| | |
|---|---|
| Document Title: | Development and Evaluation of a Contrastive Learning Framework with Applications to Micromorphometry of Aluminum Powder used in Explosives |
| Author(s): | Danica Ommen, Christopher Saunders, JoAnn Buscaglia |
| Document Number: | 308221 |
| Date Received: | December 2023 |
| Award Number: | 2019-DU-BX-4011 |

# Development and Evaluation of a Contrastive Learning Framework with Applications to Micromorphometry of Aluminum Powder used in Explosives

Danica Ommen[1]
Chris Saunders[2]
JoAnn Buscaglia[3]

[1]Iowa State University
[2]South Dakota State University
[3]FBI Laboratory Division

13 June 2023

## Funding

## Disclaimer

## Acknowledgements

# Preamble


## Funding

This research was supported in part by appointments to the Visiting Scientist Program at the Federal Bureau of Investigation Laboratory Division, administered by the Oak Ridge Institute for Science and Education, through an interagency agreement between the US Department of Energy and the FBI.

This research was supported in part by a grant from the NIJ, Award # 2019-DU-BX-4011.

Travel was supported by NIJ Contract # GS-00F-439GA, Order # 15NJD21F00000011.


## Disclaimer


## Acknowledgements

# Preamble


## Funding


## Disclaimer

Names of commercial products are provided for identification purposes only, and inclusion does not imply endorsement of the manufacturer, or its products or services by the FBI. The views expressed are those of the author(s) and do not necessarily reflect the official policy or position of the FBI, the U.S. Government, or the NIJ.


## Acknowledgements

## Funding

## Disclaimer

## Acknowledgements

The authors gratefully acknowledge the support of Visiting Scientists and Honors Interns at the FBI Laboratory, especially our colleagues Jack Hietpas, Jena Baldaino, and Kayla Moquin for their contributions to this research project.

# Introduction

- Al powder is often used as a fuel in IEDs.
- Individuals attempting to make IEDs often obtain it from legitimate commercial products or make it themselves using readily available Al starting materials.
- The characterization and differentiation between sources of Al powder may provide investigative and intelligence value.
- Our goal is to use micromorphometric features of Al powder particles from a variety of different source types and apply statistics!

Considerations include:

## Hierarchy of Propositions

## Common-source vs. Specific-source

## Closed-set vs. Open-set

# Forensic Identification of Source

## Hierarchy of Propositions

- **Offense Level**: The individual(s) made the IED.
- **Activity Level**: The individual(s) supplied the Al powder to the IED maker.
- **Source Level**: The Al powder in the IED was produced by manufacturer X.
- **Sub-Source Level**: The Al powder in the IED came from this particular bulk Al powder sample.

## Common-source vs. Specific-source

## Closed-set vs. Open-set

# Forensic Identification of Source

## Hierarchy of Propositions

## Common-source vs. Specific-source

- **Specific-source**: The Al powder in the IED came from this particular bulk sample.
- **Common-source**: The Al powder in the first and second IED come from the same bulk sample.

## Closed-set vs. Open-set

## Hierarchy of Propositions

## Common-source vs. Specific-source

## Closed-set vs. Open-set

- **Closed-set**: The alternative population consists of a finite number of known sources with observed samples.
- **Open-set**: It is not possible to list and/or observe samples from all possible sources in the alternative population.

Choices for this presentation:

**Hierarchy of Propositions**

Sub-Source Level

**Common-source vs. Specific-source**

Common-source

**Closed-set vs. Open-set**

Open-set

# Al Powder Data

| Sample Type | # of Samples |
|---|---|
| Ball-milled Al Foil | 29 |
| Al-containing Spray Paint | 36 |
| Binary Exploding Targets | 40 |
| Industrial Manufacturers | 47 |
| Other | 2 |
| **Total** | **154** |

# Feature Extraction Method

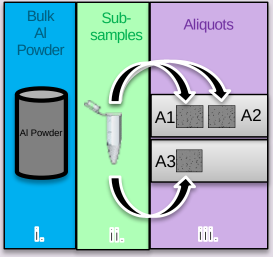
A three step process consisting of:

**Sample Preparation**

**Automated Imaging**

**Particle Micromorphometry**

# Sample Preparation




i. Bulk Al powder was thoroughly mixed/re-distributed in sample tube
ii. A micro-spatula of sample was placed into a microtube containing Permount®
iii. A 10μL aliquot was taken from the subsample and placed on a microscope slide with an 18x18mm cover slip

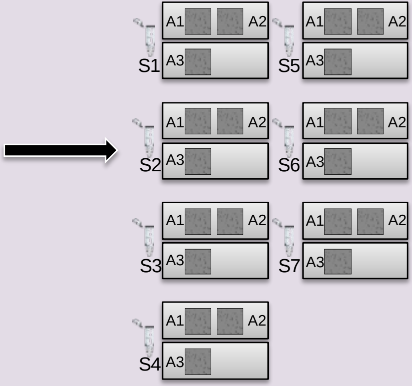



**Figure:** reprinted with permission from Ommen et al. [1]

# Automated Imaging

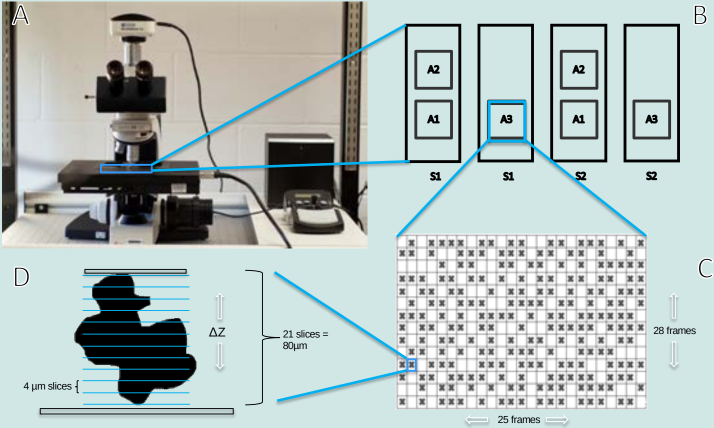



**Figure:** reprinted with permission from Ommen et al. [1]

## Particle Micromorphometry

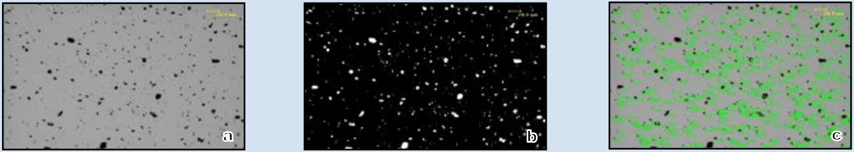



**Figure:** Each image (a) was converted to a binary image (b) to enhance edge detection. The particles were then counted (c), eliminating any particles along the border of the image, and measured. Seventeen parameters were measured for each identified particle within the image FOV: area; perimeter; feret diameter (minimum, maximum and mean); diameter (minimum, maximum, and mean); roundness; aspect ratio; box (height, width, and ratio); radii (minimum, maximum, and mean distance from particle centroid to edge); and fractal dimension. The data from thousands of particles were exported to a large text data file for further statistical analysis. (reprinted with permission from Ommen et al. [1])

Feature-based

Score-based

# Statistical Analysis

## Feature-based

Model high-dimensional raw features of evidence, $E = \{Q_1, Q_2, K\}$

$$\underbrace{\frac{P\left(H_p|E\right)}{P\left(H_d|E\right)}}_{\text{Posterior Odds}} = \underbrace{\frac{P\left(E|H_p\right)}{P\left(E|H_d\right)}}_{\text{Bayes Factor}} \times \underbrace{\frac{P\left(H_p\right)}{P\left(H_d\right)}}_{\text{Prior Odds}}$$

NOTES:

- Raw features of $E$ are $17 \times$ thousands - dimensional non-sparse matrices for each FOV
- Difficult to statistically model such large matrices
- Even if we could model them, computationally intractable to evaulate the likelihoods

## Score-based

# Statistical Analysis

## Feature-based

## Score-based

Model the low-dimensional output from a comparison function

$$\underbrace{\frac{f(\Delta(Q_1, Q_2)|H_p)}{f(\Delta(Q_1, Q_2)|H_d)}}_{\text{Score-based Likelihood Ratio}}$$

NOTES:

- $\Delta : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$ is a "black-box" comparison function where $\mathcal{X}$ is the raw feature-space
- $f$'s often need to best estimated using the $K$'s

Consists of four parts:

**Optional Dimension Reduction**

**Comparison Function**

**Contrastive Algorithm**

**SLR Computation**

# Contrastive Learning Framework

## Optional Dimension Reduction

If the data is too complex or high dimensional, simplify via

- Summary Statistics
- Principle Component Analysis (PCA)
- etc.

## Comparison Function

## Contrastive Algorithm

## Score-Based Likelihood Ratio

# Contrastive Learning Framework

## Optional Dimension Reduction

## Comparison Function

- Method for quantifying the (dis)similarity of pairs of evidential items
- Can incorporate adaptive scoring functions
- Examples:
  - Modified Wasserstein Distance Score (WDS) [2]
  - Modified ASTM-Glass Vector of Scores (VOS) [3,4]

## Contrastive Algorithm

## SLR Computation

# Contrastive Learning Framework

## Optional Dimension Reduction

## Comparison Function

## Contrastive Algorithm

- Method for determining the best separation between within-source or between-source comparisons
- Can incorporate statistical learning algorithms
- Examples:
  - Unsupervised Linear Discriminant Analysis (LDA)
  - Adaptive Random Forest Score (RFS) [5,6]

## SLR Computation

# Contrastive Learning Framework

## Optional Dimension Reduction

## Comparison Function

## Contrastive Algorithm

## SLR Computation

Outcome is an SLR for quantifying evidential value

- Traditional Kernel Density Estimation (KDE) [6]
- Bi-Normal Receiver Operating Characteristic (ROC) Curve [7]
- Pseudo-independence Ensemble System [8]

# Contrastive Learning Example

## Optional Dimension Reduction

First-level Summary Statistics

## Contrastive Algorithm

Unsupervised Linear Discriminant Analysis (LDA)

## Comparison Function

Modified Wasserstein Distance Score (WDS)

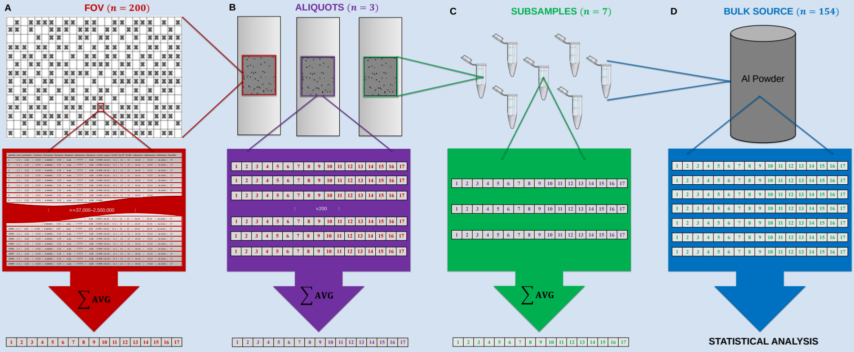
## SLR Computation

Bi-Normal Receiver Operating Characteristic (ROC) Curve

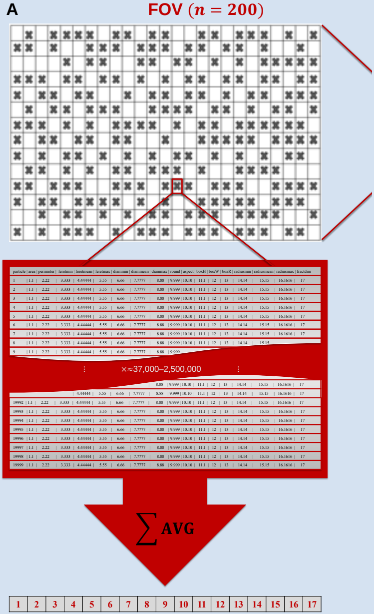
# Contrastive Learning Example

## Optional Dimension Reduction

- Several levels at which we can summarize data by taking an average:



# First Mean Summary = FOV Means



# Contrastive Learning Example

## Optional Dimension Reduction

Final Data:

- $n = 154$ Samples $\times$ 7 Subsamples
- Within-Source Subsample Pairs: $154 \times \binom{7}{2} = 3234$
- Between-Source Subsample Pairs: $\binom{154 \times 7}{2} - 3234 = 577,269$
- 200 FOV Means $\times 3$ Aliquots $= 600$ per Subsample

## Contrastive Algorithm

## Comparison Function

## SLR Computation

# Contrastive Learning Example

## Optional Dimension Reduction

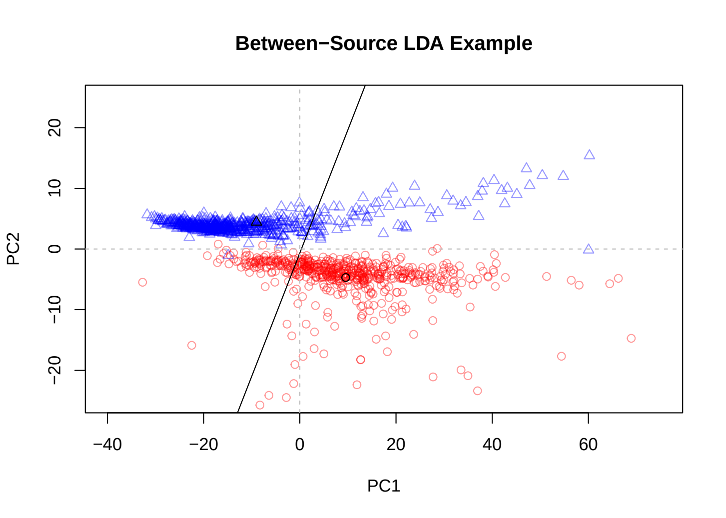
## Contrastive Algorithm

- For each pair of subsamples, perform unsupervised LDA
- Goal of LDA is binary classification (with equal priors)
- Output the vectors of posterior class probabilities
- The LDA method was implemented in R© using the `lda` and `predict` functions from the MASS package.

## Comparison Function

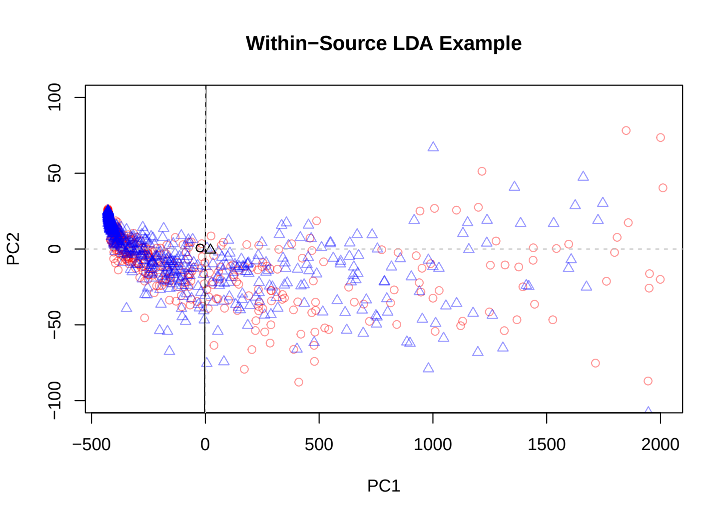
## SLR Computation

# Contrast: LDA




- Red circle class prediction: $\circ = 0.995$, $\triangle = 0.005$

# Contrast: LDA




- Red circle class prediction: $\circ = 0.494$, $\triangle = 0.493$

# Contrastive Learning Example

## Optional Dimension Reduction

## Contrastive Algorithm

## Comparison Function

- The modified Wasserstein Distance function is used to compare the distance between LDA output vectors.
- If the subsamples are from the same sample source, then the WDS should be small.
- If the subsamples are from different sample sources, then the WDS should be large.

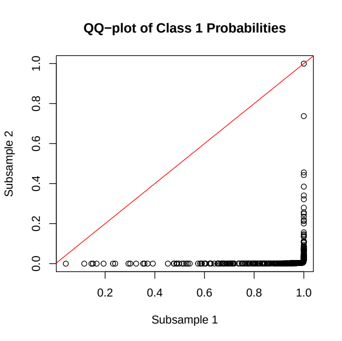
## SLR Computation

# Compare: WDS

Let's look at some quantiles for the red circle class predictions

$$WDS = \frac{1}{P}\sum_{p=1}^{P}(\circ_p - \triangle_p)^2$$

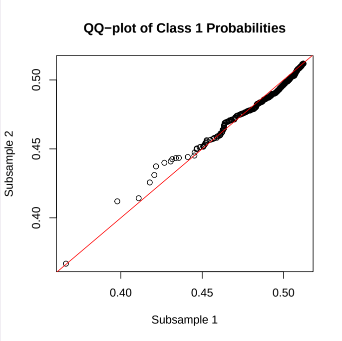
## Between-Source Example




PCA-WDS = 0.725

Full-WDS = 1

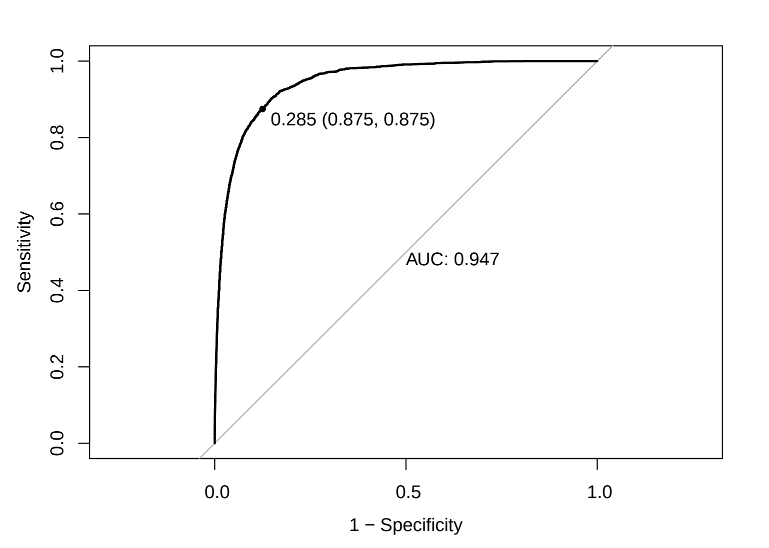
## Within-Source Example




PCA-WDS = $7 \times 10^{-6}$

Full-WDS = $7 \times 10^{-5}$

# Compare: WDS



0.285 (0.875, 0.875)
AUC: 0.947
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.5
1.0
1 − Specificity

# Contrastive Learning Example

## Optional Dimension Reduction

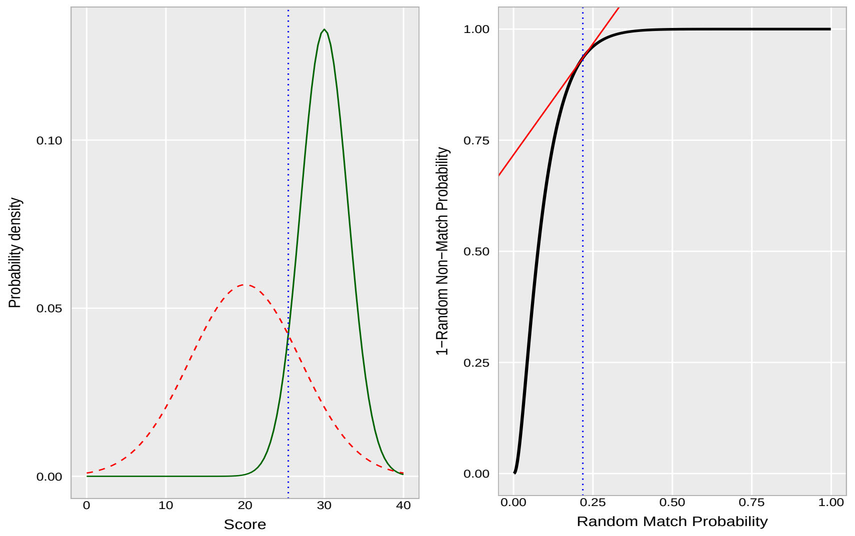
## Contrastive Algorithm

## Comparison Function

## SLR Computation

- The resulting WDS can either be a within-source (“match") or between-source comparison.
- The ROC curve for this binary classification plots the random match probability against one minus the random nonmatch probability.
- The derivative of the ROC curve is equivalent to the SLR.

# SLR: Bi-Normal ROC



## SLR: Bi-Normal ROC

- $1 - RNMP = P(\Delta(Q_1, Q_2) \leq \tau | H_p) = F_p(\tau)$ with density $f_p$
- $RMP = P(\Delta(Q_1, Q_2) \leq \tau | H_d) = F_d(\tau)$ with density $f_d$
- Statistically, the ROC is defined for $0 < x < 1$

$$R(x) = F_p(F_d^{-1}(x))$$

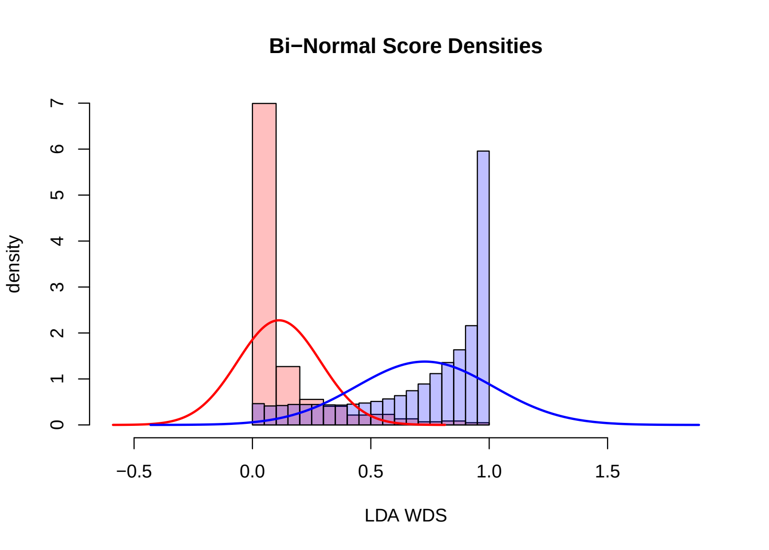
- For $F_d^{-1}(x) = \Delta(Q_1, Q_2)$, the derivative of the ROC curve is

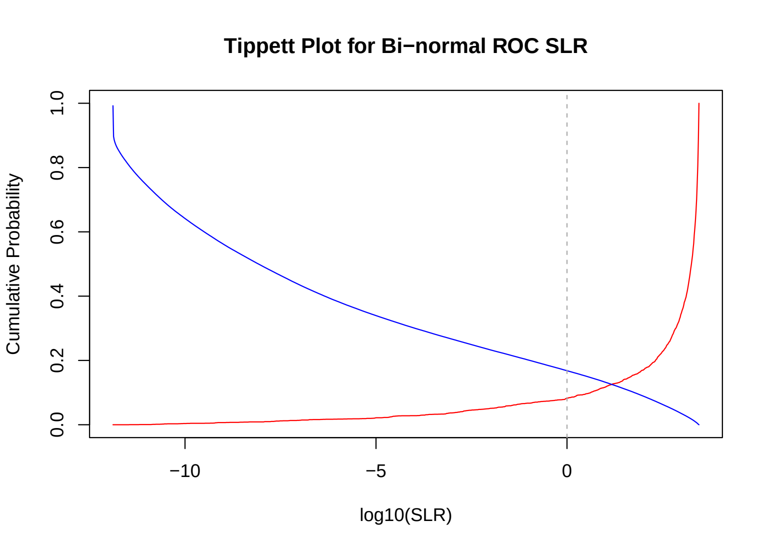
$$\frac{dR(x)}{dx} = \frac{f_p(F_d^{-1}(x))}{f_d(F_d^{-1}(x))} = \frac{f(\Delta(Q_1, Q_2)|H_p)}{f(\Delta(Q_1, Q_2)|H_d)} = SLR!$$

- To make things simple, suppose the scores are transformed using a function $T$ so $f_p \sim N(\mu_p, \sigma_p)$ and $f_d \sim N(\mu_d, \sigma_d)$

$$SLR = \frac{\hat{\sigma}_d \ \phi( \ [\hat{\mu}_p - T(\Delta(Q_1, Q_2))]/\hat{\sigma}_p \ )}{\hat{\sigma}_p \ \phi( \ [\hat{\mu}_d - T(\Delta(Q_1, Q_2))]/\hat{\sigma}_d \ )}$$

# SLR: Bi-Normal ROC




Tippett Plot for Bi−normal ROC SLR
Cumulative Probability
0.0
0.2
0.4
0.6
0.8
1.0
−10
−5
0
log10(SLR)

# Results

- log-SLRs that support the prosecution hypothesis
  - When W is ground truth, these are “correct."
  - When B is ground truth, these are misleading!

| | (0,2] | (2,4] |
|---|---|---|
| W | 0.088 | 0.82 |
| B | 0.079 | 0.089 |

- log-SLRs that support the defense hypothesis
  - When B is ground truth, these are “correct."
  - When W is ground truth, these are misleading!

| | (-12,-10] | (-10,-8] | (-8,-6] | (-6,-4] | (-4,-2] | (-2,0] |
|---|---|---|---|---|---|---|
| B | 0.358 | 0.146 | 0.112 | 0.083 | 0.068 | 0.065 |
| W | 0.004 | 0.005 | 0.009 | 0.011 | 0.023 | 0.031 |

# Conclusion

- The contrastive learning framework has the flexibility to handle quantifying evidential value for a variety of complex evidence forms.
- In the common source Al powder example, the contrastive learning framework resulted in:
  (FOV Mean Summary -> LDA -> WDS -> Bi-normal ROC SLR)
  - TPR = 91.7%  TNR = 83.2%
  - FPR = 16.8%  FNR = 8.3%
- Each module of the contrastive learning framework can be substituted for any other applicable method (example in appendix).
- The compare and contrast modules of the contrastive learning framework can switch order (example in appendix).

# Future Work

Lots of room for improvement ...

- Consider other definitions of source or sub-source
- Extend framework to address specific source problem
- Pairwise comparison approach creates complex score-dependence structure that hasn't been handled appropriately.
- Bi-normal assumption for the Al powder data is clearly inappropriate.

# Thanks for listening!

**Questions**

Email me: dmommen@iastate.edu

# References


1. Ommen DM, Baldaino J, Saunders CP, Hietpas J, Buscaglia J. https://doi.org/10.1111/1556-4029.14946
2. Ommen DM, Fuglsby C, Caligiuri M. http://dx.doi.org/10.1016/j.forsciint.2020.110644
3. ASTM International, E2927-16e1, 2022
4. ASTM International, E2330-19, 2022
5. Park S, Carriquiry A. https://doi.org/10.1214/18-AOAS1211
6. Johnson M, Ommen DM. https://doi.org/10.1002/sam.11566
7. Zhu X, Tang L, Tabassi E. https://ieeexplore.ieee.org/document/8272737
8. Ommen DM, Veneri F. https://forensicstats.org/blog/portfolio/ensemble-slrs-for-forensic-evidence-comparison/

# Contrastive Learning Framework

| | Example 1 | Example 2 |
|---|---|---|
| Dimension Reduction | FOV Means | Aliquot Means |
| Method Step 1 | Contrast: LDA | Compare: ASTM VOS |
| Method Step 2 | Compare: WDS | Contrast: RFS |
| SLR Computation | Bi-Normal ROC | Traditional KDE |

# Contrastive Learning Example 2

## Optional Dimension Reduction

Second-level Summary Statistics

## Comparison Function

ASTM Vector of Scores (VOS)

## Contrastive Algorithm

Random Forest Score (RFS)

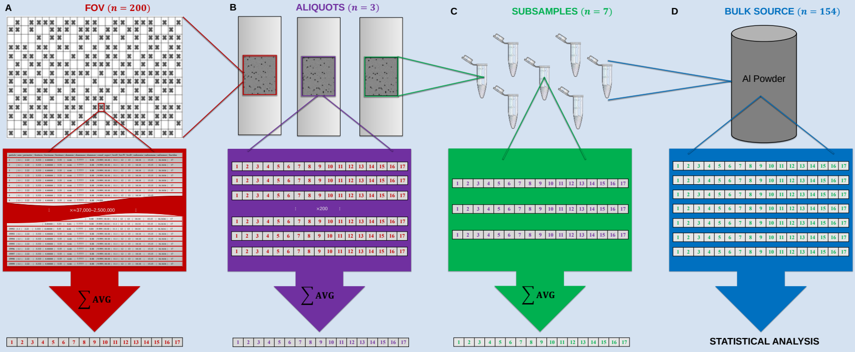
## SLR Computation

Traditional Kernel Density Estimation (KDE)

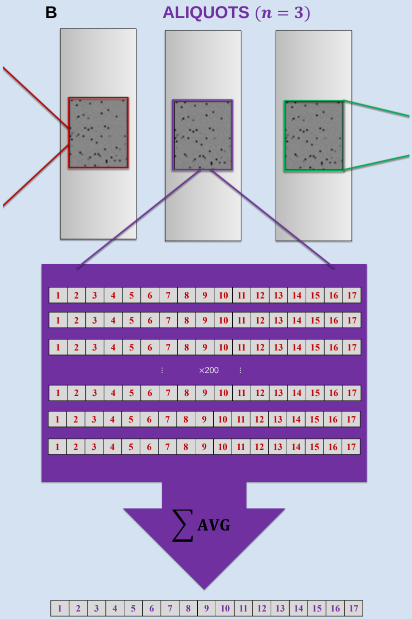
# Contrastive Learning Example 2

## Optional Dimension Reduction

- Several levels at which we can summarize data by taking an average:



# Second Mean Summary = MoM



# Contrastive Learning Example 2

## Optional Dimension Reduction

Final Data:

- $n = 154$ Samples $\times$ 7 Subsamples
- Within-Source Subsample Pairs: $154 \times \binom{7}{2} = 3234$
- Between-Source Subsample Pairs: $\binom{154 \times 7}{2} - 3234 = 577,269$
- 3 Aliquot Means per Subsample

## Comparison Function

## Contrastive Algorithm

## SLR Computation

# Contrastive Learning Example 2

## Optional Dimension Reduction

## Comparison Function

- The ASTM Method was designed to infer whether glass fragments in a query sample are indistinguishable from glass fragments in the known sample according to their trace elemental compositions as measured by ICP-MS [3,4].
- Goal: modify the method to determine whether two questioned subsamples of Al powder share a source according to the 17 micromorphometric parameters.

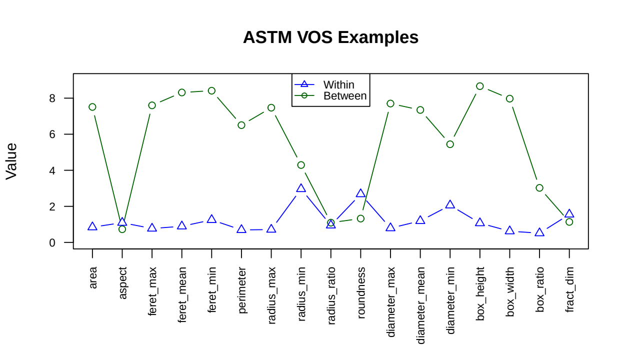
## Contrastive Algorithm

## SLR Computation

# Compare: ASTM VOS

- Let $Q_{ij}$ denote the vector of aliquot means for the $j^{th}$ aliquot of the $i^{th}$ subsample in the pair.
- Compute

$$\Delta(Q_1, Q_2) = \frac{\left|\bar{Q}_1 - \bar{Q}_2\right|}{\sqrt{\frac{1}{2}\sum_{j=1}^{3}\left(Q_{2j} - \bar{Q}_2\right)^2}}$$



# Contrastive Learning Example 2

## Optional Dimension Reduction

## Comparison Function

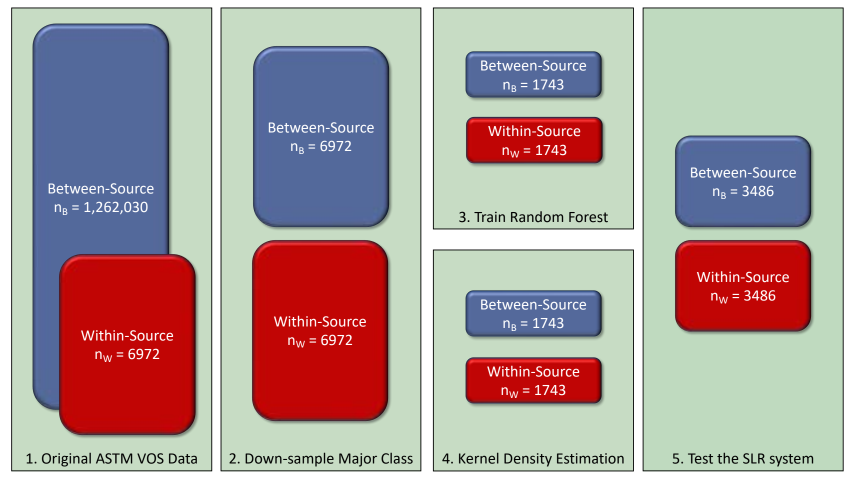
## Contrastive Algorithm

- A random forest algorithm is used to compare the similarity of ASTM VOS via binary classification probabilities.
- If the subsamples are from the same sample source, then the RFS should be large.
- If the subsamples are from different sample sources, then the RFS should be small.

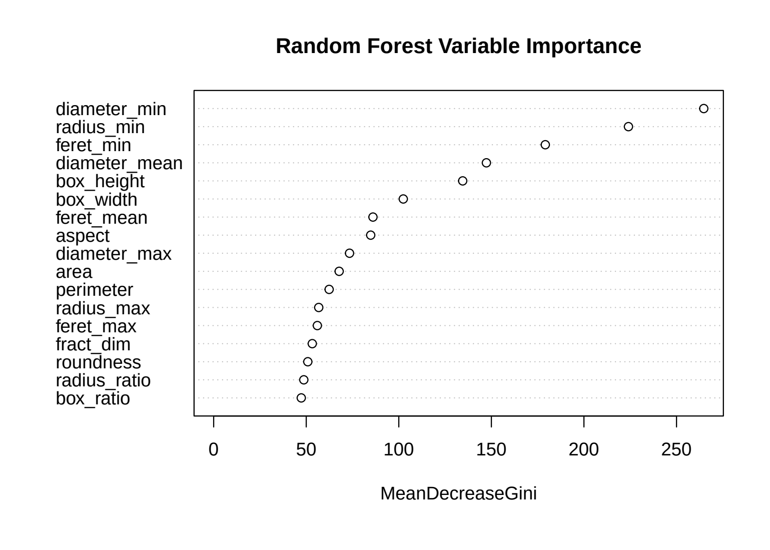
## SLR Computation

# Contrast: RFS

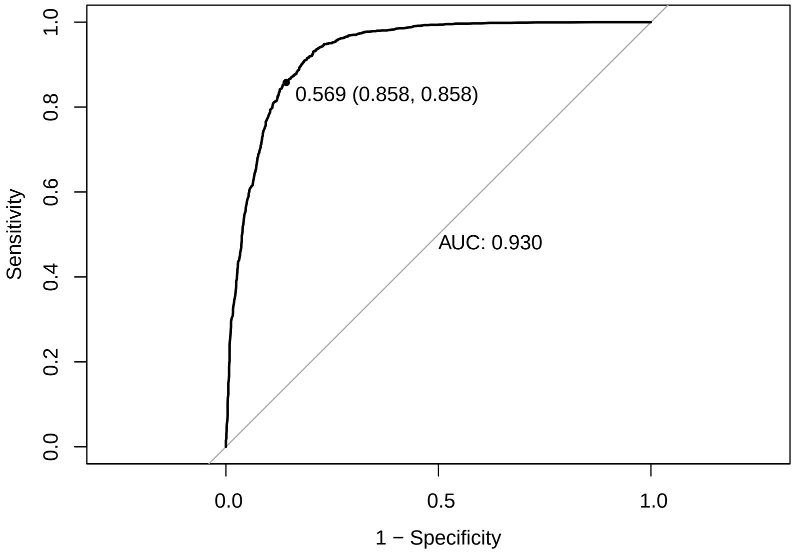
- The RF algorithm serves as an *adaptive* scoring function, which means that it needs data to train …



# Contrast: RFS



# Contrast: RFS



# Contrastive Learning Example 2

## Optional Dimension Reduction

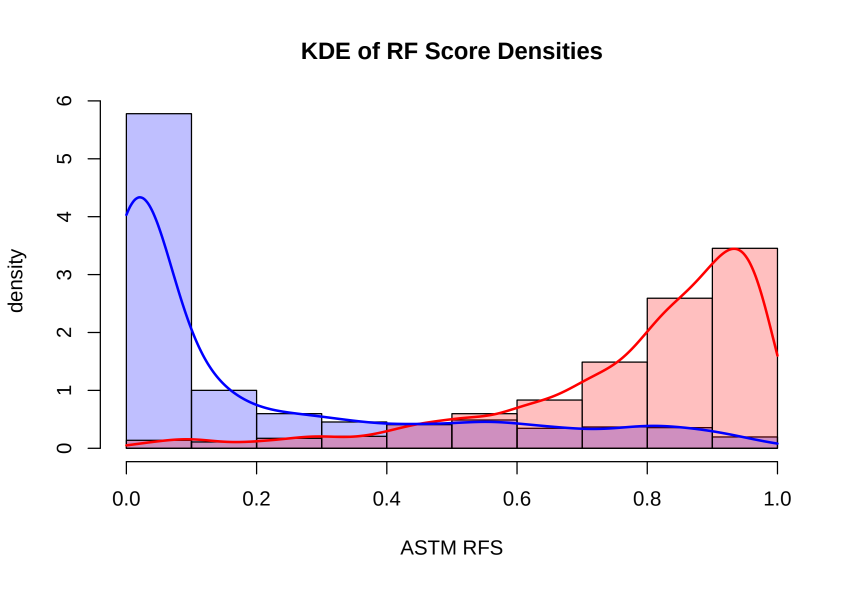
## Comparison Function

## Contrastive Algorithm

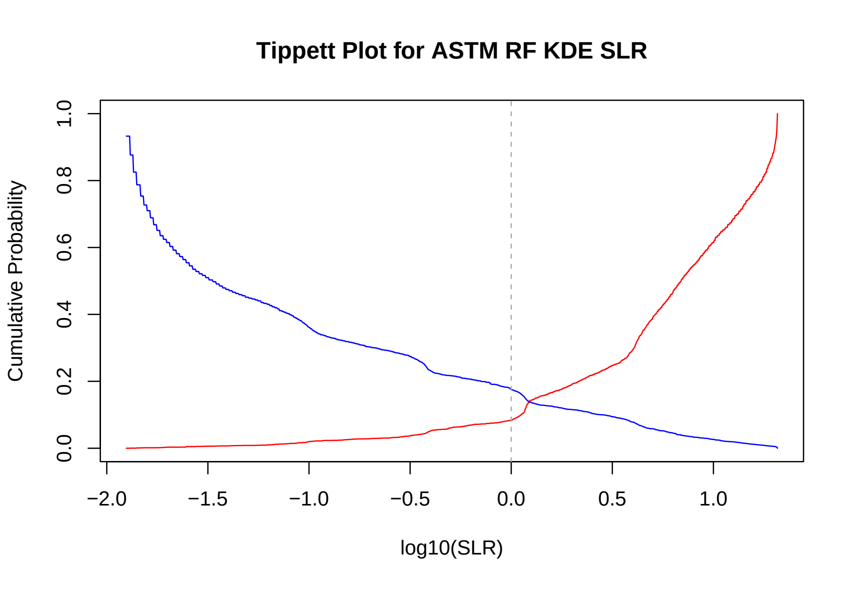
## SLR Computation

- The resulting RFS is either a within-source or between-source comparison.
- KDE of known within-source RFS used to estimate $f_p$
- KDE of known between-source RFS used to estimate $f_d$
- The SLR is the ratio of $\Delta(Q_1, Q_2)$ evaluated at $f_p$ over $f_d$.

# SLR: Traditional KDE



# Results



Tippett Plot for ASTM RF KDE SLR
Cumulative Probability
0.0
0.2
0.4
0.6
0.8
1.0
−2.0
−1.5
−1.0
−0.5
0.0
0.5
1.0
log10(SLR)

# Results

- log-SLRs that support the prosecution hypothesis
  - When W is ground truth, these are “correct" (TPR=91.6%).
  - When B is ground truth, these are misleading (FPR=17.6%)!
- log-SLRs that support the defense hypothesis
  - When B is ground truth, these are “correct" (TNR=82.4%).
  - When W is ground truth, these are misleading (FNR=8.4%)!

| | (-2,-1] | (-1, 0] | (0,1] | (1,2] |
|---|---|---|---|---|
| B | 0.638 | 0.186 | 0.150 | 0.026 |
| W | 0.020 | 0.064 | 0.531 | 0.385 |

**INTERNATIONAL CONFERENCE ON FORENSIC INFERENCE AND STATISTICS (ICFIS)**
**10-15 JUNE 2023**
**LUND, SWEDEN**

# Development and Evaluation of a Contrastive Learning Framework with Applications to Micromorphometry of Aluminum Powder used in Explosives

Authors: Danica Ommen (Iowa State University)*, Christopher Saunders (South Dakota State University), JoAnn Buscaglia (Federal Bureau of Investigation Laboratory Division)

Abstract:
The identification of source framework can be used within forensic evidence interpretation to compare a pair of items and determine whether they have come from a common unknown source or from two different unknown sources. In comparing aluminum (Al) powder particles recovered from two pre-blast improvised explosive devices (IEDs), the goal is to determine whether the powder sources are associated, potentially providing investigative between-case linkages. These problems can be addressed using a variety of statistical techniques, including the Two-Stage, Likelihood Ratio and Bayes Factor approaches. Unfortunately, the complex nature of the evidence, such as replicate measurements taken on different levels of substructure within the source powder, creates difficulties in applying the usual approaches in a straightforward manner. For characterizing features of the Al powder, we take several subsamples from the bulk Al powder, several aliquots from each subsample, several fields of view are imaged on each aliquot, and then multiple micromorphometric parameters are measured for each particle in a field of view. The hierarchical nature of this type of data creates a complex dependency structure that is difficult to directly incorporate into the traditional statistical methods for source identification. In this presentation, a contrastive learning algorithm framework is developed for complex evidence types like Al powder micromorphometry. The contrastive learning methods consist of two major components: a method for quantifying the similarity (or dissimilarity) of pairs of evidential items and a method for determining the best separation between within-source or between-source comparisons. During this presentation, we explore several different methods of quantifying pairwise similarity and several different methods of classifying pairs as within- or between-source comparisons. We will also present our approaches for evaluating the performance of the resulting score functions using micromorphometric data from Al powders.

Keywords: scores, machine learning, trace evidence, common source