

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DOS MATERIAIS

GUILHERME CIOCCIA NEVES

**ELABORAÇÃO DE PROTOCOLO PARA IDENTIFICAÇÃO DE
RESÍDUOS DE DISPARO POR ARMA DE FOGO UTILIZANDO
ESPECTROSCOPIA LIBS E APRENDIZADO DE MÁQUINA**

CAMPO GRANDE, MS

2023

GUILHERME CIOCCIA NEVES

**ELABORAÇÃO DE PROTOCOLO PARA IDENTIFICAÇÃO DE
RESÍDUOS DE DISPARO POR ARMA DE FOGO UTILIZANDO
ESPECTROSCOPIA LIBS E APRENDIZADO DE MÁQUINA**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência dos Materiais da Universidade Federal de Mato Grosso do Sul, como requisito à obtenção do título de Mestre em Ciência dos Materiais.

Orientador: Prof. Dr. Bruno Spolon Marangoni

Coorientador: Prof. Dr. Cícero Rafael Cena da Silva

CAMPO GRANDE, MS

2023

AGRADECIMENTOS

A minha mãe, que mesmo morando distante envia mensagens de carinho e atenção todos os dias.

A minha companheira Lethycia Anjos, pelo apoio incondicional nas batalhas do dia-a-dia.

Ao meu orientador Bruno Marangoni, que sempre foi flexível para me ouvir e auxiliar a desenvolver novos conceitos.

E ao meu coorientador Cícero Cena, por clarear e organizar minhas ideias.

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES), por meio do Programa de Cooperação Acadêmica em Segurança Pública e Ciências Forenses (PROCAD - SPCF) – Número do Processo 88887.612535/2021-00.

RESUMO

A violência por arma de fogo causa milhares de vítimas anualmente no Brasil. Isso ocorre, dentre outros fatores, devido à baixa taxa de resolução de homicídios em diversas regiões do país. A eficiência é fundamental para investigação, pois quase a totalidade dos crimes que são solucionados tem sua investigação finalizada no primeiro ano após o ocorrido. Um dos métodos para gerar evidências periciais sobre os suspeitos de realizarem disparos por arma de fogo é a identificação de resíduos de disparo em suas mãos na forma de partículas inorgânicas, geralmente compostas por chumbo (Pb) e antimônio (Sb). O método padrão ouro para identificação de GSR (Gunshot Residues) é o MEV-EDS que, apesar de ser extremamente preciso, demanda muito tempo de trabalho para identificar cada amostra. Este trabalho propõe um novo protocolo de coleta e análise de GSR como alternativa ao MEV-EDS, baseando-se em técnicas que associam espectroscopia LIBS com algoritmos de aprendizado de máquina. Com a metodologia proposta, o modelo treinado em com base no algoritmo SVM foi capaz de prever com precisão amostras positivas e negativas, estabelecendo probabilidades para cada previsão. Com base nas probabilidades, um limite de previsão foi definido para descartar previsões pouco confiáveis. O modelo final foi capaz de prever com precisão as amostras verdadeiras de cada classe e eliminar as amostras contaminadas e aberrantes (fora do padrão esperado), o que demonstrou o potencial do novo protocolo para analisar inúmeras amostras simultaneamente e em condições de campo, onde não é possível controlar elementos contaminantes.

Palavras-Chave: GSR, arma de fogo, espectroscopia, LIBS, aprendizado de máquina.

ABSTRACT

Gun violence causes thousands of victims annually in Brazil. This occurs, among other factors, due to the low homicide resolution rate in several regions of the country. Efficiency is essential for investigation, as almost all crimes that are solved have their investigation completed within the first year after the incident. One of the methods for generating evidence on those suspected of shooting a firearm is the identification of gunshot residue on their hands in the form of inorganic particles, generally composed of lead (Pb) and antimony (Sb). The standard method for identifying GSR (Gunshot Residues) is SEM-EDS which, despite being extremely accurate, requires a lot of work time to identify each sample. This work proposes a new GSR collection and analysis protocol as an alternative to SEM-EDS, based on techniques that combine LIBS spectroscopy with machine learning algorithms. With the proposed methodology, the model trained based on the SVM algorithm was able to predict with positive and negative samples, establishing probabilities for each prediction. Based on the probabilities, a prediction threshold was set to rule out somewhat reliable outliers. The final model was able to accurately predict the true samples of each class and eliminate contaminated samples and outliers (outside the expected standard), which demonstrated the potential of the new protocol to analyze numerous samples simultaneously and in field conditions, where no it is possible to control contaminating elements.

Key-Words: GSR, firearm, spectroscopy, LIBS, machine learning.

LISTA DE FIGURAS

Figura 1 - a) revólver calibre .38 SPL Amadeo Rossi S.A. utilizado na etapa de disparos e b) perfil e inscrições da base da munição .38 SPL.	19
Figura 2 - a) esquema informativo sobre as regiões das mãos que a amostra foi coletada e b) foto do processo de coleta de acordo com esquema informativo.	20
Figura 3 - Modelo esquemático do processo de captura de dados realizado com sistema LIBS. No total, 15 pulsos igualmente espaçados incidiram sobre a fita.	21
Figura 4 - Média por classe dos espectros brutos. O desvio padrão da média é indicado pelo sombreado abaixo da média espectral.....	24
Figura 5 - Espectros da classe “Não Atirador” considerados <i>outliers</i> pois possuem transições relacionadas à GSR.	25
Figura 6 - Média por classe dos espectros processados. O desvio padrão da média é indicado pelo sombreado abaixo da média espectral.....	26
Figura 7 - Média espectral e desvio padrão da fita adesiva dupla-face utilizada para a coleta das amostras.	27
Figura 8 - Média espectral de cada classe sobreposta em linhas com as indexações de transições.....	28
Figura 9 - Espectro coletado sobre o projétil de mesmo tipo utilizado na coleta das amostras.....	29
Figura 10 - Gráfico de dispersão das três primeiras PCs, seccionado em duas dimensões para melhor visualização	29
Figura 11 - Loadings das cinco primeiras componentes principais para o intervalo espectral completo.	30
Figura 12 - Gráfico de dispersão das três primeiras PCs para o Região 1, seccionado em duas dimensões para melhor visualização.....	31
Figura 13 - Gráfico de dispersão das três primeiras PCs para o Região 2, seccionado em duas dimensões para melhor visualização.....	32
Figura 14 - Matriz confusão para os melhores resultados obtidos em validação externa para cada região espectral. a) Região completo; b) Região 1; c) Região 233	
Figura 15 - Amostras “Não Atirador” selecionadas para validação externa. Nota-se grande quantidade de contaminação por Pb e Sb nas amostras “0” e “2”, já nas amostras “1” e “3” podem ser identificadas pouca ou nenhuma contaminação.	35

Figura 16 - Amostras com pouca ou sem GSR rotuladas como “Atirador”, portanto, Outliers, presentes no conjunto de validação externa.....36

LISTA DE TABELAS

Tabela I - Modelos treinados e otimizados sob a métrica "Acurácia" para cada intervalo espectral. Resultados obtidos em validação externa.	33
Tabela II - Probabilidades de classificação para a classe real "Não Atirador" no intervalo espectral completo.....	37
Tabela III - Probabilidades de classificação para a classe real Atirador no intervalo espectral completo.	37
Tabela IV - Probabilidades de classificação para a classe real "Não Atirador" no intervalo espectral "Região 1".....	38
Tabela V - Probabilidades de classificação para a classe real "Atirador" no intervalo espectral "Região 1".	38
Tabela VI - Probabilidades de classificação para a classe real Não Atirador no intervalo espectral "Região 2".....	39
Tabela VII - Probabilidades de classificação para a classe real Atirador no intervalo espectral "Região 2".	39

SUMÁRIO

RESUMO.....	1
ABSTRACT.....	2
LISTA DE FIGURAS.....	3
LISTA DE TABELAS.....	5
1. INTRODUÇÃO.....	7
2. OBJETIVOS.....	9
2.1 Objetivo geral.....	9
2.2 Objetivos específicos.....	9
3. REVISÃO DA LITERATURA.....	10
3.1. Resíduo de disparo por arma de fogo.....	10
3.2. Métodos para identificação de GSR.....	10
3.3. Espectroscopia óptica para identificação de GSR.....	11
3.4. Análise multivariada e aprendizado de máquina.....	12
3.4.1. Pré-processamento espectral.....	12
3.4.2. <i>Standard Normal Variate</i> (SNV).....	13
3.4.3. <i>Rolling Balls</i> (RB).....	13
3.5. Análise de componentes principais.....	14
3.6. Support Vector Machines (SVM).....	14
3.7. Validação cruzada.....	16
3.8. Previsão probabilística.....	17
4. MATERIAIS E MÉTODOS.....	18
4.1. Coleta das amostras.....	18
4.2. Pré-processamento.....	21
4.3. Análise não-supervisionada e seleção de transições.....	22
4.4. Análise supervisionada.....	22
4.5. <i>Softwares utilizados</i>	23
5. RESULTADOS E DISCUSSÃO.....	24
5.1. Análise exploratória.....	24
5.2. Análise não-supervisionada (PCA).....	29
5.3. Desempenho dos modelos de aprendizado de máquina.....	33
5.4. Análise detalhada de probabilidades.....	34
6. CONCLUSÃO.....	40
7. REFERÊNCIAS BIBLIOGRÁFICAS.....	43

1. INTRODUÇÃO

A violência é um problema grave e persistente no Brasil, apenas na última década 560 mil pessoas perderam a vida em decorrência de lesão provocada de forma violenta (homicídio, suicídio, acidente ou morte suspeita). Desse total, cerca de 80% das mortes foram causadas por armas de fogo [1, 2].

Além das ocorrências fatais, é importante salientar que ferimentos e lesões provocados por armas de fogo podem causar sequelas irreversíveis à vítima, que impossibilitam o trabalho e geram internações que pressionam o Sistema Único de Saúde (SUS) [3].

Em 2019, último ano disponível da série histórica do Atlas da Violência, destaca-se que os estados brasileiros que mais registraram mortes por arma de fogo foram Bahia, Pernambuco, Pará, Rio de Janeiro e Ceará. Juntos, eles concentraram quase metade do total de homicídios. Proporcionalmente, a cada 100 mil habitantes, os estados que mais registraram mortes por arma de fogo foram Sergipe, Rio Grande do Norte, Bahia, Amapá e Pernambuco [1].

Estes mesmos estados estão presentes nas últimas posições do ranking da impunidade do Instituto Sou da Paz, que é elaborado com base na taxa de homicídios solucionados em relação aos homicídios cometidos no ano. Similarmente, observa-se que estados com taxas de homicídio abaixo da média nacional, como Roraima, Mato Grosso do Sul e Santa Catarina, possuem maiores taxas de solução de homicídios [2].

Dessa forma, é possível observar uma correlação entre a taxa de solução de homicídios e a redução desses crimes. O relatório “Aprendizados sobre o esclarecimento de homicídios no Brasil: Práticas promissoras e caminhos a seguir” [4], do Instituto Sou da Paz, elenca os possíveis motivos para esse fenômeno e a importância do esclarecimento desses crimes. Nesse documento, diversos especialistas argumentam que a prisão de homicidas:

1. Previne novos homicídios causados por reincidência;
2. Desencoraja outros crimes na região devido a responsabilização do criminoso;
3. Pode interromper um ciclo de violência causado por vingança;
4. Fornece segurança para que novas testemunhas colaborem.

É importante notar que a maior parte das soluções de homicídio ocorrem no primeiro ano após o fato, e dificilmente os crimes são solucionados em períodos seguintes. Portanto, o fornecimento de provas científicas é essencial para a investigação do crime de forma rápida e confiável, agilizando uma fase importante dos processos penais.

Convencionalmente, a *Gun-shot Residue* (GSR, ou “Resíduo por Arma de Fogo”) é identificada por microscopia eletrônica de varredura (MEV-EDS), considerada “padrão ouro”. Apesar deste método ser muito preciso, ele é oneroso e demanda um elevado período de tempo de trabalho. Em contrapartida, é possível encontrar na literatura estudos de caso que demonstram a eficácia de técnicas de espectroscopia óptica na identificação dos elementos que compõem a GSR [5, 7, 8, 9, 13, 14].

Quando associados à análise de espectroscopia óptica, algoritmos de aprendizado de máquina podem ser utilizados para identificar as transições eletrônicas mais importantes para identificação de GSR, além de automatizar o processo de classificação das amostras fornecendo alta confiabilidade e eficiência, conforme já demonstrado na literatura [7, 13, 14].

Portanto, este trabalho propõe a criação de um novo protocolo para identificação de resíduos de disparo por arma de fogo (GSR) nas mãos de suspeitos. De forma simplificada, esse método consiste na coleta de amostras utilizando fita adesiva dupla-face e analisá-las por meio de *Laser-Induced Breakdown Spectroscopy* (LIBS). Em seguida, é empregado um modelo de aprendizado de máquina para prever a classe de espectros desconhecidos e calcular a probabilidade de pertencerem a um atirador ou não atirador. Caso a probabilidade de previsão seja considerada muito baixa, recomenda-se a aplicação de outro método de análise ou o descarte da amostra.

Essa abordagem integrada combina as vantagens da espectroscopia óptica com a potência dos algoritmos de aprendizado de máquina, proporcionando uma metodologia eficiente, confiável e escalável para a identificação de GSR.

2. OBJETIVOS

2.1 Objetivo geral

Elaboração de novo protocolo de análise para identificação confiável e escalável de resíduo por arma de fogo nas mãos de indivíduos, utilizando espectroscopia LIBS associada a análise exploratória de dados e modelos de aprendizado de máquina.

2.2 Objetivos específicos

- Identificar e selecionar transições características de resíduos por arma de fogo em espectros coletados das mãos de indivíduos;
- Avaliar metodologia de coleta das amostras a partir da capacidade de identificação das transições em análise exploratória e do desempenho dos modelos de aprendizado de máquina;
- Analisar e comparar o desempenho dos modelos de aprendizado de máquina treinados na região espectral completa e em regiões de transições selecionadas;
- Implementar a análise de probabilidades de previsão com base no algoritmo SVM;
- Treinar, otimizar e validar pelo modelos de aprendizado de máquina que sejam acurados, precisos e confiáveis no contexto da análise de probabilidades;
- Elaborar definição de “classe indeterminada” com base na análise das probabilidades de classificação de cada amostra em validação externa.

3. REVISÃO DA LITERATURA

3.1. Resíduo de disparo por arma de fogo

O resíduo de disparo por arma de fogo, ou *Gunshot Residue* (GSR), é composto por partículas produzidas no disparo de uma arma, decorrentes da combustão da mistura iniciadora e da carga propelente. A GSR é formada principalmente por pó propelente queimado ou parcialmente queimado, fumaça, metais da arma e do projétil, graxa e lubrificantes [6].

Dentre os resíduos, é possível identificar compostos orgânicos e inorgânicos, sendo que o primeiro tem origem principalmente de lubrificantes utilizados na manutenção do equipamento e da mistura propelente, a depender do processo de fabricação e da marca comercial do tipo de munição, enquanto o segundo se origina predominantemente da mistura iniciadora e do projétil [6, 7].

Os elementos inorgânicos que são comumente relacionados à GSR são Chumbo (Pb), Antimônio (Sb) e Bário (Ba), provenientes da mistura iniciadora presente na espoleta (também chamado de *primer*). Estes elementos são vaporizados pelo ambiente de alta temperatura e pressão do disparo, sendo expelidos rapidamente pelo ar, acabando por se depositarem nas mãos, cabelos e roupas do atirador na forma de GSR [6, 8].

As partículas inorgânicas e orgânicas de GSR podem ser coletadas das superfícies em que se depositam utilizando métodos de *stub* (tipicamente fita adesiva) ou *swab* (haste flexível com algodão na extremidade). Atualmente, técnicas de coleta com aplicação sucessiva de fita adesiva no local são amplamente utilizadas pela eficiência e simplicidade, podendo ser aplicadas em vários tipos de materiais, como pele, roupas e cabelo [6, 9].

3.2. Métodos para identificação de GSR

Descrito inicialmente por Teodoro Gonzalez, em 1933, o “teste de parafina” tem aplicação baseada na coleta da amostra na mão do suspeito utilizando cera quente de parafina que, após resfriada e solidificada, é retirada da mão carregando os resíduos do disparo. O molde então é submetido a uma pulverização de difenilamina, que assume cor azulada caso existam partículas residuais de disparo [10].

O teste de parafina apresenta bom nível de informação sobre a distribuição de partículas, mas apresenta um alto nível de falsos positivos. Como ele apenas indica

nitritos presentes nas amostras, partículas relacionadas com o uso de fertilizantes, fármacos e urina podem ser indicadas pelo teste [10].

Posteriormente, outros testes colorimétricos foram desenvolvidos em escala, como o teste de Griess, o teste de Harrison e Gilroy e o teste de rodizonato de sódio. O primeiro, descrito por Peter Griess, era utilizado principalmente para identificar íons de nitrato provenientes de GSR e tinha seletividade levemente maior ao teste de parafina [10,11].

Já o teste de Harrison e Gilroy tem foco principal na detecção de partículas metálicas que constituem a GSR e utiliza *swab* encharcado em solução de dil. HCl para coletar a amostra das mãos do suspeito. Após a secagem do *swab*, a amostra é tratada com tri-phenylarsonium iodide e, por fim, o reagente rodizonato de sódio é adicionado. Dependendo da cor identificada, pode indicar a presença de antimônio, chumbo e bário. O principal problema desta técnica é que o resultado não é quantificável, além disso, a presença de outros materiais podem causar interferências na reação [10, 11].

Atualmente, o método considerado “padrão ouro” para identificação de GSR é a inspeção por Microscopia de Varredura Eletrônica (MEV). Por meio deste equipamento, é possível identificar a topologia característica das partículas de GSR, que combinado com uma análise espectral de dispersão de raios-X (EDS) indica a sua composição metálica [11].

Apesar da alta precisão, a demora para entregar o resultado é uma grande desvantagem do método, pois a busca pela partícula de GSR é desempenhada pelo perito técnico de forma visual e cautelosa, devendo analisar uma grande área amostral buscando separar as partículas úteis para a identificação das descartáveis [10].

3.3. Espectroscopia óptica para identificação de GSR

Diversos trabalhos demonstram o potencial de técnicas espectroscópicas associadas à análise multivariada de dados na identificação ou classificação de inúmeros tipos de amostra. Dentre as técnicas espectroscópicas disponíveis, destaca-se a *Laser-Induced Breakdown Spectroscopy* (LIBS) pela eficiência, boa acurácia de detecção, pouca necessidade de preparação e mínima abrasão da amostra [17].

O método LIBS é baseado na interação entre um *laser* potente de curta duração e a superfície da amostra. Este processo induz a formação de um plasma, que ao ser

resfriado emite radiação eletromagnética em múltiplos comprimentos de onda detectáveis por um espectrômetro acoplado ao sistema. Dessa forma, é possível identificar as transições características dos átomos que compõem a amostra por meio da análise das linhas espectrais.

Nota-se o impacto de várias aplicações da técnica LIBS e análise multivariada na literatura, por exemplo: na medicina, com o diagnóstico de células cancerígenas [15, 16]; na agricultura, com a detecção de elementos específicos em água, fertilizantes e sementes [18, 19, 20]; e nas ciências forenses, com a identificação de resíduos de disparo por arma de fogo [7, 8, 21]. Nesta última, a técnica LIBS é utilizada para identificar principalmente transições dos elementos chumbo (Pb), bário (Ba) e antimônio (Sb) [8].

3.4. Análise multivariada e aprendizado de máquina

Métodos espectroscópicos tendem a gerar uma grande quantidade de dados, com matrizes que podem alcançar milhares de pontos amostrais. Dessa forma, a análise eficiente e precisa deve ser realizada com o advento de métodos computacionais, aplicando técnicas específicas para melhorar a performance geral do processo. Neste trabalho, o ciclo de análise dos dados constituiu-se do pré-processamento espectral, do aprendizado de máquina não-supervisionado e, por fim, do aprendizado de máquina supervisionado.

3.4.1. Pré-processamento espectral

A espectroscopia LIBS é particularmente suscetível a estímulos externos à emissão do plasma, que causam efeitos aleatórios e não negligenciáveis nos resultados da análise. Devido a isso, experimentos com LIBS geralmente possuem pouca reprodutibilidade caso métodos de processamento espectrais não sejam aplicados [22, 23].

Geralmente, estes efeitos são caracterizados por ruídos aleatórios e variações sistemáticas no conjunto espectral, como a deriva e o deslocamento da linha de base devido a efeitos do espalhamento da luz e de variações instrumentais. Além disso, pode ocorrer diferenças globais na intensidade do sinal dos espectros [22].

Portanto, o pré-tratamento é essencial para a análise subsequente, pois intensifica o sinal sobre o ruído aleatório e auxilia na separação de picos sobrepostos

[22]. Além disso, alguns algoritmos de aprendizado são diretamente dependentes do método de pré-processamento aplicado, incrementando ou reduzindo a performance dos modelos [22, 23]. Neste trabalho, as técnicas de pré-tratamento utilizadas foram: média espectral, Standard Normal Variate (SNV), Rolling Balls (RB).

3.4.2. Standard Normal Variate (SNV)

A normalização SNV é um método aplicado para atenuar efeitos multiplicativos e diferenças globais na intensidade do sinal, tendendo a reduzir a variação entre os pontos amostrais [22, 24]. Esta transformação é aplicada em cada espectro de forma independente por meio da sua média espectral e da normalização pelo seu desvio padrão:

$$x_{i,j}^{SNV} = \frac{(x_{ij} - \bar{x}_i)}{\sqrt{\frac{\sum_{y=1}^p (x_{ij} - \bar{x}_i)^2}{p-1}}} \quad (1)$$

Onde $x_{i,j}^{SNV}$ é o ponto espectral transformado, $x_{i,j}$ é o ponto espectral original do espectro i na variável j , \bar{x}_i é a média do espectro i , e p é o número de variáveis ou comprimentos de onda no espectro.

A capacidade de corrigir o deslocamento da curva de base e atenuar a variação da intensidade global do sinal são pontos positivos, além disso, há incremento de performance em modelos de aprendizado de máquina baseados em distância quando os espectros são transformados pelo método [22].

3.4.3. Rolling Balls (RB)

Para qualquer técnica espectroscópica, é ideal que as curvas de base espectrais sejam niveladas a zero (sem sinal), mas devido a influências externas e sistemáticas é comum observar um deslocamento em certas regiões dos espectros [22, 25]. Diversos algoritmos foram desenvolvidos para atenuar este problema, onde destaca-se a Rolling Balls (RB) pela simplicidade e eficácia. A correção por este método se baseia na delimitação da trajetória de uma esfera com raio fixo rolando abaixo de cada espectro que é utilizada para subtrair a curva de base [23, 25].

Proposta originalmente para transformar espectros de raios-X, a técnica foi simplificada e pode ser utilizada amplamente por outros métodos espectroscópicos. Além da correção de curva base, a RB possui um efeito suavizador nos espectros e

sua otimização de parâmetros geralmente é realizada por meio de inspeção visual [25].

3.5. Análise de componentes principais

A Análise de Componentes Principais, mais conhecida pela sigla *PCA* (*Principal Components Analysis*), é uma técnica de redução de dimensionalidade tipicamente utilizada para visualizar padrões em agrupamentos amostrais, oferecendo um indicativo se duas ou mais amostras são separáveis.

Dessa forma, o conjunto de dados original, geralmente composto de milhares de amostras e variáveis, é submetido a uma transformação linear que reduz a sua dimensão em um novo conjunto de coordenadas, chamadas Componentes Principais (PC). A variância do sistema tende a se concentrar nas PCs de menor valor, tipicamente PC1, PC2 e PC3, e a distribuição das amostras pode ser visualizada em duas ou três dimensões em qualquer dupla ou trio de PCs, pois cada dimensão é linearmente independente.

Os Componentes Principais são os autovetores da matriz de covariância dos dados, portanto, geralmente são calculados pelo método de decomposição em valores singulares (*Single Value Decomposition* - SVD) da matriz de covariância. Com os PCs calculados, é possível determinar os autovalores, que se referem à escala da matriz de covariância, enquanto os autovetores se referem à direção da matriz de covariância. É possível acoplar as duas informações (escala e direção) a partir da relação entre os autovetores e autovalores, chamada de *loadings*, que são as correlações entre as variáveis originais e os componentes principais em escala unitária.

3.6. Support Vector Machines (SVM)

O Support Vector Machines (SVM) é um algoritmo de aprendizado de máquina utilizado em tarefas supervisionadas de classificação e regressão. De forma resumida, após a separação do conjunto amostral original nos subconjuntos de “treinamento” e “validação”, o modelo é treinado sobre a relevância de cada ponto de dados do conjunto de treinamento a fim de criar a fronteira de decisão entre as classes em análise.

Para fazer previsões de pontos amostrais desconhecidos, o SVM calcula a distância até cada um dos vetores de suporte, que são os pontos do conjunto de treinamento localizados na fronteira entre as classes. Com base nas distâncias até esses vetores de suporte e na importância que foram aprendidas durante o treinamento, o SVM estabelece a fronteira de decisão para classificação.

Essa classificação ocorre através de uma função que mapeia a posição das amostras de modo a maximizar a distância entre as amostras de diferentes classes. O SVM busca identificar um hiperplano de $(n-1)$ dimensões, em relação à dimensão n das amostras originais, que melhor separa as classes distintas. Essa abordagem torna o SVM adequado para lidar com problemas complexos de classificação, especialmente em espaços de alta dimensão, onde sua capacidade de encontrar o hiperplano ótimo entre as classes se destaca.

Além disso, outra vantagem do SVM é a capacidade de lidar com dados não-lineares por meio do uso de “kernels”. Os kernels permitem mapear as amostras para um espaço de maior dimensão onde a separação das classes pode ser realizada de maneira mais eficiente. Isso possibilita que o SVM trabalhe em cenários em que os dados não podem ser separados linearmente no espaço original, expandindo sua aplicabilidade para uma variedade maior de problemas complexos de classificação e regressão [26].

Por fim, outro ponto de destaque no algoritmo SVM são os hiperparâmetros, que são parâmetros internos ajustáveis que permitem controlar o comportamento e a complexidade do modelo. Dessa forma, um dos hiperparâmetros mais importantes é o parâmetro C , que controla a regularização do modelo. Ao ajustar o valor de C , podemos modificar o equilíbrio entre a suavidade da fronteira de decisão e a precisão da classificação [26].

Quando o valor de C é pequeno, o modelo se torna mais restrito, resultando em uma fronteira de decisão quase linear, na qual cada ponto de dados tem uma influência limitada. Conforme o valor de C aumenta, a influência dos pontos de dados mal classificados no modelo também aumenta, o que faz com que a fronteira de decisão se curve para corretamente classificar esses pontos. Portanto, ajustar o hiperparâmetro C permite encontrar o equilíbrio adequado entre a flexibilidade do modelo e sua capacidade de generalização para obter melhores resultados de classificação [26].

3.7. Validação cruzada

A criação de um modelo de aprendizado de máquina depende de um conjunto de dados para treinamento e outro para teste ou validação. Dessa forma, é possível avaliar a performance do modelo utilizando métricas como acurácia e precisão, ou ainda visualizar os erros e acertos do modelo por meio de uma matriz confusão.

Nesse contexto, o presente trabalho utilizou o processo de validação cruzada conhecida como *Leave-One-Out Cross Validation* (LOO-CV), uma técnica de validação cruzada iterativa. Nesse método, o algoritmo de aprendizado de máquina é treinado diversas vezes, sendo que em cada iteração, uma única amostra é excluída do conjunto de treinamento para servir como conjunto de teste. Esse processo é repetido para todas as amostras disponíveis, garantindo que cada exemplo tenha a oportunidade de ser utilizado como dado de teste.

A utilização do LOO-CV é particularmente benéfica em conjuntos de dados com tamanho reduzido, pois permite uma avaliação mais precisa da performance do modelo. Em comparação a outros métodos de validação cruzada, o LOO-CV tende a oferecer uma menor taxa de erro de classificação, uma vez que cada amostra é testada exatamente uma vez.

Entretanto, é importante ressaltar que, à medida que o tamanho do conjunto de treinamento aumenta, o tempo de processamento também se eleva significativamente. Isso ocorre porque o modelo precisa ser treinado repetidamente para cada uma das amostras, o que pode se tornar computacionalmente custoso em conjuntos de dados grandes.

Em suma, a criação de um modelo de aprendizado de máquina é um processo que envolve etapas cruciais de treinamento e validação. O uso adequado de técnicas como o LOO-CV proporciona uma avaliação mais precisa da performance do modelo, permitindo a seleção de abordagens mais eficazes para problemas específicos. No entanto, é fundamental considerar o tamanho do conjunto de dados e o tempo de processamento requerido para escolher a estratégia mais adequada em cada caso.

Neste trabalho, o processo de treinamento foi realizado por meio de validação cruzada *Leave-One-Out Cross Validation* (LOO-CV), no qual o algoritmo é treinado iterativamente sobre todo o conjunto de dados pelo menos uma vez, com exceção de uma amostra que é separada do processo para teste. Em conjuntos amostrais

menores, o LOO-CV oferece menor taxa de erro de classificação em comparação a outros métodos de validação cruzada, mas o tempo de processamento aumenta conforme a quantidade de amostras de treinamento.

3.8. Previsão probabilística

Os modelos de aprendizado de máquina para tarefas de classificação geralmente fornecem apenas uma matriz de valores inteiros como saída, isso significa que eles só podem prever se uma amostra é negativa ou positiva. Por exemplo, um modelo de classificação de câncer de mama pode prever se um tumor é benigno ou maligno, mas não pode fornecer uma estimativa da probabilidade de um tumor ser maligno.

Por fornecer informações limitadas, a matriz de valores inteiros pode não ser a mais indicada para a tomada de decisão em aplicações práticas, principalmente se a qualidade dos dados de treinamento e teste não são compatíveis entre si ou, se por algum erro experimental, as amostras de validação externa não são correspondentes às amostras de treinamento.

É possível solucionar este problema aplicando algum tipo de calibração no algoritmo que retorne uma matriz de probabilidades juntamente com a matriz de previsão binária, fornecendo informações mais completas ao usuário. Dessa forma, uma amostra de validação externa com características muito diferentes das amostras de treinamento pode ser descartada ou rotulada como “Indeterminada”, conforme um limite de corte de probabilidade (ou “*threshold*”) definido pelo operador.

Algoritmos baseados em margem de decisão, como o SVM, não calculam nativamente as probabilidades de classificação da matriz de previsão. No entanto, existem vários tipos de calibração que podem ser aplicadas para calcular a matriz de probabilidades [27].

Proposta por John C. Platt em 1999, a calibração de Platt é um dos métodos mais utilizados para calcular a distribuição de probabilidades das classes previstas de um classificador. Essa transformação foi elaborada com foco em SVM, mas pode ser aplicada em outros modelos de classificação.

A calibração de Platt é aplicada após o treinamento e a validação externa do modelo, pois utiliza a classe prevista das amostras e as distâncias entre as amostras

previstas e a margem de decisão do algoritmo para treinar um modelo de regressão logística [27].

Seja $P(y = 1|x)$ a probabilidade da classe da amostra ser $y = 1$, dadas suas características (ou variáveis) originais. Assim, a transformação logística pode ser escrita da seguinte forma:

$$P(y = 1|x) = \frac{1}{1 + \exp(Af(x) + B)} \quad (2)$$

Onde $f(x)$ é função da margem de decisão, e A e B são parâmetros estimados a partir do método de máxima verossimilhança, que é uma otimização sobre o mesmo conjunto de dados de treinamento do classificador original. Esta transformação tem como efeito gerar uma curva sigmoideal de probabilidades no conjunto de dados previstos. Para não ocorrer *overfitting* (quando o modelo se ajusta excessivamente aos dados de treinamento) na previsão de probabilidades, o sistema realiza uma validação cruzada internamente [27].

Dessa forma, define-se a classe $y = 1$ da amostra como $P(y = 1|x) > 0,5$ e vice-versa. No entanto, é necessário ter cautela com a previsão probabilística resultante por este método, pois nem sempre ela será compatível com a previsão do classificador, principalmente quando a probabilidade é mais próxima do limite de 0,5. Dessa forma, existe a possibilidade de uma amostra classificada como “1” ter probabilidade de 0,49. Este problema é mais comum quando há poucas amostras disponíveis para validação externa.

A calibração de Platt demonstra ser especialmente eficiente para modelos de margem de decisão, como o SVM, e algoritmos que produzem modelos de previsão com distribuições de probabilidades distorcidas. Para modelos já bem calibrados, ou seja, aqueles que já retornam previsões probabilísticas nativamente, o efeito da transformação é reduzido [28].

4. MATERIAIS E MÉTODOS

4.1. Coleta das amostras

O conjunto de dados constituiu-se de dois subconjuntos amostrais: o primeiro, de amostras coletadas das mãos de voluntários com as mãos limpas, antes de

realizarem disparos com arma de fogo, e o segundo, de amostras coletadas das mãos do mesmo grupo de voluntários após dispararem com arma de fogo.

A coleta de amostras foi baseada na metodologia proposta pelo Procedimento Operacional Padrão (POP) nº1.4, publicado em setembro de 2013 pela Secretaria Nacional de Segurança Pública (SENASP), que padroniza o procedimento de coleta de resíduos por arma de fogo [27]. O processo de coleta foi realizado com o auxílio dos peritos criminais lotados no Núcleo Regional de Criminalística (NRC/CR), da Unidade Regional de Perícias e Identificação de Costa Rica/MS (URPI/MS), da Coordenadoria Geral de Perícias do MS (CGP).

Para as amostras de disparo, a arma de fogo utilizada foi um revólver da marca Amadeo Rossi S.A, calibre .38SPL (*Special*), de fabricação brasileira [28]. A munição utilizada foi do tipo cartuchos de munição, de fabricação nacional, própria para arma de fogo de cano de alma raiada, intactos, constituídos por estojo e cápsula de espoletamento em liga metálica dourada, carregada com projétil calibre .38 SPL manufaturado em liga de chumbo nu do tipo ogival [28]. A Figura 1 mostra em detalhe a arma e munição utilizadas:

Figura 1 - a) revólver calibre .38 SPL Amadeo Rossi S.A. utilizado na etapa de disparos e b) perfil e inscrições da base da munição .38 SPL.

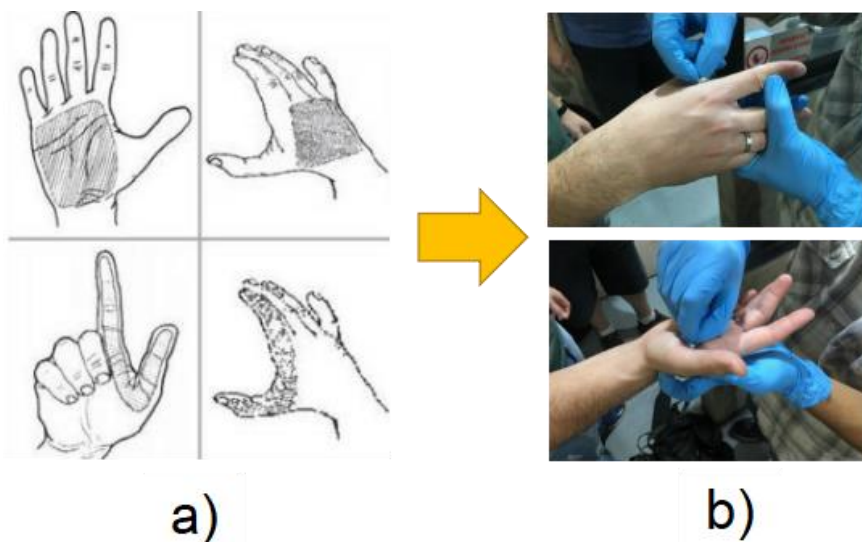


Fonte: Rodrigo Wenceslau, 2020.

A coleta de amostras foi baseada na metodologia proposta pelo Procedimento Operacional Padrão (POP) nº1.4, publicado em setembro de 2013 pela Secretaria Nacional de Segurança Pública (SENASP), que padroniza o procedimento de coleta de resíduos por arma de fogo.

Após a realização dos disparos, as amostras de GSR foram coletadas diretamente das mãos do atirador utilizando fita adesiva dupla-face, transparente, marca 3M, seguindo o protocolo demonstrado na Figura 2:

Figura 2 - a) esquema informativo sobre as regiões das mãos que a amostra foi coletada e b) foto do processo de coleta de acordo com esquema informativo.

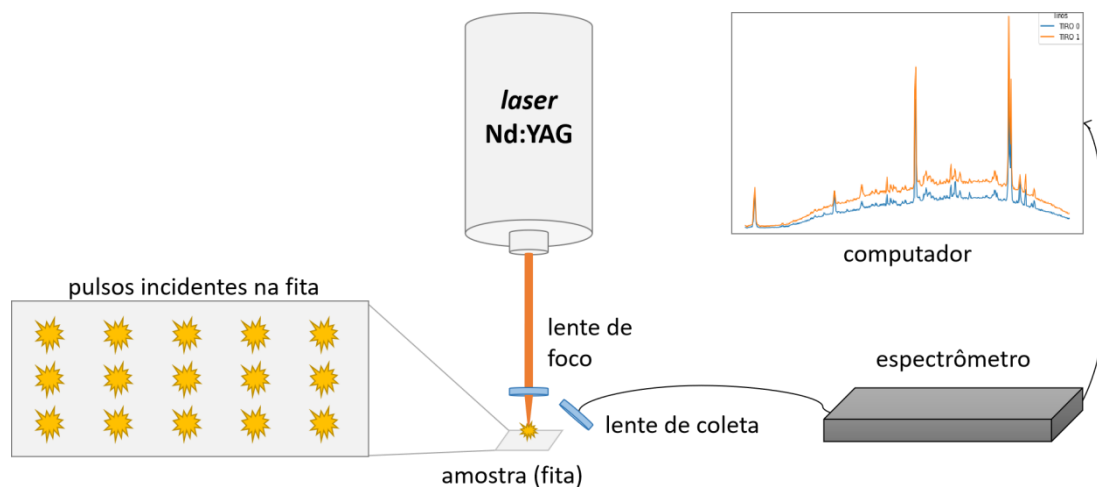


Fonte: (a) Acevedo et al., 2013; (b) Do Autor.

Em seguida, as amostras foram rotuladas e analisadas pelo método de *Laser Induced Breakdown Spectroscopy* (LIBS) na região do UV (190-300 nm). O equipamento LIBS utilizado possui *laser* principal do tipo Nd:YAG *Q-switched*, marca *Quanta-Ray®* (*Spectra Physics*), operando em um comprimento de onda de 1064 nm e largura de pulso de 8 ns. A sincronização do sistema foi realizada utilizando gerador de delay comercial com precisão de nanossegundo. O delay entre o pulso e a aquisição foi de 500 ns.

Os espectros foram registrados pelo sistema LIBS seguindo uma ordem de 15 leituras espectrais por amostra, realizadas sobre espaçamentos regulares na superfície da fita. Um único pulso foi disparado a cada leitura do equipamento LIBS para que as partículas de GSR depositadas superficialmente fossem detectadas pelo espectrômetro. O procedimento de coleta de dados espectrais está ilustrado esquematicamente na Figura 3:

Figura 3 - Modelo esquemático do processo de captura de dados realizado com sistema LIBS. No total, 15 pulsos igualmente espaçados incidiram sobre a fita.



Fonte: Do autor.

Dessa forma, garantiu-se que uma amostragem representativa de resíduos por disparo de arma de fogo fosse detectada pelo espectrômetro. No total, 240 espectros foram coletados, 60 correspondentes à classe “Não Atirador” e 180 correspondentes à classe “Atirador”. Desses 180 espectros “Atirador”, 60 correspondem a um disparo, outros 60 a dois disparos e os últimos 60 a três disparos.

4.2. Pré-processamento

Os seguintes métodos de pré-processamento foram aplicados, nesta ordem: média espectral, SNV e RB.

A média espectral foi aplicada agregando cada três espectros consecutivos de acordo com a sequência de pulsos do *laser* em cada amostra. Dessa forma, as 240 leituras espectrais brutas foram agregadas em 80 amostras, 20 amostras da classe “Não Atirador” e 60 amostras da classe “Atirador”.

Em seguida, o conjunto de dados foi normalizado de forma padrão com o método SNV, utilizando a função *StandardScaler* disponível no pacote *Scikit-Learn* (versão 1.2.2).

Na última etapa de pré-processamento, foi aplicada a correção de curva base utilizando o algoritmo RB, com raio da circunferência definido por inspeção visual em 150 *pixels*. Este algoritmo está disponível no pacote *Scikit-Image* (versão 0.19.3).

4.3. Análise não-supervisionada e seleção de transições

A dimensão do conjunto espectral foi reduzida pelo método PCA, resultando em um conjunto reduzido de variáveis ordenadas não correlacionadas. Assim, os *loadings* das 5 primeiras PCs foram calculados e inspecionados visualmente para identificação de transições espectrais relacionadas com GSR.

Dessa forma, definiu-se dois cortes espectrais com transições selecionadas que deveriam conter GSR:

- 210 a 230 nm e 255 a 265nm;
- 210 a 230 nm, 255 a 265 nm e 278 a 289 nm

Para a identificação de agrupamentos (*clusters*) entre as amostras, os *score-plots* (gráficos de dispersão) de cada conjunto espectral (completo e de transições selecionadas) foram analisados com relação as três primeiras PCs.

4.4. Análise supervisionada

Os conjuntos amostrais foram divididos aleatoriamente em subconjuntos de treino e validação, de forma que, do total de 80 amostras processadas, 70 foram selecionadas para o treinamento do modelo e 10 foram dedicadas à validação, onde 4 pertenceram à classe “Não Atirador” e 6 à classe “Atirador”.

O treinamento foi realizado pelo processo de busca extensiva (*gridsearch*), no qual automaticamente vários modelos de aprendizado de máquina são treinados sequencialmente variando parâmetros específicos definidos previamente, como variáveis internas do algoritmo e a dimensionalidade do conjunto espectral definida pela quantidade de PCs.

Neste trabalho, o algoritmo *Support Vector Machines* (SVM) foi aplicado, com a quantidade de PCs dos conjuntos variando de 1 a 20 e o parâmetro interno de regularização “C” configurado aos seguintes valores: 0.1, 1.0, 10, 100 e 1000. O resultado de cada iteração da busca extensiva foi avaliado por meio de validação cruzada *leave-one-out* (LOOCV), os modelos com a melhores acurácia e precisão no conjunto de treino foram selecionados para a etapa de validação externa.

Os três conjuntos de *ranges* espectrais foram treinados e validados para determinar a acurácia, precisão e probabilidade de previsão em cada configuração. O corte de probabilidade de acerto foi definido em 75%, portanto, amostras previstas com probabilidade abaixo desse limite foram reclassificadas como “Indeterminadas”.

4.5. Softwares utilizados

As funções e pacotes utilizados neste trabalho foram escritos em linguagem de programação *Python* (versão 3.10.12) em ambiente de desenvolvimento integrado *Google Colaboratory*.

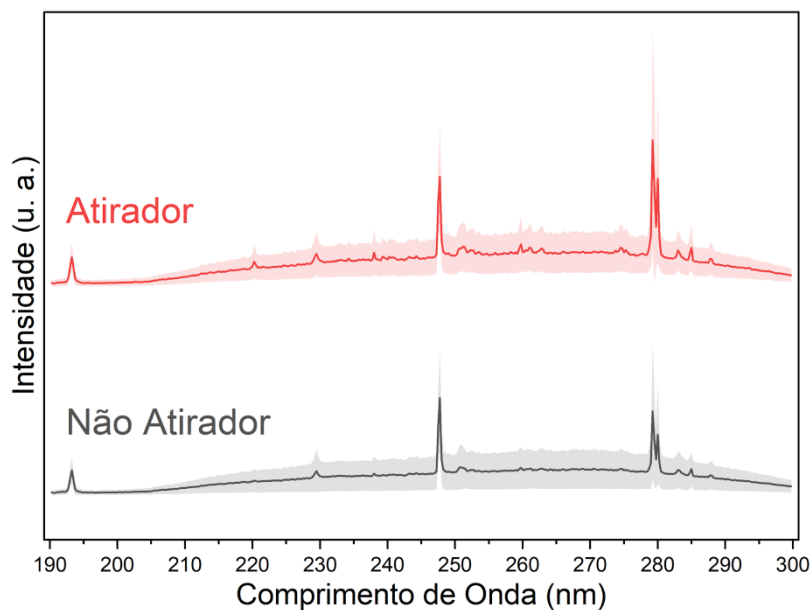
O software *Origin* (versão 2021b) foi utilizado para confecção dos gráficos de resultados.

5. RESULTADOS E DISCUSSÃO

5.1. Análise exploratória

Inicialmente realizada uma análise exploratória sobre espectros brutos sem processamento, centralizados pela média por classe, conforme mostrado na Figura 4:

Figura 4 - Média por classe dos espectros brutos. O desvio padrão da média é indicado pelo sombreado da média espectral.

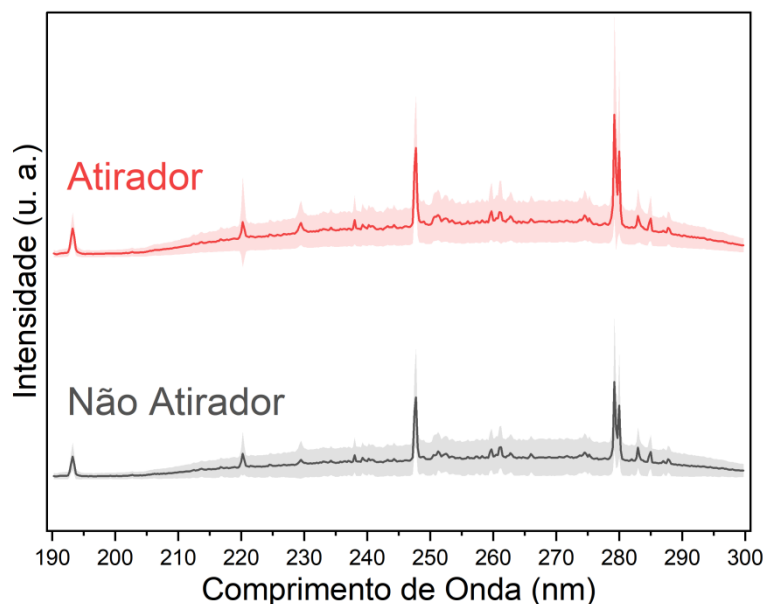


Fonte: Do autor.

Nota-se que os espectros possuem alta variabilidade, que é indicado pelo alto intervalo de desvio padrão em ambos conjuntos espectrais, visível principalmente na região sem emissões (*continuum*). É possível observar uma curva de base pronunciada a partir do comprimento de onda de 205 nm que percorre todo o conjunto de dados, provavelmente relacionada a uma alta variação de temperatura do plasma.

Além disso, foram identificadas algumas amostras com a presença de transições não relacionadas ao conjunto “Não Atirador”, provavelmente devido à erros experimentais no processo de coleta de dados. Essas amostras, também chamadas de “*outliers*” ou “*aberrantes*” ao conjunto, podem ser observadas na Figura 5 :

Figura 5 - Espectros da classe "Não Atirador" considerados *outliers* pois possuem transições relacionadas à GSR.



Fonte: Do autor.

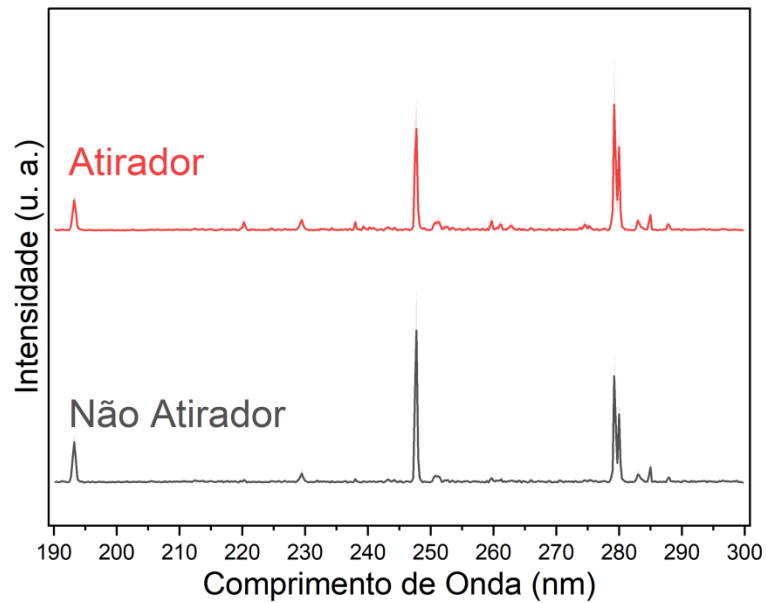
Especialmente nas regiões de 220 nm e 261 nm, ambas relacionadas ao chumbo (Pb I), nota-se a presença de picos pronunciados em ambas as classes. Essas transições não são observadas na Figura 4, pois foram omitidas para demonstrar os espectros esperados para ambas as classes.

Uma possível explicação para a presença desses valores discrepantes na classe "Não Atirador" pode ser atribuída ao ambiente altamente poluído por GSR no qual as amostras foram coletadas. Observa-se que disparos foram realizados previamente no estande de tiro, portanto, é provável que a atmosfera do local estivesse contaminada por partículas de GSR suspensas. Além da contaminação pelo ar, é possível que o simples contato das mãos com superfícies contaminadas, como paredes, papéis e bancadas, tenha sido suficiente para contaminá-las com GSR.

Devido ao baixo número de amostras coletadas para a classe "Não Atirador", foi decidido não excluir os valores discrepantes durante a análise supervisionada. O modelo de aprendizado de máquina mostrou ser sensível à escassez de amostras, resultando em um desempenho inferior em relação ao conjunto completo. De qualquer forma, os efeitos negativos desses *outliers* foram reduzidos após as etapas de pré-processamento.

A Figura 6 mostra o conjunto espectral após o pré-processamento:

Figura 6 - Média por classe dos espectros processados. O desvio padrão da média é indicado pelo sombreado abaixo da média espectral.

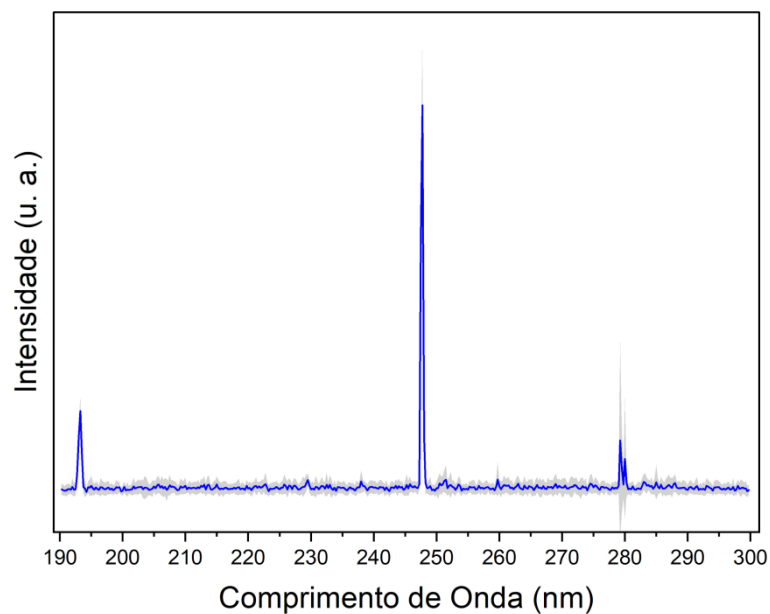


Fonte: Do autor.

Após o pré-processamento, a variabilidade dos espectros foi relativamente controlada e os picos das transições ficaram mais proeminentes. Dessa forma, o objetivo do pré-processamento foi alcançado, pois os dados heterogêneos foram normalizados, ajustados e o desvio padrão foi reduzido.

A Figura 7 apresenta a média espectral da fita adesiva dupla-face isolada utilizada na coleta:

Figura 7 - Média espectral e desvio padrão da fita adesiva dupla-face utilizada para a coleta das amostras.

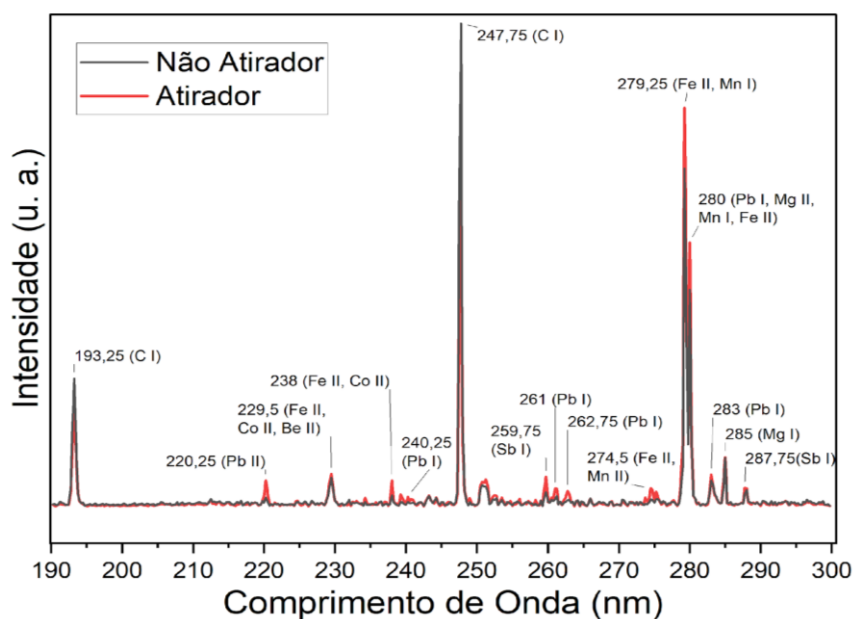


Fonte: Do autor.

Ao analisar a Figura 7, é notável que a fita dupla-face é principalmente composta por carbono, exibindo picos de emissão em 193,25 nm e 247,75 nm. No entanto, também é observado um pico em 279,25 nm, que provavelmente está relacionado a uma possível contaminação/presença de outros materiais. Importante mencionar que essas mesmas transições são identificadas tanto nos conjuntos "Atirador" quanto nos conjuntos "Não Atirador", conforme ilustrado na

Figura 8:

Figura 8 - Média espectral de cada classe sobreposta em linhas com as indexações de transições.



Fonte: Do autor.

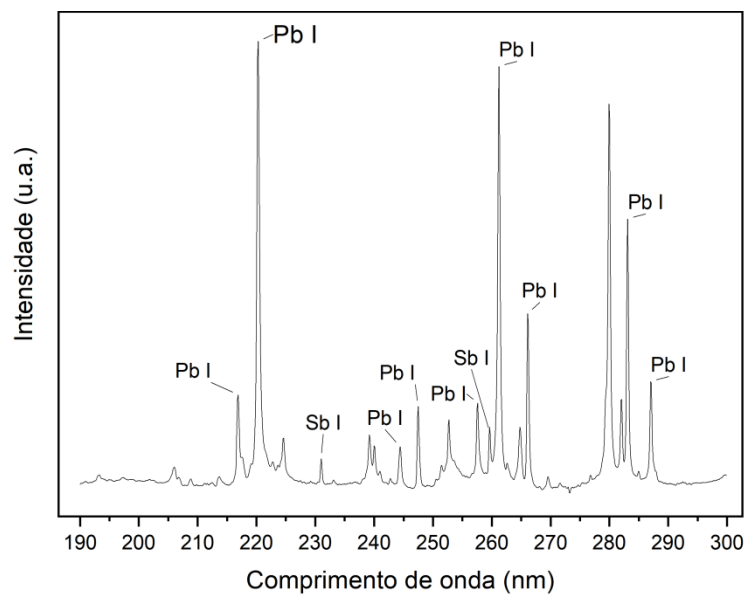
As transições relacionadas ao carbono foram detectadas em ambos conjuntos espectrais, que pode ser decorrência da composição da fita e de resíduos presentes nas mãos dos voluntários.

Nota-se a presença de várias transições relacionadas ao chumbo (Pb), como em 220,25 nm, 240,25 nm, 261 nm, 262,75 nm e 283 nm. Em menores intensidades, pode-se perceber antimônio (Sb) em 259,75 nm e 287,75 nm. Além disso, podemos

notar a presença de contaminantes (Sb e Pb) no espectro “não atirador”. Esse fato é devido a provável contaminação explicada anteriormente.

Todos esses elementos são relacionados ao projétil e, conseqüentemente, à GSR, conforme mostrado na Figura 9:

Figura 9 - Espectro coletado sobre o projétil de mesmo tipo utilizado na coleta das amostras.

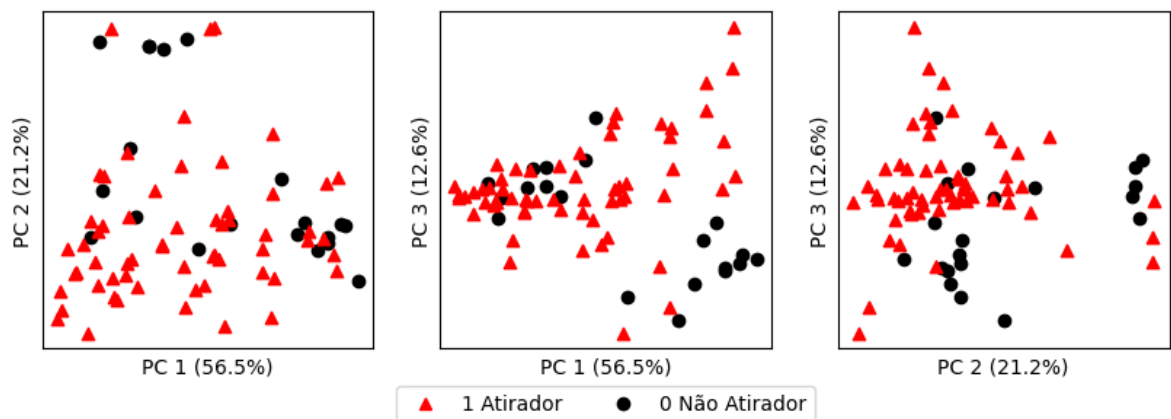


Fonte: Do autor.

5.2. Análise não-supervisionada (PCA)

Para análise de aprendizado de máquina não-supervisionado, este *range* espectral espectral foi reduzido a três PCs . A Figura 10 mostra o gráfico de dispersão (*score plot*) das amostras em dimensionalidade reduzida:

Figura 10 - Gráfico de dispersão das três primeiras PCs, seccionado em duas dimensões para melhor visualização



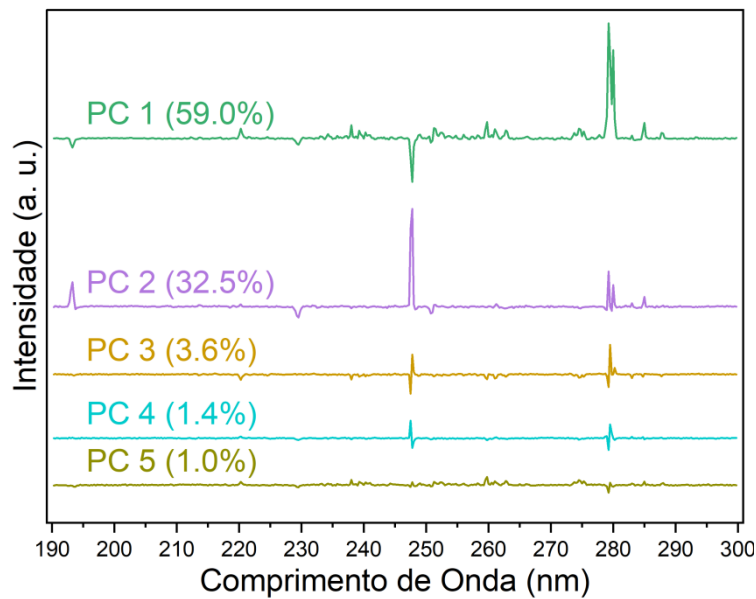
Fonte: Do autor.

Não é evidente identificar agrupamentos distintos entre as amostras ou qualquer tendência de separação na região espectral analisada. No entanto, é possível destacar os conjuntos PC1 vs. PC3 e PC2 vs. PC3, pois apresentaram uma leve tendência de separação com relação ao eixo PC3.

É relevante ressaltar que há a presença de 5 amostras "Não Atirador" que "invadiram" o *cluster* de "Atirador" nos três conjuntos. Isso é um indicativo de que essas amostras podem estar contaminadas por partículas de GSR devido a erros no processo de coleta conforme discutido anteriormente.

Apesar de ser uma análise feita sobre poucos pares de Componentes Principais, ela indica que há transições por todo o *range* espectral que estão impactando positivamente e negativamente na separação entre as classes. Assim, a análise de *loadings* foi aplicada para observar a variância das transições, conforme apresentado na Figura 11.

Figura 11 - Loadings das cinco primeiras componentes principais para o intervalo espectral completo.



Fonte: Do autor.

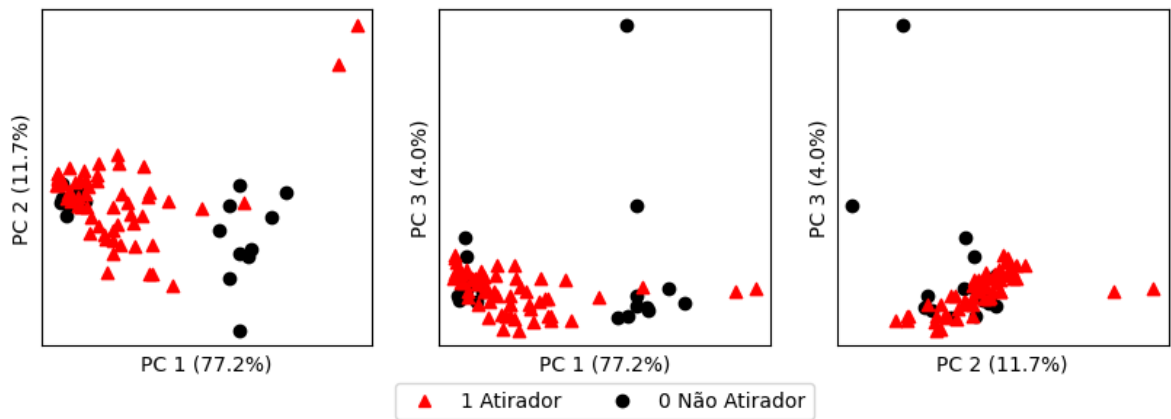
Conforme observou-se na Figura 10, a tendência de separação ficou mais concentrada na PC3. Esse resultado é corroborado pela Figura 11, pois nota-se uma transição importante em 220 nm (Pb) nas PC1 e PC3 que não está presente na PC2. Além disso, nota-se três transições (259,75 nm, 261 nm e 262,75 nm) relacionadas à GSR presentes principalmente em PC1 e PC3.

Dessa forma, após observar os espectros processados e os *loadings* das componentes principais, dois subconjuntos de regiões espectrais foram selecionados com o objetivo de alcançar uma melhor performance na separação supervisionada:

- Região 1: Seleção de 210 nm a 230 nm e 255 nm a 265 nm, com foco na transição intensa de Pb I.
- Região 2: Seleção de 210 nm a 230 nm, 255 nm a 265 nm e 278 nm a 289 nm, com ênfase adicional nas transições de Pb I e Sb I e de Pb I na extremidade final do *range*

A Figura 12 mostra o gráfico de dispersão da distribuição de amostras das três primeiras PCs para o Região 1:

Figura 12 - Gráfico de dispersão das três primeiras PCs para o Região 1, seccionado em duas dimensões para melhor visualização.

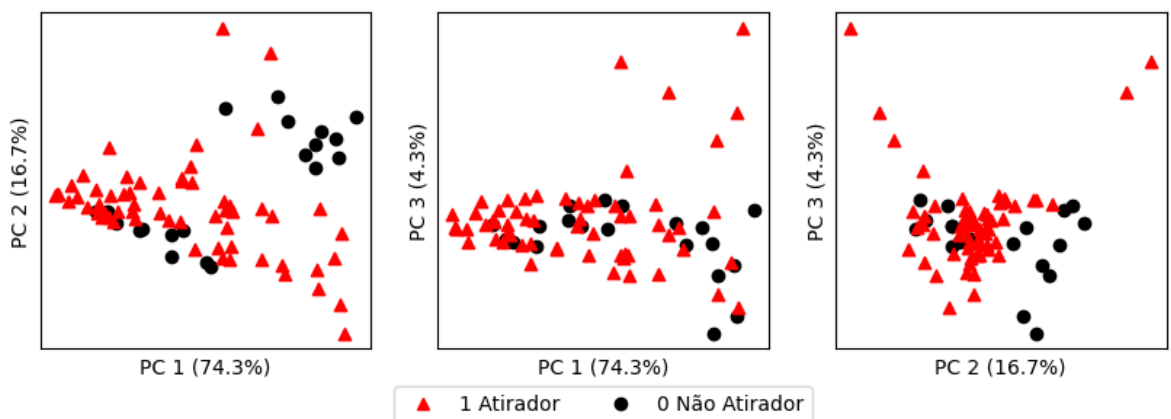


Fonte: Do autor.

Na Análise de Componentes Principais, o Região 1 não demonstrou tendência de agrupamento entre as classes, com a PC1 desempenhando um pouco melhor a separação entre as variáveis.

A Figura 13 mostra a o gráfico de dispersão da distribuição de amostras das três primeiras PCs para o Região 2:

Figura 13 - Gráfico de dispersão das três primeiras PCs para o Região 2, seccionado em duas dimensões para melhor visualização.



Fonte: Do autor.

Observa-se que os agrupamentos são menos pronunciados neste conjunto, com tendências mais aparentes nas distribuições da PC1 com a PC3 e PC2 com a PC3. Assim, nota-se que a PC3 desempenhou um papel mais importante na separação das amostras.

Todos os conjuntos de dados foram submetidos à etapa de treinamento e validação do modelo de aprendizado de máquina. Os modelos foram treinados utilizando o algoritmo SVM, variando a dimensionalidade das regiões espectrais pelo número de PCs e o parâmetro de regularização "C" do SVM. O melhor modelo selecionado em validação interna *LOO-CV* para cada região espectral foi testado em validação externa.

5.3. Desempenho dos modelos de aprendizado de máquina

A Tabela I mostra as métricas de avaliação e os parâmetros otimizados dos modelos de aprendizado de máquina treinados:

Tabela I - Modelos treinados e otimizados sob a métrica "Acurácia" para cada intervalo espectral.

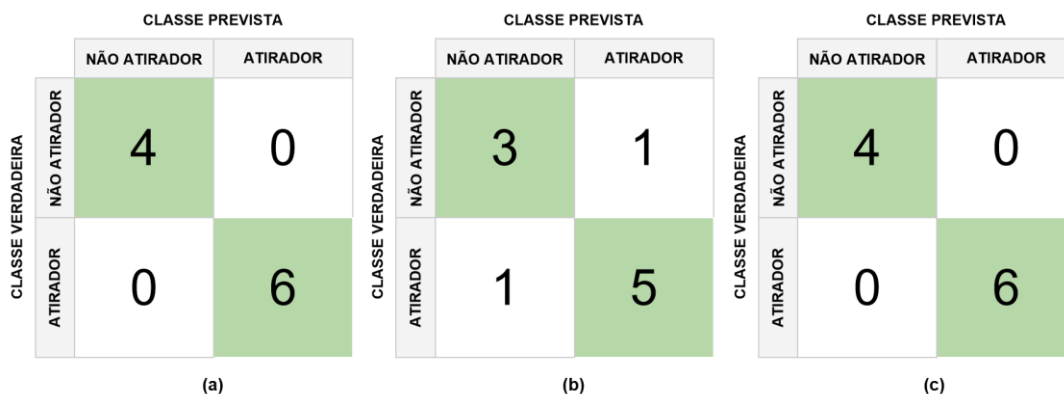
Resultados obtidos em validação externa.

Espectros	PCs	Parâmetro "C"	Acurácia	Precisão
Completo	7	1	100%	100%
Região 1	18	100	80%	83%
Região 2	2	100	100%	100%

Cada linha da Tabela I representa de forma geral o melhor modelo treinado para cada região espectral, de acordo com hiperparâmetros otimizados "PCs" e "C" correspondentes.

A matriz confusão de classificação para cada modelo pode ser avaliada, como mostrado na Figura 14:

Figura 14 - Matriz confusão para os melhores resultados obtidos em validação externa para cada região espectral. a) Região completo; b) Região 1; c) Região 2



Fonte: Do autor.

As seleções (a) e (c) maximizaram a métrica de precisão, alcançando 100%. Portanto, em uma análise inicial, esses modelos tiveram sucesso em eliminar completamente os falsos positivos, ou seja, as amostras "limpas" de GSR que poderiam ser erroneamente classificadas como "Atirador".

Sob a mesma perspectiva, o modelo (b) obteve a pior performance, com uma precisão de 83%, o que significa que classificou erroneamente uma amostra sem GSR como "Atirador". Além disso, este modelo precisou de mais PCs (18) para capturar as nuances entre as amostras, o que não foi necessário nos outros modelos. Isso pode indicar que a fronteira de decisão entre as amostras pode ter sido mais complexa para o modelo (b).

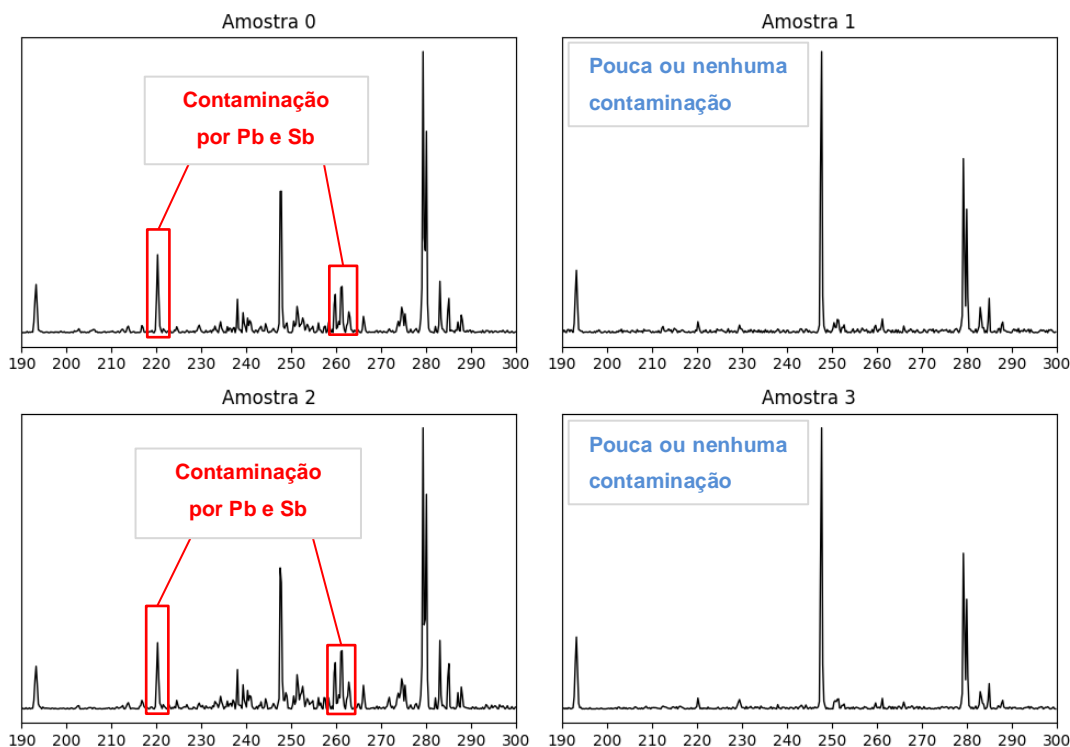
Mas, é importante lembrar que as amostras utilizadas no treinamento e validação externa dos modelos possuem um certo grau de contaminação, tanto para o conjunto "Não Atirador" com transições de Pb e Sb presentes, quanto para o conjunto "Atirador", com amostras limpas de GSR devido a incidência do laser em locais vazios da fita adesiva.

5.4. Análise detalhada de probabilidades

A análise de probabilidades possibilita um estudo mais detalhado sobre o desempenho dos modelos de aprendizado de máquina, principalmente com relação a generalização de amostras com características fora do padrão. Para facilitar essa análise, as 10 amostras de validação externa foram numeradas unicamente de 0 a 9.

Dessa forma, a Figura 15 mostra os espectros de emissão das amostras de validação externa da classe "Não Atirador":

Figura 15 - Amostras “Não Atirador” selecionadas para validação externa. Nota-se grande quantidade de contaminação por Pb e Sb nas amostras “0” e “2”, já nas amostras “1” e “3” podem ser identificadas pouca ou nenhuma contaminação.

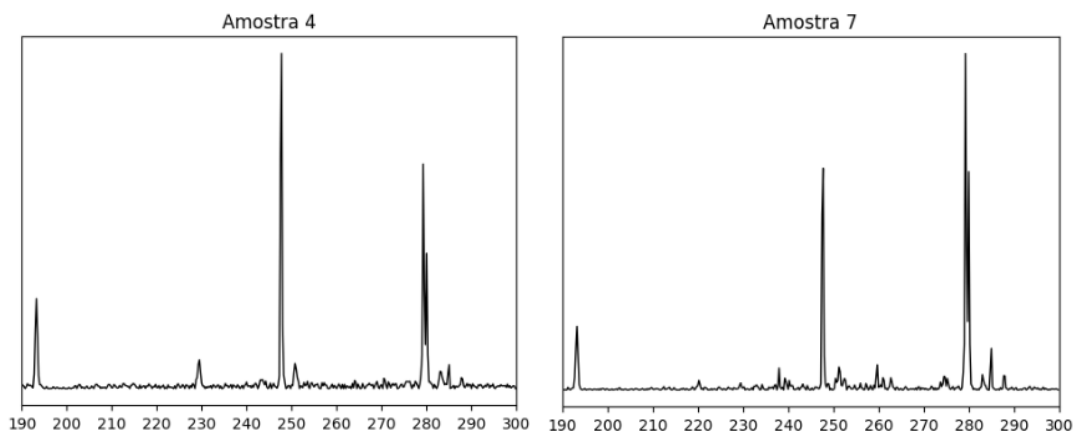


Fonte: Do autor.

Como era esperado, é possível identificar uma alta proporção de contaminação por Pb e Sb em amostras das classes “Não Atirador”. Um classificador flexível deve ser capaz de limitar os efeitos desses contaminantes nos resultados e não interpretar essas amostras como verdadeiras de sua classe, mas sim “Indeterminadas”.

Além disso, destacam-se as amostras “4” e “7”, também do conjunto de validação externa, que são rotuladas “Atirador”, mas possuem pouca ou nenhuma contaminação por GSR, conforme mostrado na Figura 16:

Figura 16 - Amostras com pouca ou sem GSR rotuladas como "Atirador", portanto, Outliers, presentes no conjunto de validação externa.



Fonte: Do autor.

É esperado que algumas amostras do conjunto "Atirador" apresentassem ausência de GSR devido ao fato do *laser* incidir sobre partes "vazias" da fita adesiva, onde não havia partículas de GSR. Sendo assim, ao treinar o modelo com esses *outliers*, espera-se que as previsões sejam melhor calibradas com relação às probabilidades.

A presença de conjuntos de dados altamente contaminados pode acarretar problemas significativos no processo de modelagem. Por esse motivo, diversas abordagens foram testadas com o intuito de reduzi-los ou solucioná-los.

Inicialmente, os modelos foram testados após realizar a remoção de todas as amostras contaminadas, mas isso resultou em um desempenho muito inferior em todas as seleções espectrais ("*underfitting*") no conjunto de validação externa devido à baixa quantidade de amostras disponíveis no conjunto "Não Atirador".

De maneira similar, também foram testados modelos utilizando amostras selecionadas livres de GSR, no conjunto de validação externa, visando minimizar as ocorrências de falsos negativos altamente confiáveis (previsões incorretas com alta probabilidade). Entretanto, essa abordagem resultou em uma considerável redução no desempenho geral de todos os modelos. Provavelmente, isso ocorreu devido à remoção dessas amostras do conjunto de treinamento acarretar na diminuição da capacidade de generalização, já que menos amostras "limpas" foram utilizadas como referência.

Por fim, observou-se que em alguns modelos treinados, as seleções espectrais que foram escolhidas para destacar as diferenças entre as amostras também

evidenciaram as faixas contaminadas, resultando em uma redução no desempenho e em altas taxas de falsos negativos.

Na Tabela II, são mostradas probabilidades de previsão das amostras de validação externa para a classe real “Não Atirador”, no intervalo espectral completo:

Tabela II - Probabilidades de classificação para a classe real “Não Atirador” no intervalo espectral completo.

Amostra	Classe Prevista	Probabilidade	Aprovação P > 75%
2	Não Atirador	96%	SIM
3	Não Atirador	92%	SIM
1	Não Atirador	83%	SIM
0	Não Atirador	77%	SIM

Fonte: Do autor.

A Tabela III mostra as probabilidades de classificação das amostras previstas em validação externa para a classe real “Atirador” no intervalo espectral completo:

Tabela III - Probabilidades de classificação para a classe real Atirador no intervalo espectral

Amostra	Classe Prevista	Probabilidade	Aprovação P > 75%
7	Atirador	94%	SIM
9	Atirador	92%	SIM
5	Atirador	91%	SIM
8	Atirador	91%	SIM
6	Atirador	90%	SIM
4	Atirador	79%	SIM

completo.

Fonte: Do autor.

Nota-se que a acurácia geral de 100% é um resultado enganador em primeiro momento, pois as amostras contaminadas foram classificadas incorretamente com alto grau de confiabilidade inclusive passando pela faixa de aprovação de P > 75%.

Esse é um indicativo de que, para o conjunto sem seleção de transições, as amostras contaminadas ganham um peso expressivo na delimitação da margem de decisão do SVM. Dessa forma, este não deve ser o modelo final escolhido, pois ele não tem capacidade de limitar o efeito de amostras contaminadas sobre a margem de decisão do algoritmo.

As probabilidades de previsão obtidas pelo modelo treinado com o intervalo espectral “Região 1”, para a classe real “Não Atirador” estão mostrados na Tabela IV:

Tabela IV - Probabilidades de classificação para a classe real “Não Atirador” no intervalo espectral “Região 1”.

Amostra	Classe Prevista	Probabilidade	Aprovação P > 75%
3	Não Atirador	75%	SIM
1	Não Atirador	70%	NÃO
0	Não Atirador	67%	NÃO
2	Atirador	39%	NÃO

Fonte: Do autor.

A Tabela V mostra as probabilidades de classificação das amostras previstas para a classe real “Atirador” no intervalo espectral completo:

Tabela V - Probabilidades de classificação para a classe real “Atirador” no intervalo espectral “Região 1”.

Amostra	Classe Prevista	Probabilidade	Aprovação P > 75%
9	Atirador	100%	SIM
5	Atirador	98%	SIM
8	Atirador	97%	SIM
6	Atirador	96%	SIM
4	Atirador	74%	NÃO
7	Atirador	51%	NÃO

Fonte: Do autor.

Este modelo, diferentemente do anterior, lidou melhor com as amostras contaminadas. As amostras 3 e 1, obtiveram as maiores probabilidades de previsão do conjunto “Não Atirador”, o que é um sinal positivo.

Apesar do corte de probabilidade ter reprovado a amostra 1, que é realmente do tipo "Não Atirador", esse mesmo modelo também reprovou as amostras 0 e 2, que estão altamente contaminadas por GSR, chegando a classificar a segunda como "Atirador", o que é um resultado esperado devido às suas características. Este é um indicativo de que, para esta região selecionada, amostras contaminadas possuíram um peso menor na delimitação da margem de decisão do SVM.

Além disso, ao analisar o conjunto "Atirador", observa-se uma baixa confiabilidade nas previsões das amostras 4 e 7, que são outliers do grupo e não atingiram o limite mínimo de probabilidade estabelecido. Consequentemente, as amostras 1, 0, 2, 4 e 7 foram consideradas "Indeterminadas" neste modelo.

A Tabela VI mostra as probabilidades de classificação obtidas pelo modelo treinado sobre a “Região 2” de transições:

Tabela VI - Probabilidades de classificação para a classe real Não Atirador no intervalo espectral “Região 2”.

Amostra	Classe Prevista	Probabilidade	Aprovação P > 75%
0	Não Atirador	96%	SIM
3	Não Atirador	94%	SIM
2	Não Atirador	88%	SIM
1	Não Atirador	53%	NÃO

Fonte: Do autor.

A Tabela VII mostra as probabilidades de classificação das amostras previstas em validação externa para a classe real “Atirador” no intervalo espectral “Região 2”:

Tabela VII - Probabilidades de classificação para a classe real Atirador no intervalo espectral “Região 2”.

Amostra	Classe Prevista	Probabilidade	Aprovação P > 75%
5	Atirador	99%	SIM
9	Atirador	96%	SIM
7	Atirador	94%	SIM
8	Atirador	84%	SIM
6	Atirador	79%	SIM
4	Atirador	63%	NÃO

Fonte: Do autor.

Os resultados obtidos por este modelo foram semelhantes aos obtidos pelo primeiro, ou seja, apesar das boas métricas de acurácia e precisão, ele não teve um bom desempenho com amostras contaminadas e *outliers*. No caso da classe "Não Atirador", este modelo reprovou uma das amostras sem GSR e aprovou com alta confiança as amostras contaminadas.

Já para a classe "Atirador", um dos *outliers* foi reprovado, enquanto o outro foi aprovado com alta confiança. Uma hipótese para esse comportamento é que o intervalo de 278 nm a 289 nm tenha menor importância na caracterização das amostras com a presença de contaminações.

Dessa forma, o modelo de previsão final mais indicado neste conjunto de dados é o modelo treinado sob a seleção da Região 1 do espectro, pois foi o único modelo que apresentou boa generalização com a capacidade de desconsiderar as amostras contaminadas e *outliers* nos resultados.

6. CONCLUSÃO

A partir da análise exploratória, foi possível identificar os elementos químicos característicos da GSR, principalmente Pb e Sb. Portanto, o protocolo de coleta de amostras com fita adesiva demonstrou ter qualidade satisfatória para coletar as partículas de GSR das mãos de pessoas que dispararam com arma de fogo.

Foi demonstrado que os espectros obtidos por meio do sistema LIBS reduzido apresentaram a qualidade necessária para a análise multivariada, tanto na etapa de aprendizado de máquina não-supervisionado quanto na etapa de aprendizado de máquina supervisionado. No entanto, para melhorar a qualidade da separação entre as classes, foi necessário submeter o conjunto de dados a um processo de pré-processamento, que envolveu média espectral, SNV e RB.

Além disso, ao analisar as três regiões espectrais propostas, foi possível observar que a seleção de transições características de GSR desempenhou um papel importante no aumento da confiança dos modelos. Essa seleção permitiu evidenciar transições que contribuíram para a melhoria do desempenho do modelo, ao mesmo tempo em que revelaram faixas de contaminação e outliers.

A partir da análise de probabilidades e *outliers*, foi demonstrado que as métricas de acurácia e precisão nem sempre são os melhores indicadores para se escolher um modelo de classificação. No conjunto de dados analisado neste trabalho, haviam muitas amostras contaminadas e *outliers*, que causou enviesamento dos modelos. Dessa forma, dos três modelos treinados, dois alcançaram 100% de acurácia e precisão em validação externa (Região completa e Região 2), mas foram extremamente enviesados pelas amostras contaminadas, reduzindo as suas capacidades de generalização.

Por outro lado, o modelo de aprendizado de máquina que “mais errou” as previsões (Região 1), também foi o que forneceu mais flexibilidade com relação a análise de probabilidades, dando poder ao operador do sistema de identificar as amostras “Indeterminadas”.

Claramente, é desejável ter menos amostras classificadas como "Indeterminadas", pois isso significa que elas serão excluídas da decisão final. No entanto, devido à falta de disponibilidade de muitas amostras nesse conjunto de dados e à alta contaminação por GSR, a melhor abordagem, que não seja a necessidade de coletar novos dados e refazer o experimento, é treinar um modelo suficientemente

flexível para excluir amostras pouco confiáveis. No entanto, modelos que simulam contaminação por Pb e variabilidade no espectros, estariam mais perto da realidade e simulariam melhor um ambiente real de coleta. Dessa forma, caso o protocolo seja estendido a aplicações e treinamento em campo, um realidade similar a encontrada nos conjuntos de dados analisados aqui seria esperado.

Conforme foi demonstrado tanto neste trabalho quanto na literatura disponível, o método de coleta de partículas de GSR com fita adesiva é eficiente e confiável. Inclusive, observou-se que esse método foi capaz de coletar GSR das mãos de pessoas que não efetuaram disparos no dia, provavelmente devido à contaminação pela atmosfera do local e ao toque em superfícies contaminadas.

O modelo de aprendizado de máquina pode atingir uma acurácia ainda maior se for treinado em dados mais precisos e não contaminados. Isso permitiria aumentar o limite de probabilidade para as classes indeterminadas. O limite definido em 75% foi escolhido com base na qualidade das amostras coletadas, mas com dados melhores esse limite poderia ser aumentado, incrementando a confiança do perito ao realizar a decisão final.

Assim, o protocolo apresentado demonstrou ser uma alternativa promissora para a identificação de elementos químicos característicos da GSR. Sua vantagem de permitir a análise de muitas amostras de uma só vez, ao contrário do método padrão ouro MEV-EDS, destaca-se como uma contribuição significativa para a eficiência e escalabilidade dessa técnica em aplicações práticas em laboratórios de perícia.

Para futuros trabalhos, é fundamental aprimorar o protocolo de coleta de amostras com fita adesiva, com o intuito de minimizar a contaminação e assegurar a qualidade dos dados obtidos. A coleta dessas amostras em locais externos aos ambientes onde ocorreram os disparos por arma de fogo é crucial para estabelecer um banco de dados confiável de “controle”.

Adicionalmente, para simular um ambiente operacional realista, torna-se necessário treinar o modelo utilizando diversas variedades de contaminantes para calibrar o limite de probabilidade necessário para a definição da classe “Indeterminada”, minimizando a taxa de falsos positivos.

Além disso, explorar outros algoritmos de aprendizado de máquina que possuam a previsão probabilística por padrão, como redes neurais e *random forest*, poderia contribuir para uma avaliação mais robusta das amostras classificadas como

"Indeterminadas", permitindo a definição de limites de probabilidade mais altos, o que consequentemente aumentaria a confiabilidade das classificações e otimizaria a capacidade de generalização do modelo.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Fórum Brasileiro de Segurança Pública; ARMAS DE FOGO E HOMICÍDIOS NO BRASIL. Disponível em: <<https://forumseguranca.org.br/wp-content/uploads/2022/09/informe-armas-fogo-homicidios-no-brasil.pdf>> . Acesso em: 01/12/2022.
- [2] INSTITUTO SOU DA PAZ. Onde Mora a Impunidade?. 2022.
- [3] Ribeiro, et al. “Lesões Provocadas Por Armas de Fogo Atendidas Em Serviços de Urgência e Emergência Brasileiros”. *Ciência & Saúde Coletiva* 22, nº 9 (setembro de 2017): 2851–60. <https://doi.org/10.1590/1413-81232017229.16492017>.
- [4] INSTITUTO SOU DA PAZ. **APRENDIZADOS SOBRE O ESCLARECIMENTO DE HOMICÍDIOS NO BRASIL: PRÁTICAS PROMISSORAS E CAMINHOS A SEGUIR**. 2021.
- [5] López-López, María, César Alvarez-Llamas, Jorge Pisonero, Carmen García-Ruiz, e Nerea Bordel. “**An Exploratory Study of the Potential of LIBS for Visualizing Gunshot Residue Patterns**”. *Forensic Science International* 273 (2017): 124–31. <https://doi.org/10.1016/j.forsciint.2017.02.012>.
- [6] Dalby, Oliver, David Butler, e Jason W. Birkett. “**Analysis of Gunshot Residue and Associated Materials-A Review**”. *Journal of Forensic Sciences* 55, nº 4 (2010): 924–43. <https://doi.org/10.1111/j.1556-4029.2010.01370.x>.
- [7] Tarifa, Anamary, e José R. Almirall. “**Fast Detection and Characterization of Organic and Inorganic Gunshot Residues on the Hands of Suspects by CMV-GC-MS and LIBS**”. *Science & Justice* 55, nº 3 (2015): 168–75. <https://doi.org/10.1016/j.scijus.2015.02.003>.
- [8] Dockery, Christopher R., e Scott R. Goode. “**Laser-Induced Breakdown Spectroscopy for the Detection of Gunshot Residues on the Hands of a Shooter**”. *Applied Optics* 42, nº 30 (2003): 6153. <https://doi.org/10.1364/AO.42.006153>.
- [9] Gassner, Anne-Laure, e Céline Weyermann. “**LC-MS Method Development and Comparison of Sampling Materials for the Analysis of Organic Gunshot Residues**”. *Forensic Science International* 264 (2016): 47–55. <https://doi.org/10.1016/j.forsciint.2016.03.022>.
- [10] Romolo, Francesco Saverio, e Pierre Margot. “**Identification of Gunshot Residue: A Critical Review**”. *Forensic Science International*, 2001, 17.

[11] Shrivastava, Priya, V. K. Jain, e Suman Nagpal. 2021. "**Gunshot Residue Detection Technologies—a Review**". Egyptian Journal of Forensic Sciences 11 (1): 11. <https://doi.org/10.1186/s41935-021-00223-9>.

[12] Wolten, et al. **PARTICLE ANALYSIS FOR THE DETECTION OF GUNSHOT RESIDUE, 1 - SCANNING ELECTRON MICROSCOPY/ENERGY DISPERSIVE X-RAY CHARACTERIZATION OF HAND DEPOSITS FROM FIRING**. Journal of Forensic Sciences Volume: 24 Issue: 2. (1979) Pages: 409-422.

[13] Bueno, Justin, e Igor K. Lednev. 2014. "**Attenuated Total Reflectance-FT-IR Imaging for Rapid and Automated Detection of Gunshot Residue**". Analytical Chemistry 86 (7): 3389–96. <https://doi.org/10.1021/ac4036718>.

[14] Bueno, Justin, Lenka Halámková, Alexander Rzhetskii, e Igor K. Lednev. 2018. "**Raman Microspectroscopic Mapping as a Tool for Detection of Gunshot Residue on Adhesive Tape**". Analytical and Bioanalytical Chemistry 410 (28): 7295–7303. <https://doi.org/10.1007/s00216-018-1359-1>.

[15] Yue, Zengqi, Chen Sun, Fengye Chen, Yuqing Zhang, Weijie Xu, Sahar Shabbir, Long Zou, et al. 2021. "**Machine Learning-Based LIBS Spectrum Analysis of Human Blood Plasma Allows Ovarian Cancer Diagnosis**". Biomedical Optics Express 12 (5): 2559. <https://doi.org/10.1364/BOE.421961>.

[16] Sherbini, Ashraf M EL, Mohamed M Hagra, Hania H Farag, e Mohamed R M Rizk. 2013. "**Diagnosis and Classification of Liver Cancer Using LIBS Technique and Artificial Neural Network**" 4 (5): 6.

[17] Hussain Shah, Syed Kifayat, Javed Iqbal, Pervaiz Ahmad, Mayeen Uddin Khandaker, Sirajul Haq, e Muhammad Naeem. 2020. "**Laser Induced Breakdown Spectroscopy Methods and Applications: A Comprehensive Review**". Radiation Physics and Chemistry 170 (maio): 108666. <https://doi.org/10.1016/j.radphyschem.2019.108666>.

[18] Carvalho, Alexandrina A.C., Daniel M. Silvestre, Flávio O. Leme, Juliana Naozuka, Danielle P. Intima, e Cassiana S. Nomura. 2019. "**Feasibility of Measuring Cr(III) and Cr(VI) in Water by Laser-Induced Breakdown Spectroscopy Using Ceramics as the Solid Support**". Microchemical Journal 144 (janeiro): 33–38. <https://doi.org/10.1016/j.microc.2018.08.031>.

[19] Morais, et al. 2018. "**Laser-Induced Breakdown Spectroscopy Determination of K in Biochar-Based Fertilizers in the Presence of Easily Ionizable Element**". *Talanta* 188: 199–202. <https://doi.org/10.1016/j.talanta.2018.05.089>.

[20] Larios, Gustavo S., Gustavo Nicolodelli, Giorgio S. Senesi, Matheus C. S. Ribeiro, Alfredo A. P. Xavier, Débora M. B. P. Milori, Charline Z. Alves, Bruno S. Marangoni, e Cícero Cena. 2020. "**Laser-Induced Breakdown Spectroscopy as a Powerful Tool for Distinguishing High- and Low-Vigor Soybean Seed Lots**". *Food Analytical Methods* 13 (9): 1691–98. <https://doi.org/10.1007/s12161-020-01790-8>.

[21] Vander Pyl, C., Keith Morris, Luis Arroyo, e Tatiana Trejos. 2020. "**Assessing the Utility of LIBS in the Reconstruction of Firearm Related Incidents**". *Forensic Chemistry* 19 (junho): 100251. <https://doi.org/10.1016/j.forc.2020.100251>.

[22] Zeaiter, M, e D Rutledge. "**3.04 Preprocessing Methods**", [s.d.], 111.

[23] Franco, Marco Aurélio de Menezes, Débora Marcondes Bastos Pereira Milori, e Paulino Ribeiro Villas Boas. "**Comparison of Algorithms for Baseline Correction of LIBS Spectra for Quantifying Total Carbon in Brazilian Soils**". arXiv, 19 de abril de 2018. [<http://arxiv.org/abs/1805.03695>]

[24] Syvilay, D., N. Wilkie-Chancellier, B. Trichereau, A. Texier, L. Martinez, S. Serfaty, e V. Detalle. "**Evaluation of the Standard Normal Variate Method for Laser-Induced Breakdown Spectroscopy Data Treatment Applied to the Discrimination of Painting Layers**". *Spectrochimica Acta Part B: Atomic Spectroscopy* 114 (dezembro de 2015): 38–45. <https://doi.org/10.1016/j.sab.2015.09.022>.

[25] Liland, Kristian Hovde, Trygve Almøy, e Bjørn-Helge Mevik. "**Optimal Choice of Baseline Correction for Multivariate Calibration of Spectra**". *Applied Spectroscopy* 64, no 9 (2010): 1007–16. <https://doi.org/10.1366/000370210792434350>.

[26] MÜLLER, Andreas C.; GUIDO, Sarah. **Introduction to Machine Learning with Python**. Sebastopol, CA: O'Reilly Media, Inc., 2016.

[27] PLATT, John C. **Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods**. Microsoft Research (1999).

[28] NICULESCU-MIZIL, Alexandru; CARUANA, Rich. **Predicting Good Probabilities With Supervised Learning.** In Proceedings of the 22nd international conference on Machine learning (ICML 2005). Association for Computing Machinery, New York, NY, USA, 625–632. <https://doi.org/10.1145/1102351.1102430>