellow ink
yielded values just below the alpha level (0.0083), and the signal-to-noise values for nickel were close in value to the limit of detection (<5 S/N).

4.2.2.3. Intra-Source Counterfeit Sample Variance

The intra-source variation for the counterfeit samples was expected to be lower than the authentic, as the counterfeit samples were known to be printed consecutively. Boxplots were created to examine the variation between replicate copies from the same printing source (separate ID given, e.g. the second replicate copy from printing source 2 is C2.2), as well as ANOVA pairwise comparisons. The variation is displayed graphically in Figure 24. The interquartile plots display large overlap between the samples of the same source, indicating (preliminary) relatively little variation.

The false exclusion rates (as calculated by the Steel-Dwass test, alpha = 0.01) are explicitly calculated in Error! Reference source not found., ranging from 0% - 18%. It was observed that most of these false exclusions arose from a singular printing source (C3, commercial printing company), indicating larger heterogeneity in their printing process and ink composition. All other printing sources (C1,C2, C3 and C5) presented lower false exclusion rates ranging from 0.0 - 10.0%.
Figure 24. Top: Boxplot of variance of cadmium content in the blue ink of counterfeit counterpart 1 within counterfeit printing sources. Bottom: ANOVA plot of variance of cadmium content in blue ink of counterfeit counterpart 1, C2.
Figure 25. False exclusion rates for two authentic templates (A1 and A2) from five counterfeit printing sources.
Figure 26. False exclusion rates for two authentic templates (A3 and A4) from five counterfeit printing sources.
Figure 27. False exclusion rates for two authentic templates from five (A5) or three (A6) counterfeit printing sources.
In summary, it was determined that both authentic and counterfeit samples exhibited lower intra-source variance as compared to inter-source variance. It was predicted that inter-class variance (variance between authentic and counterfeit sources, as well as between different counterfeit printing sources) would be much higher than intra-class variance.

4.2.3. Binary Classification: Authentic and Counterfeit

First, a two-class system was tested with the six machine learning methods. All counterfeit samples were grouped as the ‘counterfeit’ class (n = 100 for black barcodes), and all authentic samples were grouped as the ‘authentic’ class (n = 35). Most pharmaceutical packaging cartons have a black barcode printed on them as part of the manufacturing process. Black ink was also predicted to be the most difficult ink to discriminate for LIBS due to the mostly organic nature of the carbon black pigment. This two-class system was simulated to be the ‘worst case scenario’ for pharmaceutical package discrimination, and LIBS was tested as a screening method to determine whether it could quickly discriminate between counterfeit and authentic samples without regard to counterfeit sample printing source.

The resulting misclassification rates are shown in Table 19. Overall, the misclassification rates for this system were within acceptable parameters (<7.2%), rendering it extremely suitable to quickly screen between authentic and counterfeit ink and paper samples. Parametric (NB, QDA, LDA) classifiers performed similarly to nonparametric classifiers (RF, NN, KNN). If LDA is excluded from analysis, the misclassification rates for the model are all below 3.0%. LDA’s higher misclassification rate may be attributed to large differences in covariance between the questioned and authentic class.
Table 19. Misclassification rates for a two-class system using combined black ink and paperboard substrate LIBS data (RF: Random Forest, NB: Naïve Bayes, NN: Neural Network, KNN: \(k\)-Nearest Neighbor, QDA: Quadratic Discriminant Analysis, LDA: Linear Discriminant Analysis).

<table>
<thead>
<tr>
<th>ML Technique</th>
<th>Misclassification Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>2.6%</td>
</tr>
<tr>
<td>NB</td>
<td>2.8%</td>
</tr>
<tr>
<td>NN</td>
<td>0.7%</td>
</tr>
<tr>
<td>KNN</td>
<td>1.2%</td>
</tr>
<tr>
<td>QDA</td>
<td>2.1%</td>
</tr>
<tr>
<td>LDA</td>
<td>7.8%</td>
</tr>
</tbody>
</table>

The accompanying LDA canonical plot is shown in Figure 28, showing separation between the counterfeit and authentic class.

![LDA Canonical plot for 2-class system using LIBS black ink and paperboard substrate.](image)

Figure 28. LDA Canonical plot for 2-class system using LIBS black ink and paperboard substrate.

As indicated by the results of the classification algorithms and the clear separation of the two classes on the canonical plot, it is possible to utilize the inorganic elemental data given from a LIBS spectra to quickly identify pharmaceutical packaging as either counterfeit or authentic, without noting counterfeit source.
4.2.4. Six Class Classification: Authentic and Counterfeit Sources

Subsequently, the counterfeit samples from five different sources were compared against their authentic counterpart, resulting in a six-class system. This particular system is advantageous for investigators attempting to find a link, or common source, between counterfeit pharmaceutical packaging. First, most of the variance of the data was captured using PCA. An example of a PCA plot displaying class separation for green ink on A2 is shown in Figure 29.

![Figure 29. PCA plot displaying maximized variance between the six classes, utilizing the second principal component of the green ink on A2.](image)
The resultant misclassification rates are shown in Table 20. They are promisingly low for a screening method. The average misclassification rates all fall below 30%, with five out of the six datasets falling below 22%.

In analyzing the different machine learning methods, it is worthwhile to note that there were no extreme outliers in performance. For each authentic sample, the misclassification rates typically fall within 15.0% - 20.0% of one another. The misclassification rates are more dependent on the authentic sample rather than the machine learning method. Additionally, there is no obvious performance difference between parametric and non-parametric methods within this system.

Table 20. Six-class system: misclassification rates of LIBS data using six machine learning methods.

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>16.4%</td>
<td>5.9%</td>
<td>28.4%</td>
<td>18.9%</td>
<td>17.9%</td>
<td>29.5%</td>
<td>19.5%</td>
</tr>
<tr>
<td>NB</td>
<td>9.0%</td>
<td>4.5%</td>
<td>32.8%</td>
<td>13.5%</td>
<td>19.4%</td>
<td>31.4%</td>
<td>17.9%</td>
</tr>
<tr>
<td>NN</td>
<td>10.4%</td>
<td>0.0%</td>
<td>20.8%</td>
<td>14.9%</td>
<td>13.4%</td>
<td>18.5%</td>
<td>13.0%</td>
</tr>
<tr>
<td>KNN</td>
<td>20.9%</td>
<td>3.0%</td>
<td>22.3%</td>
<td>16.2%</td>
<td>17.9%</td>
<td>25.9%</td>
<td>17.7%</td>
</tr>
<tr>
<td>QDA</td>
<td>23.8%</td>
<td>15.0%</td>
<td>40.3%</td>
<td>21.6%</td>
<td>34.3%</td>
<td>27.8%</td>
<td>27.1%</td>
</tr>
<tr>
<td>LDA</td>
<td>13.4%</td>
<td>6.0%</td>
<td>29.8%</td>
<td>18.9%</td>
<td>31.3%</td>
<td>27.8%</td>
<td>21.2%</td>
</tr>
</tbody>
</table>

Figure 30 displays the QDA canonical plot showing separation for samples from authentic source A2 and the corresponding counterfeit sources. Although some classes exhibit wide separation from one another (C3 and A), others (C2 and C4) were unable to be fully separated through the use of inorganic elemental data alone.
Although the misclassification rates were higher for the six-class system than the two-class system (as to be expected, given the additional opportunities for confusion of between counterfeit classes), the average misclassification rates across the datasets fell below 30%. Most of these misclassifications occurred due the confusion of different counterfeit sources, rather than the confusion of an authentic sample as a counterfeit sample or vice versa. For a screening method, these results indicate that LIBS is appropriate for the quick screening of packaging materials.
LIBS provided an efficient, fast, sensitive technique for the quick screening of pharmaceutical packaging. The technique was optimized through Box Behnken to yield the maximum signal-to-noise while causing minimal destruction to the paperboard sample. Intra-lot and inter-lot variability were assessed for authentic sources, while intra-printing source variability was assessed for counterfeit samples. All three sources of intra-sample variability were determined to be much smaller than inter-sample variability between different classes (authentic and counterfeits), demonstrating the utility of elemental profiles for characterization and classification of packages. The dimensionality of the system was reduced using Principal Component Analysis. For classification, six machine learning methods were used: Random Forest, Naïve Bayes, Neural Network, k-Nearest Neighbors, Quadratic Discriminant Analysis, and Linear Discriminant Analysis. The machine learning algorithms were trained in a 60/40 train/test split. Initially, a two-class system was used consisting of all the authentic and counterfeit samples collected that discriminated between authentic and counterfeit samples using black ink and paperboard substrate. As significant variation between inter-product samples was not expected for black ink and paperboard substrate due to their formulation, it was considered the “worst case scenario” for pharmaceutical screening analysis. Over 2000 spectra were analyzed in this system, with resulting misclassification rates below 8.0%. This indicates that, even in a non-ideal situation for ink analysis, LIBS is still a quick, effective screening tool for pharmaceutical products.

Finally, a six-class system was used that focused on the comparisons of the counterfeit sources to their respective authentic counterparts. Seven ink colors (red, blue, yellow, green, pink, brown, and black) were tested across all samples. LIBS was generally capable of distinguishing between both authentic and counterfeit samples as well as different counterfeit sources. Over 2500 spectra were analyzed in this system, with resulting misclassification rates generally less than 30.0% for all ink colors.
5. Method Validation of Attenuated Total Reflectance – Fourier Transform Infrared Spectroscopy

5.1. Methodology

5.1.1. Instrument Specifications

The instrument used in the study was a Perkin Elmer FT-IR Spectrum Two Spectrometer (Massachusetts, USA) with a micro-ATR accessory attached to the base. The relevant instrument parameters are the range (400 - 4000 cm\(^{-1}\)) and resolution (1 cm\(^{-1}\)). The Spectrum Two spectrometer has a lithium tantalate detector. The experimental parameters used in the study are shown in Table 21.

<table>
<thead>
<tr>
<th>Scan Speed</th>
<th>1 cm/s</th>
<th>Accumulations</th>
<th>4 Scans</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution</td>
<td>4 cm(^{-1})</td>
<td>Bounces</td>
<td>1</td>
</tr>
<tr>
<td>Crystal Type</td>
<td></td>
<td>Diamond</td>
<td></td>
</tr>
</tbody>
</table>

5.1.2. Data Pre-Processing and Statistical Analysis.

When the spectra were collected, it was visually examined through the Perkin Elmer Spectrum software (v 10.5.4) to note major qualitative differences between the counterfeit and authentic samples. Each spectrum consisted of roughly 3600 datapoints (one %T reading (y-axis) for every wavenumber (1 cm\(^{-1}\), x-axis)). After visual examination, it was determined that most of the inter-source variation in the spectra took place in the 400 cm\(^{-1}\) to 1400 cm\(^{-1}\) region. This portion of the spectrum was therefore included in the data analysis. Each spectrum was also transformed into its respective second derivative through RStudio (vers. 1.0). In order to reduce dimensionality of the data, principal component analysis was applied through JMP Pro 14. As there were 1000 original variables (one wavenumber in the 400 cm\(^{-1}\) to 1400 cm\(^{-1}\) region), the data was reduced to one
principal component, while capturing over 80% of the original variance. Each principal component for its respective ink color was combined per each sample prior to classification.

After the data was reduced, the machine learning algorithms were applied as described in 4.1.3. However, unlike LIBS, the ATR-IR beam does not penetrate into the ink layer to the substrate, generating an ink spectrum free of contribution from the carton. As a result, the paperboard substrate was analyzed separately from the ink and its elemental profile included as an additional variable in analysis.

5.2. Results and Discussion
5.2.1. PCA Analysis and Spectral Comparison

Examples for ATR-FTIR spectra of black ink are shown in Figure 31, where a full spectrum is shown for authentic reference A1 and magnified portions are shown for A2 – A5. There was typically little variation or peaks of interest between wavenumbers 2000 cm\(^{-1}\) – 2800 cm\(^{-1}\), as well as 3200 – 4000 cm\(^{-1}\). More variation occurred in the 2800 cm\(^{-1}\) – 3200 cm\(^{-1}\), or, most of all, the 400 cm\(^{-1}\)-1500 cm\(^{-1}\) (generally known as the ‘fingerprint region’ of ATR-FTIR spectra). Additionally, spectra comparing A1 and a replicate printout for C4 are shown in Figure 32. Again, a magnified portion is shown highlighting the 400 cm\(^{-1}\) to 1500 cm\(^{-1}\), where the most variation was observed between the authentic and counterfeit samples.
Figure 31. ATR-FTIR spectrum of black ink on A1 (bottom), as well as magnified portions of the fingerprint region of A1 - A5.
Figure 32. ATR-FTIR spectra of authentic template A1 (black), as well as a replicate printour from C4 (gray).
The instrument was able to analyze infrared signals from 400 cm\(^{-1}\) to 4000 cm\(^{-1}\) with a resolution of 1 cm\(^{-1}\), resulting in 3600 total data points or 3600 features to explain the data. Action was taken to reduce the feature space of the data in two ways: only using the 400 cm\(^{-1}\) to 1400 cm\(^{-1}\) region of the spectrum and reducing data dimensionality via PCA analysis. PCA was used to compress the features into one principal component per ink color, which possessed the majority of the variance in the original data (>80%). Principal component boxplots depicting maximum spread between the classes is shown in Figure 33.

![PCA plot of yellow ink for A1, depicting class separation with one feature PC1 component.](image)

5.2.2. 6 Class System

As in 4.2.4., six machine learning methods were used to classify the ATR-FTIR spectral data into six classes: as the authentic sample or a counterfeit sample from one of five printing sources. The results from this analysis is shown in Table 22. The average misclassification rates across all machine learning methods were below 15%. Like in 0., the misclassification rates within one dataset were relatively similar to one another (<10-15% difference) with few exceptions. There
was no obvious trend in method efficacy between parametric and nonparametric learning techniques. Interestingly, the datasets with overall higher misclassification rates using LIBS data (A3, A5, A6) also possessed overall higher misclassification rates using the ATR-FTIR data.

Table 22. Six-class system: misclassification rates of ATR-FTIR data using six machine learning methods.

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RF</strong></td>
<td>7.5%</td>
<td>3.0%</td>
<td>19.4%</td>
<td>3.3%</td>
<td>14.9%</td>
<td>23.7%</td>
<td>10.3%</td>
</tr>
<tr>
<td><strong>NB</strong></td>
<td>4.5%</td>
<td>4.0%</td>
<td>11.9%</td>
<td>1.1%</td>
<td>11.9%</td>
<td>9.3%</td>
<td>7.1%</td>
</tr>
<tr>
<td><strong>NN</strong></td>
<td>1.5%</td>
<td>0.0%</td>
<td>17.9%</td>
<td>3.0%</td>
<td>41.8%</td>
<td>5.1%</td>
<td>11.6%</td>
</tr>
<tr>
<td><strong>KNN</strong></td>
<td>1.5%</td>
<td>9.0%</td>
<td>11.9%</td>
<td>4.4%</td>
<td>3.0%</td>
<td>5.1%</td>
<td>5.8%</td>
</tr>
<tr>
<td><strong>QDA</strong></td>
<td>3.0%</td>
<td>0.0%</td>
<td>13.4%</td>
<td>1.5%</td>
<td>7.5%</td>
<td>6.8%</td>
<td>5.4%</td>
</tr>
<tr>
<td><strong>LDA</strong></td>
<td>9.0%</td>
<td>10.9%</td>
<td>19.4%</td>
<td>8.6%</td>
<td>19.4%</td>
<td>5.1%</td>
<td>12.1%</td>
</tr>
</tbody>
</table>
The corresponding QDA canonical plot for A1 is shown in Figure 34. As with the LIBS data, some classes show notable separation from the others (A and C2, primarily). Other classes show more overlap with one another (C1 and C3, C4 and C5). Interestingly, the data separated itself into three clusters: the authentic group cluster, the cluster of counterfeit sources originating from private printing companies (C1 – C3), and a cluster of counterfeit sources originating from home printers (C4 – C5). However, the second cluster especially exhibits high overlap with one another.

![Figure 34. QDA canonical plot, displaying class separation, of A1 using ATR-FTIR data.](image)

Like LIBS, ATR-FTIR had sufficiently low misclassification rates for the technique, especially when considering the technical advantages (e.g., speed, no sample preparation). ATR-FTIR provided organic information about the samples; through the use of the classification algorithms, the samples were relatively successfully matched with their source. In examining the clustering ability of the data, QDA plots showed that there was more clustering than observed in LIBS, though some sources still showing overlap.
5.3. Conclusion

ATR-FTIR was used to identify the presence of organic molecular functional groups within a molecule. Authentic and counterfeit pharmaceutical carton spectra were collected and visually examined through software. Most of the variation took place in the fingerprint region of the spectrum, as opposed to the region of the spectrum above $>2000 \text{ cm}^{-1}$. Some variation was observed even between authentic references of the similar ink color in the former region, indicating that different constituents can be present in inks on similar colors. For the purposes of analysis and to prevent overfitting to the system, only the region from $400 \text{ cm}^{-1}$ to $1400 \text{ cm}^{-1}$ was analyzed. It was predicted that this region alone would provide enough discriminating information to adequately classify the authentic and counterfeit samples.

PCA was performed to condense the data into the first principal components of their respective ink-colors and paper substrates. One principal component was taken for each ink color, explaining well over 80% of the variance in the original dataset. After the principal components were collected, they were combined for each authentic template (e.g. A1 had four analyzed ink colors and the paper substrate, resulting in five principal components). Once the principal components were collected, machine learning methods (Random Forest, Naïve Bayes, Neural Networks, $k$-Nearest Neighbors, Quadratic and Linear Discriminant Analysis) were used to classify the samples.

A six-class system was used to classify the samples: five counterfeit printing sources per each corresponding authentic package, for a total of six authentic product packages. The classification in general was excellent for a screening method with over 85% average successful classifications. Like in LIBS, one machine learning technique did not necessarily outperform the rest. Authentic templates with higher misclassifications in LIBS generally (A3 and A6 in particularly) also had higher misclassification rates in ATR-FTIR, suggesting an overall similar chemical profile within the ink formulations for those templates. Regardless, it was predicted that the complementary information from LIBS and ATR-FTIR would reduce the misclassification even further.
Examination of the canonical QDA plots yielded appreciable separation of some of the sample classes. A clustering effect was observed between the different printing source types. The authentic and counterfeit samples had clustered away from one another, but a separate cluster had formed within the counterfeit printing sources. The cluster of professional printing company sources (C1 – C3) and the cluster of home office printer sources (C4 – C5) were distinct from one another, yet there was overlap between the printing sources themselves.

6. Combination of Laser-Induced Breakdown Spectroscopy and Attenuated Total Reflectance – Fourier Transform Infrared Spectroscopy

6.1. Methodology

The performance in classification and identification of counterfeit packages, using LIBS and FTIR alone was compared to the performance when the data was combined. Due to the orthogonality of the methods, it was expected the LIBS and IR data would complement each other and improve the classification. The combined data would yield a more comprehensive chemical profile for each sample, further highlighting the differences between not only counterfeit and authentic samples but also within counterfeit printing sources themselves.

As discussed previously, three principal components were acquired from the LIBS data for each ink color and one principal component was acquired from the ATR-FTIR data for each ink color and paperboard substrate. In order to combine the data, the principal components from each technique were combined on one spreadsheet and analyzed through JMP Pro 14. The six machine learning techniques used previously (Random Forest, Naïve Bayes, Neural Network, k-Nearest Neighbors, Quadratic Discriminant Analysis, and Linear Discriminant Analysis) classified the data once again as either authentic or belonging to one of five counterfeit printing sources.

The misclassification rate was calculated for each authentic template. Misclassification rates were compared to those received when the data was used in isolation for each technique, as well as
analyzing the misclassification rates in general. Further, the canonical plots were also examined to determine whether more appreciable separation within the clusters occurred.

6.2. Results and Discussion

6.2.2. Combining Spectral Data Acquired from LIBS and ATR-FTIR

The LIBS and ATR-FTIR data were combined in order to determine their classification accuracy when joined. With the combined data, a more robust chemical profile was determined of the sample. For example, in Figure 35 displays an example of a LIBS peak and a magnified ATR-FTIR spectra for blue ink on the authentic sample as well as the counterfeit samples. While the LIBS peak shows a large gap in intensity between the authentic and counterfeit samples, there is more overlap observed between the different counterfeit printing sources. In examining the ATR-FTIR spectrum, however, there is less overlap between the different counterfeit printing sources (to a large degree, though some overlap between similar counterfeit printing sources such as C1 (blue) and C2 (orange) still occur). In this way, the combined data would allow for better separation of the classes than the data used alone.
Figure 35. Two emission lines (Al at 396 nm and Ca at 397 nm) and magnified ATR-FTIR Spectra of blue ink on A1 and 5 corresponding counterfeit counterparts (1 (authentic class: red, counterfeit printing source 1: blue, counterfeit printing source 2: green, counterfeit printing source 3: orange, counterfeit printing source 4: purple, counterfeit printing source 5: navy).
The same six-class system as described in previous sections was used. The misclassification rates for the combined LIBS and ATR-FTIR data are shown in Table 23. Misclassification rates of dataset using combined LIBS/ATR-FTIR data. The misclassification rates for the combined data were generally lower than either the isolated LIBS or ATR-FTIR data, with an average below 7.0%. Promisingly, the datasets that yielded relatively higher misclassification rates (A3, A5, A6) in the isolated LIBS and ATR-FTIR data indicated some of the largest decreases in misclassification rates between the data analyses. In examining the machine learning methods specifically, some generally performed better in this regard than others. The Neural Network algorithm had some of the lowest misclassification rates across all three models (LIBS, ATR-FTIR, and combined LIBS/ATR-FTIR data). This is likely due to the NN algorithm’s ability to account for interactions between feature variables, as well as having no firm distributional assumptions about the data. The machine learning methods that did make distributional assumptions (notably Naïve Bayes, Quadratic Discriminant Analysis, and Linear Discriminant Analysis) suffered poorer performance as a result of those assumptions not being meant. Accordingly, the other two machine learning methods that did not make distributional assumptions (Random Forest and k-Nearest Neighbors) had lower misclassification rates than the parametric classifiers.

Table 23. Misclassification rates of dataset using combined LIBS/ATR-FTIR data.

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>1.5%</td>
<td>0.0%</td>
<td>13.4%</td>
<td>0.0%</td>
<td>8.9%</td>
<td>5.6%</td>
<td>4.9%</td>
</tr>
<tr>
<td>NB</td>
<td>1.5%</td>
<td>3.0%</td>
<td>19.4%</td>
<td>2.7%</td>
<td>7.5%</td>
<td>7.4%</td>
<td>6.9%</td>
</tr>
<tr>
<td>NN</td>
<td>0.0%</td>
<td>0.0%</td>
<td>5.9%</td>
<td>0.0%</td>
<td>3.0%</td>
<td>5.6%</td>
<td>2.4%</td>
</tr>
<tr>
<td>KNN</td>
<td>10.4%</td>
<td>1.5%</td>
<td>10.4%</td>
<td>0.0%</td>
<td>3.0%</td>
<td>5.6%</td>
<td>5.2%</td>
</tr>
<tr>
<td>QDA</td>
<td>3.0%</td>
<td>3.1%</td>
<td>6.0%</td>
<td>0.0%</td>
<td>8.9%</td>
<td>9.4%</td>
<td>5.1%</td>
</tr>
<tr>
<td>LDA</td>
<td>1.5%</td>
<td>1.6%</td>
<td>7.4%</td>
<td>1.4%</td>
<td>14.9%</td>
<td>3.8%</td>
<td>5.1%</td>
</tr>
</tbody>
</table>

Accordingly, the LDA canonical plot for sample A1 is illustrated in Figure 36. Here, maximum class separation is seen, with little overlap occurring between any of the six classes. The three
clusters as seen with previous data still exist, but less overlap is occurring within printing sources in the same cluster.

Figure 36. LDA canonical plot, displaying class separation, of A1 using combined LIBS and ATR-FTIR data.

The misclassification rates generally decreased when the data from the two techniques were combined, indicating that complementary, discriminatory information was given by LIBS and ATR-FTIR. An average misclassification rate of <7.0% fits for purpose for a screening method, particularly when the method is highly suitable for on-site analysis. As well, the clustering observed in the LDA canonical plot indicates that enough identifying information was given by the methods to maintain the classes’ individuality.
6.3. Conclusion

Although the spectral data from LIBS and ATR-FTIR yielded acceptable misclassification rates when analyzed separately, the features were combined to determine whether the misclassification could be decreased even further. It was predicted that this would be the case due to the complementary nature of LIBS and ATR-FTIR data. LIBS data reveal mostly inorganic information about the sample of interest, while ATR-FTIR reveals mostly organic information. Using the data in conjunction would provide a more comprehensive chemical profile of the sample. In this way, differences in the composition between the authentic and counterfeit samples (as well as within counterfeit samples of different printing sources) would be highlighted during classification.

Misclassification rates from the combined data were promising. The average misclassification rates fell below 7.0%. Datasets that yielded higher misclassification rates when examining one technique in isolation were successfully classified when the data was combined. Additionally, when examining the LDA canonical plots, the same clusters occurred as seen in the ATR-FTIR data: one authentic sample cluster, one cluster of professional printing companies, and one cluster of office printers. However, the ATR-FTIR showed appreciable overlap within the counterfeit clusters. Less overlap was shown using the combined data. This indicates that more discriminating information between the counterfeit printing sources is revealed when the LIBS and ATR-FTIR data is used in combination, allowing the classification algorithms to find more differences between similar counterfeit printing sources. This would be useful in an investigation when authorities attempt to determine a difference between distinct but similar counterfeit printing source origins.

In general, these results show that the combination of organic and inorganic spectral data provide sufficient and extensive features to classify pharmaceutical packaging samples as authentic or counterfeit, as well as indicating the potential origin of a counterfeit source. When used in combination, the misclassification rates were lower than the isolated data from LIBS and ATR-FTIR. While the misclassification rates were sufficiently low in that case (LIBS generally
providing misclassification rates below 30% and ATR-FTIR below 15%), it was further decreased to less than 7% misclassification between authentic and counterfeit samples, as well as within counterfeit printing sources.

7. Conclusion Summary

One of the rising dangers in the past thirty years is the counterfeit manufacture of pharmaceutical products. The prevalence and distribution of these fake drugs differs wildly by geographic location, but areas with low pharmaceutical supply chain security are especially at risk. The creation of the Internet e-market has also provided another avenue for counterfeiters to sell their products. Several characteristics increase a counterfeiter’s likelihood to imitate drugs, including high demand, high cost, or low availability. Antimalarials and vasodilators are among the most commonly counterfeit pharmaceuticals.

There are three ways for a pharmaceutical product to be counterfeit: if the composition of the pharmaceutical itself is counterfeit (increased/decreased/absent API, presence of adulterants), if the packaging enclosing the pharmaceutical is counterfeit, or both. Although extensive studies have been conducted on the former, few have been conducted on the packaging of the pharmaceutical. This study has indicated that spectroscopic tools are effective in discriminating between authentic and counterfeit pharmaceutical packaging. In particular, Laser Induced Breakdown Spectroscopy (LIBS) and Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy (ATR-FTIR) are suitable for the discrimination of different types of printing inks and carton substrate. Additionally, several characteristics of these two techniques render them especially suitable for the identification of counterfeit packaging. Both LIBS and ATR-FTIR are fast, sensitive, portable, and have relatively negligible sample destruction.

For that reason, LIBS and ATR-FTIR were chosen to discriminate between authentic and counterfeit pharmaceutical packaging. Authentic pharmaceutical cartons (n = 54) were collected from known sources. Counterfeit pharmaceutical cartons (n = 112) were recreated from five printing sources and were based on six authentic pharmaceutical carton blueprints. For the
authentic samples, intra-lot and inter-lot variability was assessed. Likewise, intra-sample variability was examined on counterfeits using replicate printouts from each printing source and blueprint. False exclusion rates were found to be low for counterfeit and authentic samples (<20%).

When classifying the samples, six machine learning methods were used: Random Forest, Naïve Bayes, Neural Network, k-Nearest Neighbors, Quadratic Discriminant Analysis, and Linear Discriminant Analysis. All of these machine learning methods were shown to have some utility in classification models in prior research studies. A mixture of parametric and non-parametric algorithms was chosen. For all classification problems, a sixty/fourty train/test split was conducted.

Analyzing LIBS data, a two-class system was evaluated that classified black ink and paperboard as authentic or counterfeit, without focusing on specific counterfeit source. Classification rates were above 92% for combined black ink and paperboard substrate data, indicating LIBS utility as a quick screening tool between authentic and counterfeit samples.

Afterwards, a six-class system was analyzed that classified seven colors of ink (red, blue, yellow, green, pink, brown, black) and paperboard substrate into one of six classes: the authentic class and one of five questioned printing sources. PCA was used to compress the data into three explanatory principal components (explaining over 80% of total variance); the principal components from all ink colors were combined for analysis. LIBS was generally able to correctly classify more than 70% of samples, with most misclassifications derived from confusion between counterfeits with similar ink types rather than confusion between authentic and counterfeits.

ATR-FTIR data was also analyzed using the six-class system. As the analyzed portion spectra was composed of 1000 data points, principal component analysis was utilized to compress the data into 1 feature for each analyzed ink type and paperboard substrate while retaining over 80% of its original variance. ATR-FTIR was able to correctly classify 85% of the samples across all seven ink colors and substrate. On LDA canonical plots, the classes organized themselves into three clusters: the authentic class, the classes composed of the professional printing sources, and the
classes composed of home printers, demonstrating is utility for providing investigative leads in anti-counterfeit operations.

Finally, the data from ATR-FTIR and LIBS was combined to see if misclassification rates could be reduced further. Across all seven colors and paperboard substrate, misclassification rates were less than 7.0%, providing over 93% correct classification.

LIBS and ATR-FTIR had low misclassification rates for counterfeit and authentic pharmaceutical packaging samples when used separately, but the informative value and classification performance was significantly improved when the data was used in combination. This indicates that both techniques show promising capability to be used as a screening tool for the fast, sensitive, non-invasive identification of counterfeit pharmaceutical packaging. Larger datasets would be needed in the future to assess the generalizability of the proposed approach.
8. References


Appendix 1. Anti-counterfeit technologies visible under UV lighting (top left, moving right: A005, A006, A007. Bottom left: A008, A009, A010.)
Appendix 11. LIBS and ATR-FTIR spectra of pink ink on counterfeit counterpart 2 (1 (authentic class: red, counterfeit printing source 1: blue, counterfeit printing source 2: green, counterfeit printing source 3: orange, counterfeit printing source 4: purple, counterfeit printing source 5: navy)
Appendix 24. LIBS and ATR-FTIR spectra of blue ink on counterfeit counterpart 6 (1 (authentic class: red, counterfeit printing source 1: blue, counterfeit printing source 2: green, counterfeit printing source 3: orange, counterfeit printing source 4: purple, counterfeit printing source 5: navy)