

# A Cartridge Case Comparison Pipeline

*Oral Preliminary Exam*

**Joseph Zemmels**

**Iowa State University: Center for Statistics and Applications in Forensic Evidence**

**2022-06-17**



# **ACKNOWLEDGEMENTS**

## **Funding statement**

This work was partially funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) through Cooperative Agreement 70NANB20H019 between NIST and Iowa State University, which includes activities carried out at Carnegie Mellon University, Duke University, University of California Irvine, University of Virginia, West Virginia University, University of Pennsylvania, Swarthmore College and University of Nebraska, Lincoln.

# About Me

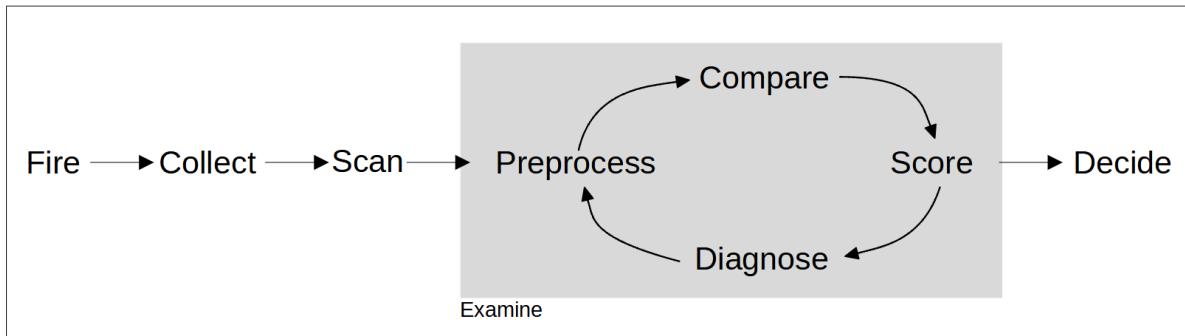
- Undergraduate at Winona State University in Winona, MN, 2013-2017
  - B.Sc. in Honors Mathematics and Mathematics: Secondary Education
- Graduate at Iowa State University, 2018-present
  - M.Sc. in Statistics, 2020
  - Creative Component: *Open-source implementation of the Congruent Matching Cells method for cartridge case identification*
- I hope to teach Statistics and Data Science at a post-secondary institution in the Denver area.

# OUTLINE

- **Introduction** Firearms and Toolmarks Identification
- **Paper #1** A Study in Reproducibility: The Congruent Matching Cells Algorithm and cmcR package
- **Paper #2** Diagnostic Tools for Cartridge Case Comparison Algorithms
- **Paper #3** Automatic Matching of Cartridge Case Impressions
- **Conclusions**

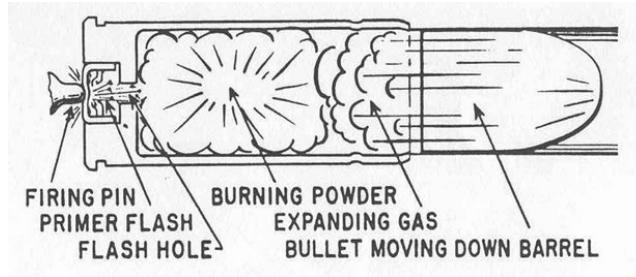
# INTRODUCTION

## Firearms and Toolmarks Identification



# The Firing Process

As the bullet travels down the barrel, the cartridge case travels backwards.



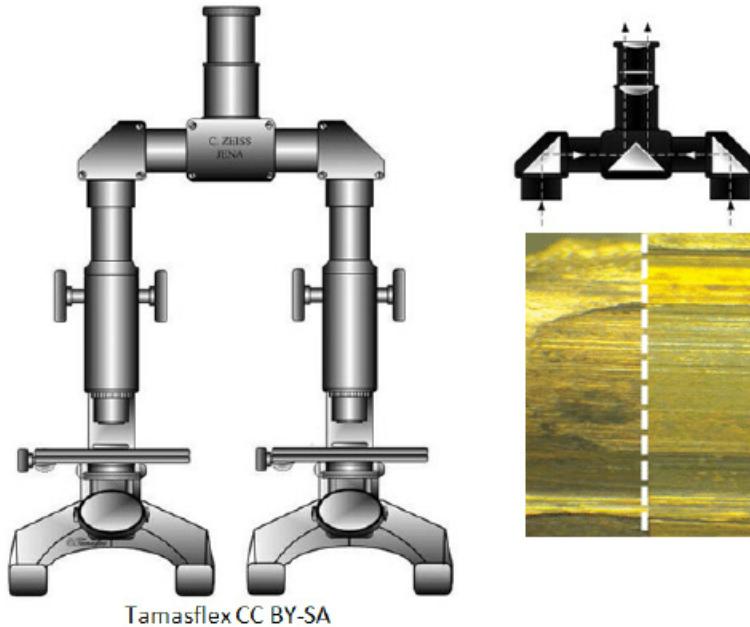
Contact with the back wall of the barrel leaves *breech face impressions*.



Similarity of breech face impressions on two cartridge cases suggests they were fired from the same gun [Tho17].

# Firearms and Toolmarks (F & T) Examinations

A comparison microscope combines two compound microscopes into a single view [Zhe+14]



Each stage is independently translated or rotated to find optimal matching position.

# Firearms and Toolmarks (F & T) Examinations

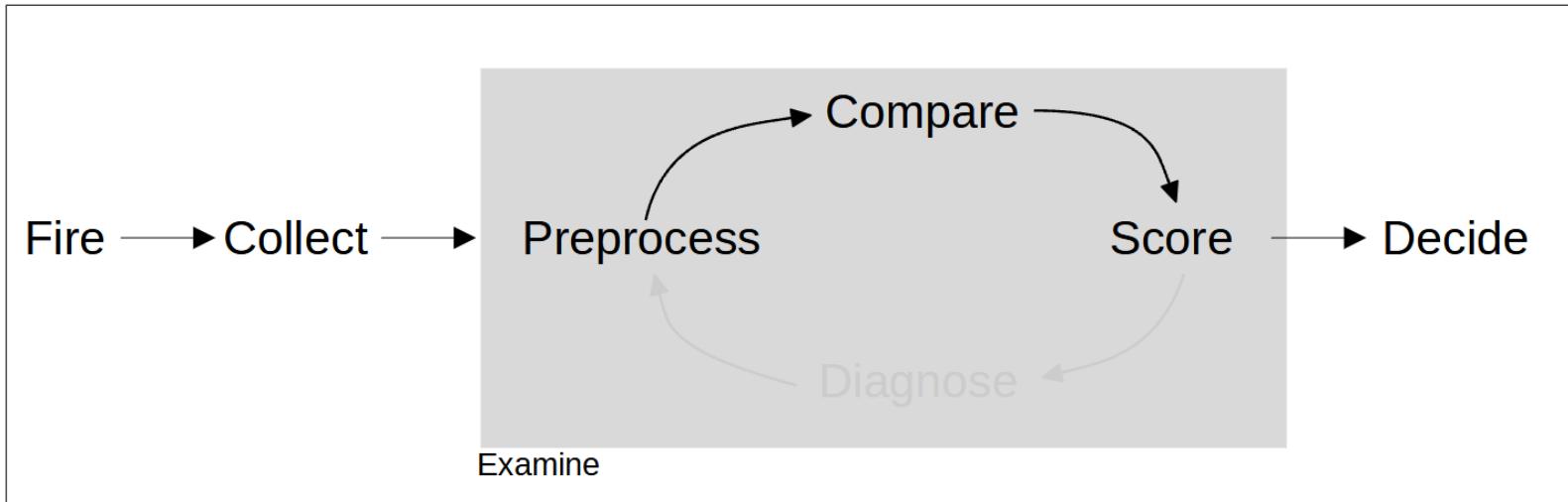
- **Class characteristics:** features associated with manufacturer of the firearm.
  - E.g., size of ammunition, width and twist direction of barrel rifling
  - Used to narrow the relevant population of potential firearms
- **Individual characteristics:** markings attributed to imperfections in the firearm surface.
- **Subclass characteristics:** markings that are reproduced across a sub-group of firearms.
  - E.g., barrels milled by the same machine may share similar markings
  - Difficult to distinguish individual from subclass characteristics
- "Sufficient" agreement of class and individual characteristics suggests that the evidence originated from the same firearm [AFT92].

# AFTE Range of Conclusions

1. **Identification:** Agreement of a combination of individual characteristics and all discernible class characteristics where the extent of agreement exceeds that which can occur in the comparison of toolmarks made by different tools and is consistent with the agreement demonstrated by toolmarks known to have been produced by the same tool.
2. **Inconclusive:**
  - 2.1 Some agreement of individual characteristics and all discernible class characteristics, but insufficient for an identification.
  - 2.2 Agreement of all discernible class characteristics without agreement or disagreement of individual characteristics due to an absence, insufficiency, or lack of reproducibility.
  - 2.3 Agreement of all discernible class characteristics and disagreement of individual characteristics, but insufficient for an elimination.
3. **Elimination:** Significant disagreement of discernible class characteristics and/or individual characteristics.
4. **Unsuitable:** Unsuitable for examination.

# F & T Comparison Pipeline

- F & T examinations can be thought of as an evidence-to-decision "pipeline."



- Some of these steps are performed implicitly by the examiner. For example:
  - "Preprocessing" includes adjusting lighting on the comparison stage.
  - The examiner determines a similarity "score" to inform their decision.
- This pipeline structure is also useful when considering automatic comparison algorithms.

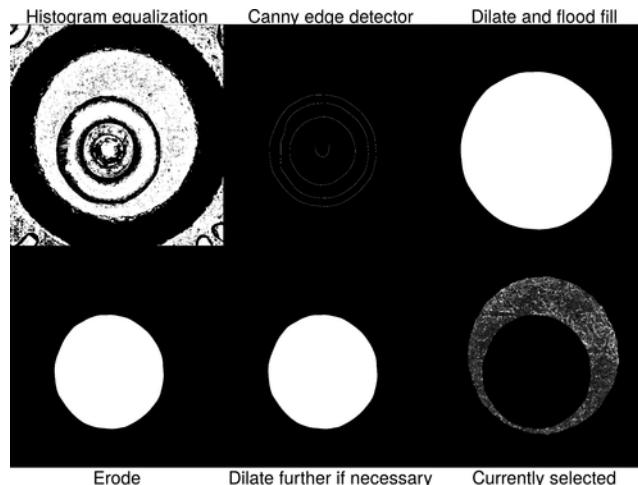
# Why Change F & T Examinations?

- "Sufficient studies have not been done to understand the reliability and repeatability of the methods."<sup>[1]</sup>
- "A fundamental problem with toolmark and firearms analysis is the lack of a precisely defined process."<sup>[1]</sup>
- "Automated techniques capable of enhancing forensic technologies" should be developed.<sup>[1]</sup>
- "Firearms analysis currently falls short of the criteria for foundational validity."  
[2]
- Efforts should be made to "convert firearms analysis from a subjective method to an objective method ... [which] would involve developing and testing image-analysis algorithms for comparing the similarity of toolmarks..."<sup>[2]</sup>

[1] National Research Council [Nat09], [2] President's Council of Advisors on Sci. & Tech. [Pre16]

# F & T Comparison Algorithms

- Algorithms are also evidence-to-decision (or score) pipelines.
- Many algorithms exist to compare F & T evidence [Son13; RC14; Rot+15; HHC17; KH18; TE18; Riv+20]
- Pipelines can be decomposed into modular "pieces."

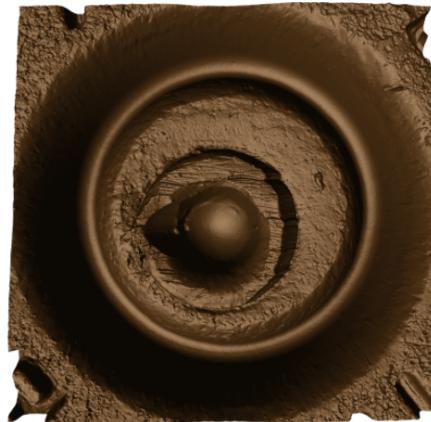
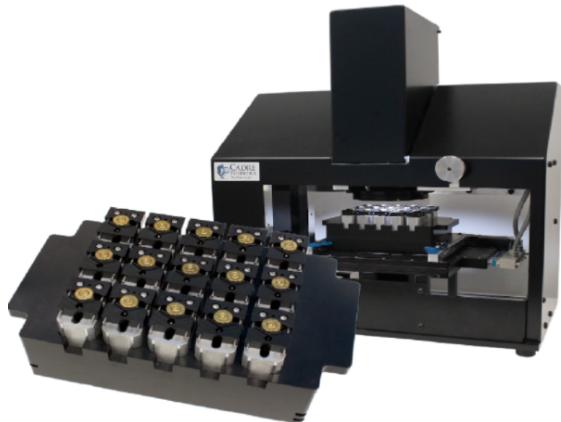


- Primarily focused on the **Congruent Matching Cells** algorithm introduced in Song [Son13].

Image source: Tai, et al. [TE18]

# Data Format

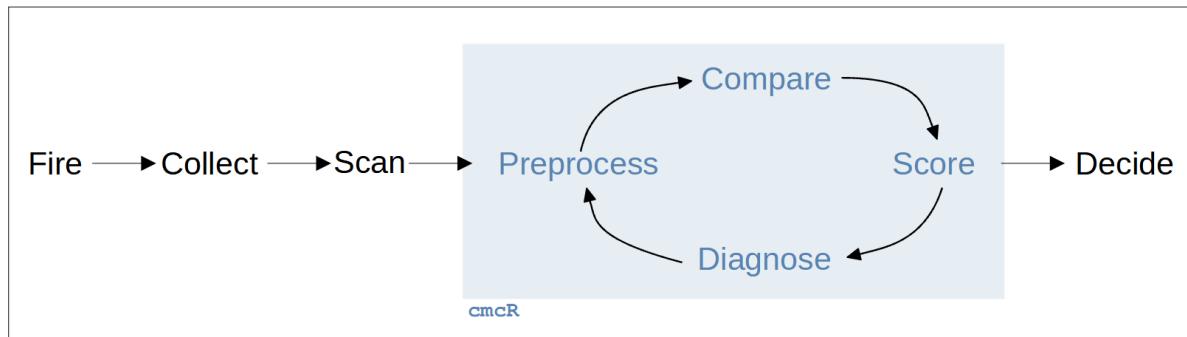
- Cartridge cases are scanned using the TopMatch-3D <sup>TM</sup> scanner from Cadre Forensics
- Scans are stored in an `.x3p` file containing:
  - a 2D array of surface values called a *surface matrix*
  - metadata about scan (size, resolution, creator, etc.)



- We consider cartridge case scans from the "Ames I" study [Bal+14].

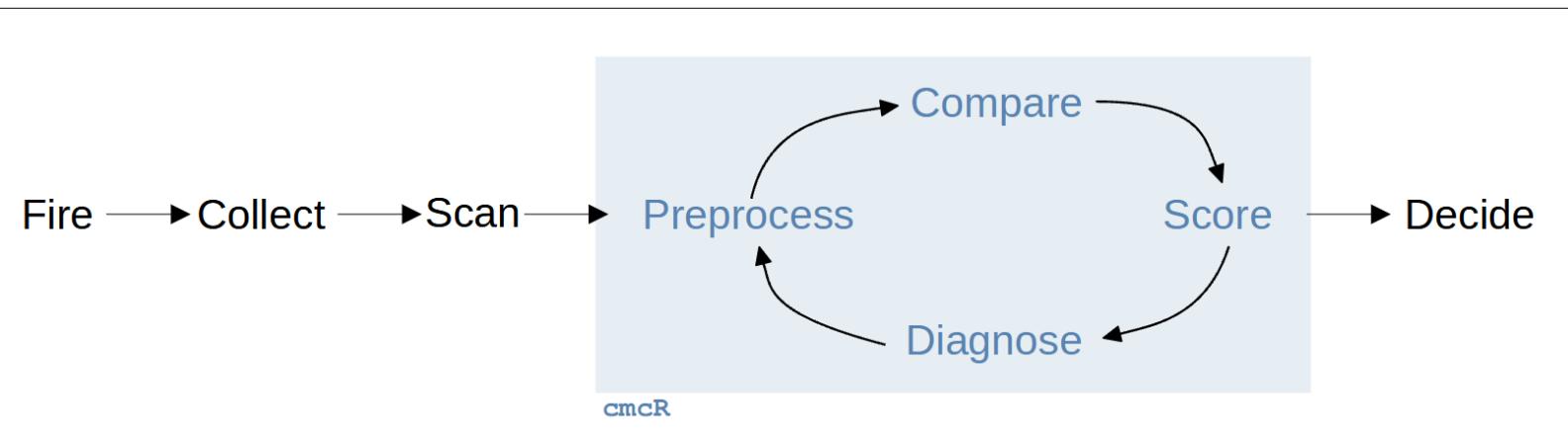
# A Study in Reproducibility: The Congruent Matching Cells Algorithm and cmcR package

*Submitted to The R Journal*



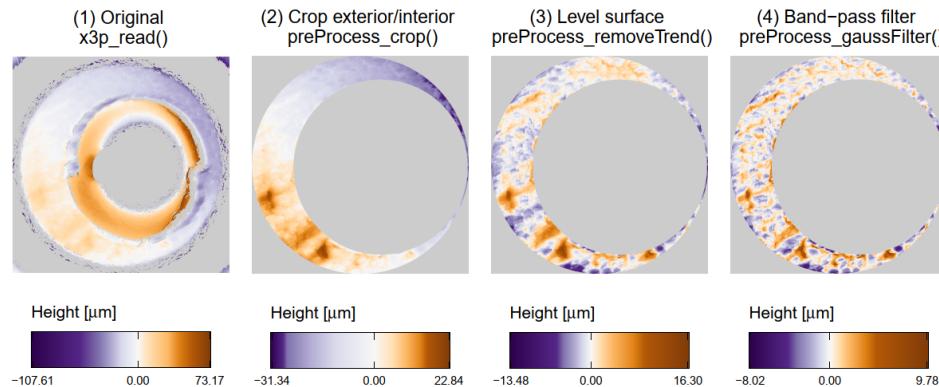
# Motivation

- Implement the **Congruent Matching Cells** (CMC) algorithm as described in Song [Son13]; Song, Chu, Tong, and Soons [Son+14]; Tong, Song, and Chu [TSC15].
- Structure implementation as a modularized pipeline.
- Each step of the algorithm is broken into exchangeable "modules."
- The open-source `cmcR` package provides a foundation upon which further improvements can be added.

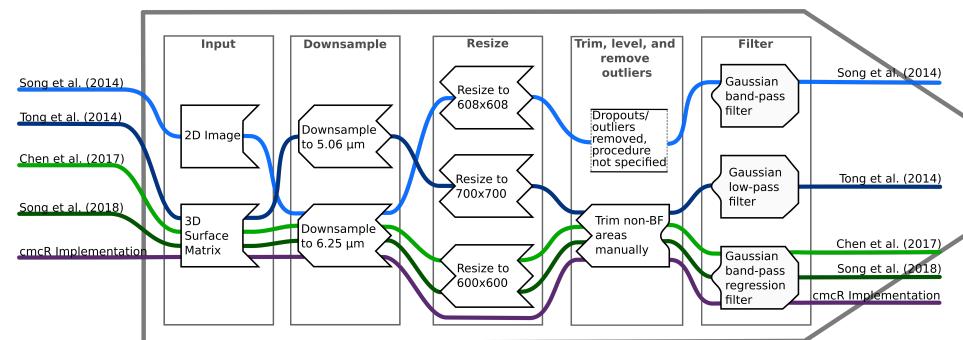


# The CMC Pipeline: Preprocess

**Goal:** Isolate breech face (BF) impressions on cartridge case scan.



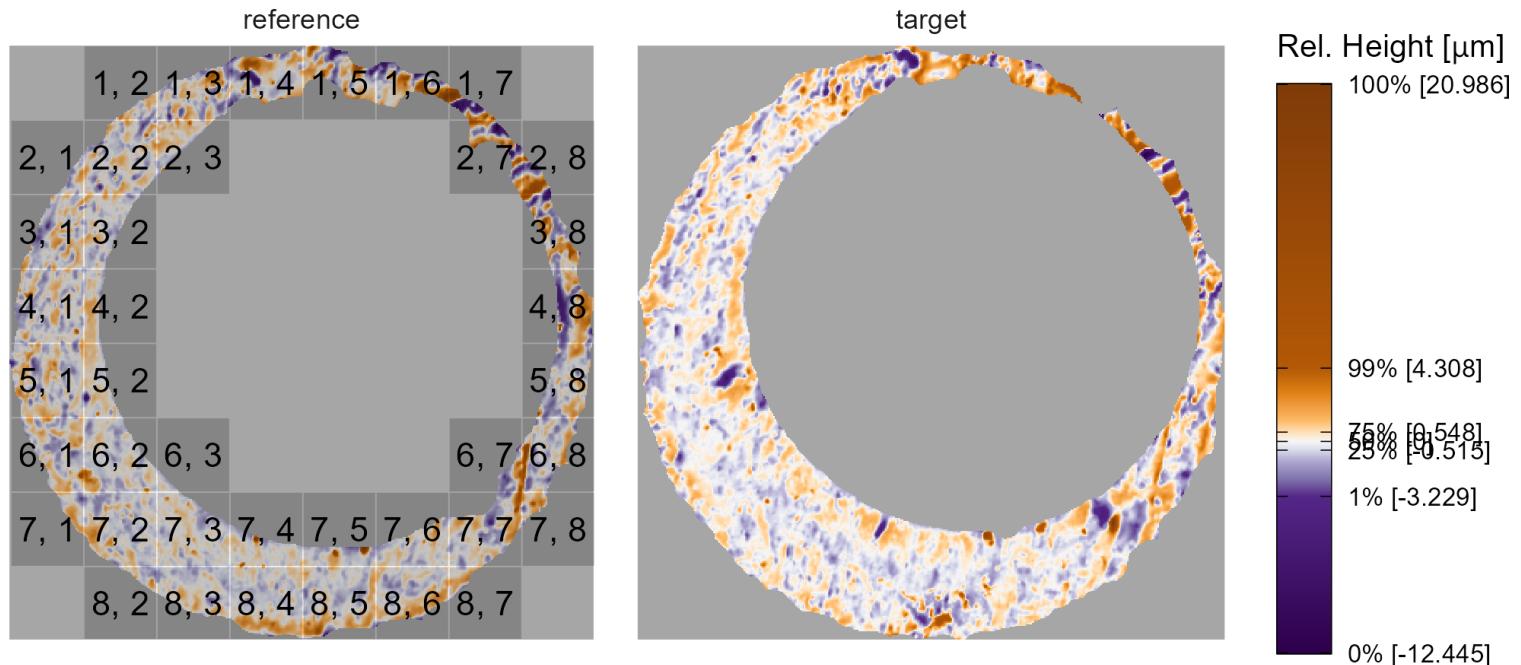
Each preprocessing step replaces non-BF "noise" with NAs.



# The CMC Pipeline: Compare

**Goal:** Extract similarity features from two cartridge case scans

- 1) Divide the *reference* scan into a grid of cells; e.g.  $8 \times 8$
  - 2) Consider the  $N$  cells that contain observed (non-NA) values; e.g.,  $N = 39$  cells



# The CMC Pipeline: Compare

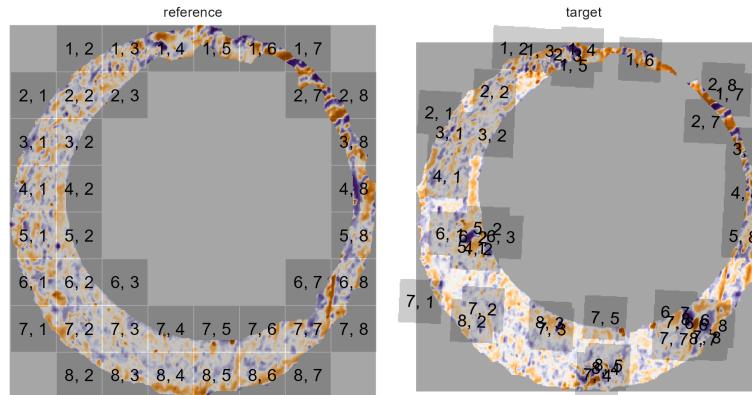
3) For a rotation  $\theta \in \Theta$ ; e.g.,  $\Theta = \{-30^\circ, -27^\circ, \dots, 27^\circ, 30^\circ\}$ :

3.1) Rotate the *target* scan by  $\theta$ .

3.2) For each reference cell  $i = 1, \dots, N$ :

3.2.1) Calculate the translation,  $[\Delta x_{\theta,i}, \Delta y_{\theta,i}]$ , at which cell maximizes the *cross-correlation function* (CCF) in the target scan.<sup>[1]</sup>

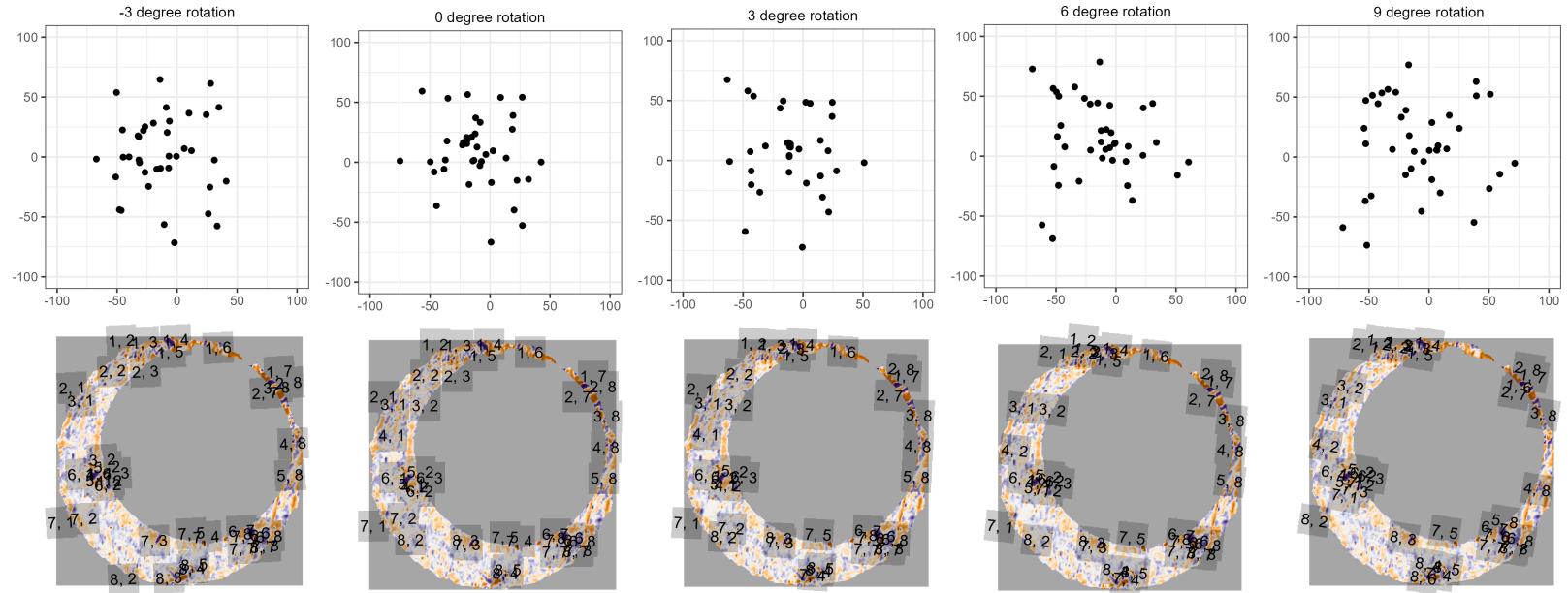
3.2.2 Return  $(\Delta x_{\theta,i}, \Delta y_{\theta,i}, CCF_{\theta,i})$



[1] See appendix for more details on CCF

# The CMC Pipeline: Compare

4) Return  $\{(\theta, \Delta x_{\theta,i}, \Delta y_{\theta,i}, CCF_{\theta,i}) : \theta \in \Theta, i = 1, \dots, N\}$

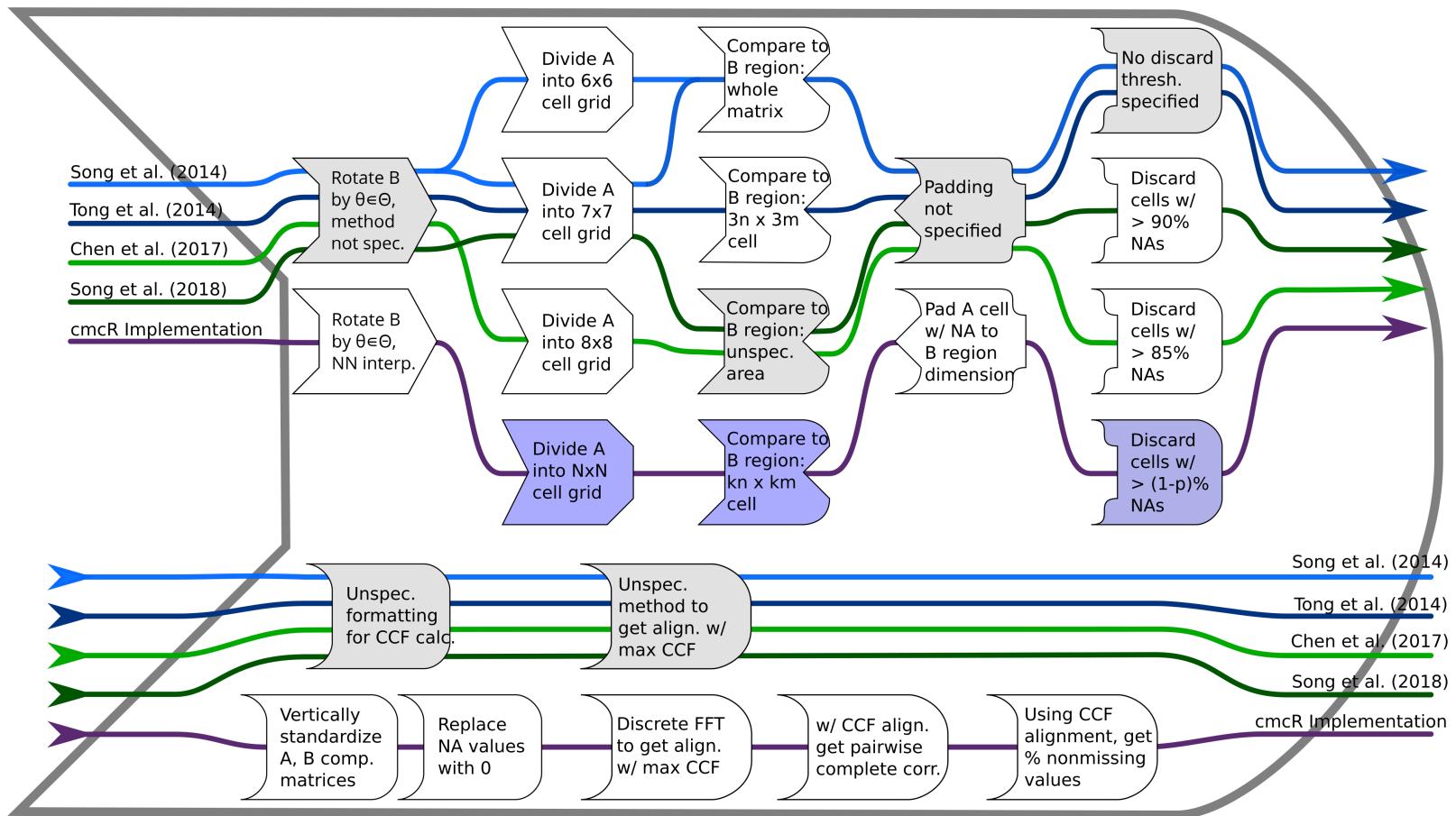


For each rotation, each cell "votes" for where it aligns best in the target scan.

For truly matching cartridge cases, the cells should "agree" on a translation at the true rotation.

In this example,  $(\theta, \Delta x, \Delta y) \approx (3, -10, 10)$  appears to be the "consensus."

# The CMC Pipeline: Compare



# The CMC Pipeline: Score

**Goal:** Calculate a similarity score based on features from previous step.

1) For each reference cell  $i = 1, \dots, N$ :

1.1) Calculate the rotation that maximizes the CCF:

$$\hat{\theta}_i = \arg \max_{\theta \in \Theta} \{CCF_{\theta,i} : \theta \in \Theta\}$$

1.2) Return this rotation with associated CCF and translation; call them  $(\hat{\theta}_i, \widehat{\Delta x}_i, \widehat{\Delta y}_i, CCF_i)$

2) Estimate the consensus rotation and translation as

$$\hat{\theta} = \text{median}(\{\hat{\theta}_i : i = 1, \dots, N\})$$

$$\widehat{\Delta x} = \text{median}(\{\Delta x_i : i = 1, \dots, N\})$$

$$\widehat{\Delta y} = \text{median}(\{\Delta y_i : i = 1, \dots, N\})$$

The consensus values are based on each cell's "top" vote.

# The CMC Pipeline: Score

3) For user-defined thresholds  $T_\theta, T_{\Delta x}, T_{\Delta y}, T_{CCF}$ , we call cell  $i$  a *Congruent Matching Cell* (CMC) if the following hold:

$$|\hat{\theta}_i - \hat{\theta}| \leq T_\theta$$

$$|\widehat{\Delta x}_i - \widehat{\Delta x}| \leq T_{\Delta x}$$

$$|\widehat{\Delta y}_i - \widehat{\Delta y}| \leq T_{\Delta y}$$

$$CCF_i \geq T_{CCF}$$

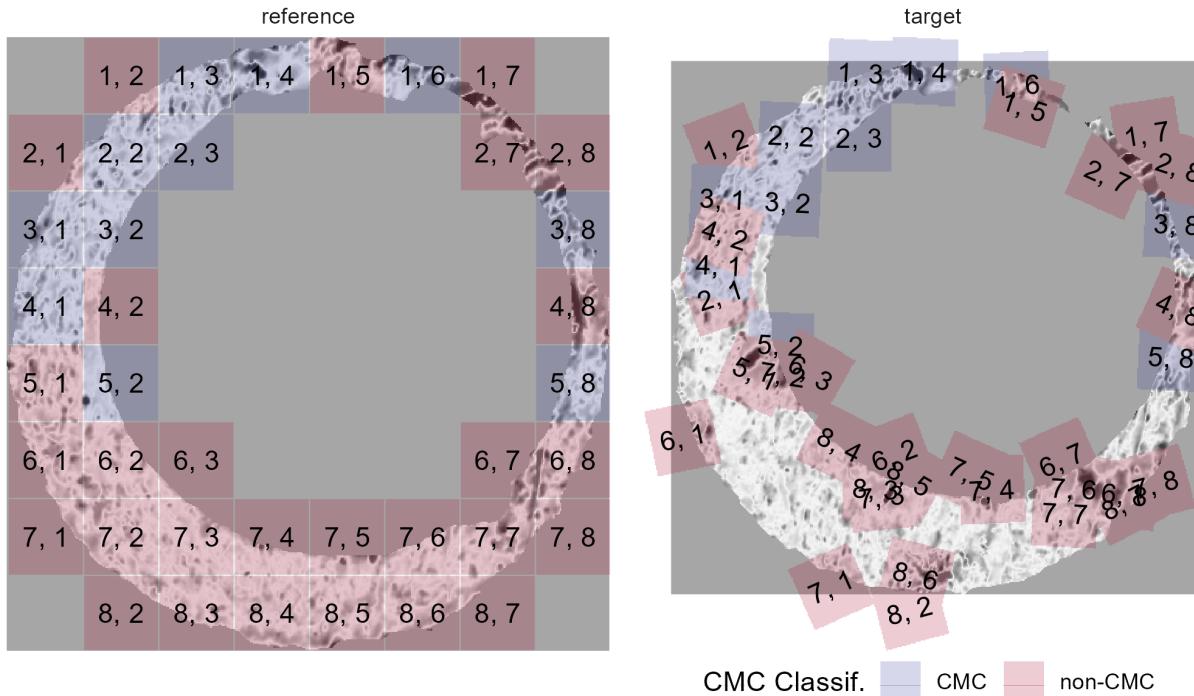
Otherwise, it is a *non-CMC*.

Cells are classified as CMCs if the estimated translations and rotation are close to the consensus **and** the associated CCF is large.

# The CMC Pipeline: Score

In this example, 11 cells are classified as CMCs and 28 as non-CMCs.

The number of CMCs is used as a similarity score.

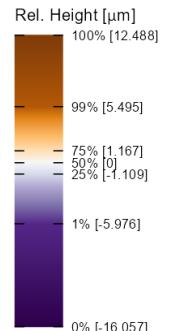
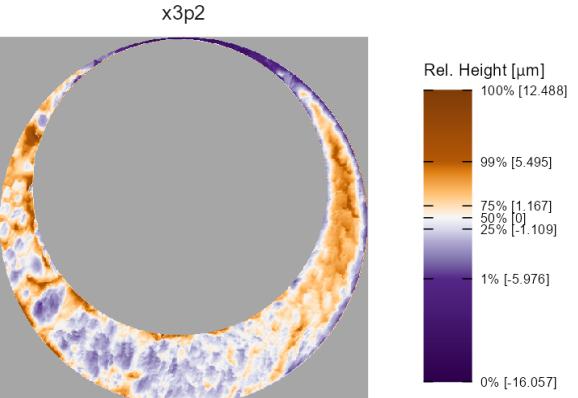
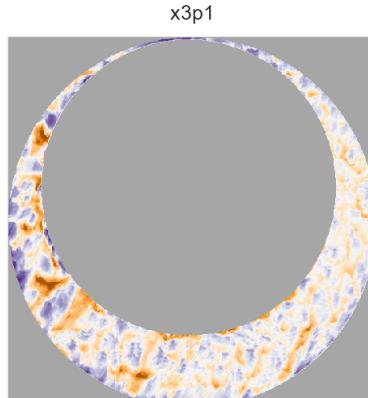
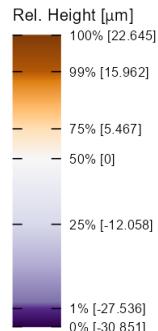
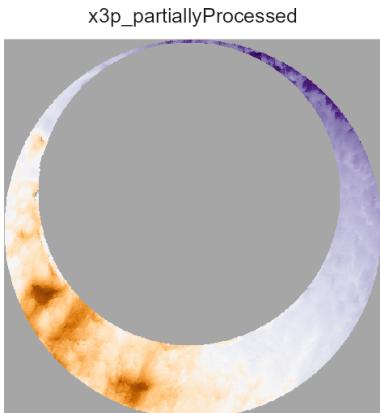


# The cmcR package

- Break down each step of the algorithm into simple "modules."
- Arrange modules in-sequence with the pipe `%>%` operator.

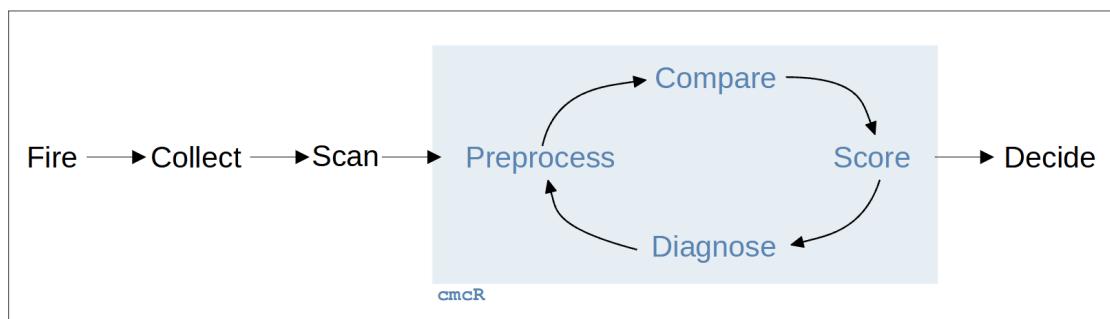
```
x3p1 <- x3p_partiallyProcessed %>%  
  preProcess_removeTrend() %>%  
  preProcess_gaussFilter()
```

```
x3p2 <- x3p_partiallyProcessed %>%  
  preProcess_gaussFilter() %>%  
  preProcess_removeTrend()
```



# Takeaways from Paper #1

- Modularization enables experimentation and improvement.
- Modules are easily reordered or replaced.
- Open-source code and data should be the standard in forensics.
- Encourages a more equitable justice system.
- `cmcR` provides a FAIR (Findable, Accessible, Interoperable, and Reusable) and "tidy" framework for future development [Wil+16; Wic+19].
  - Implementation can be incorporated into other workflows.
  - Simple functions are composed with the pipe.
  - Designed for humans.

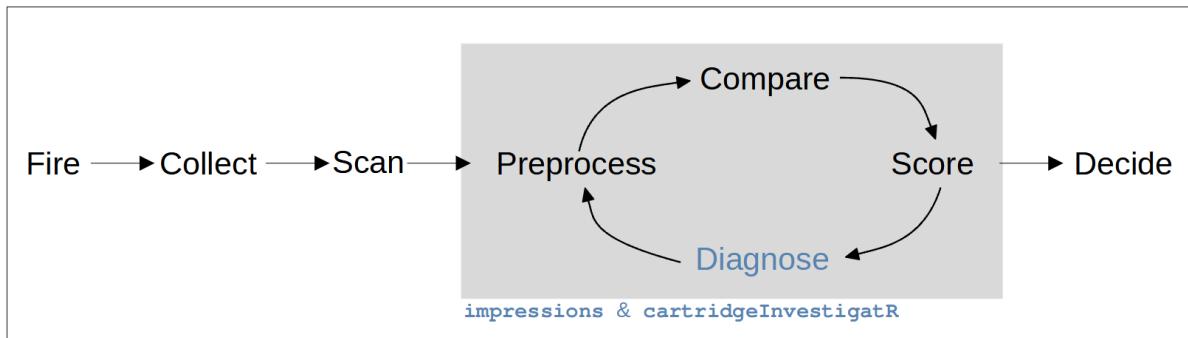


# Paper #1 To-do

- Resubmission #2
  - Previous submissions were "overly critical" of original (NIST) authors.
  - We've changed our tone to be aspirational.
  - We will resubmit within the next few weeks.
- Cartridge case data published on Figshare or ISU DataShare

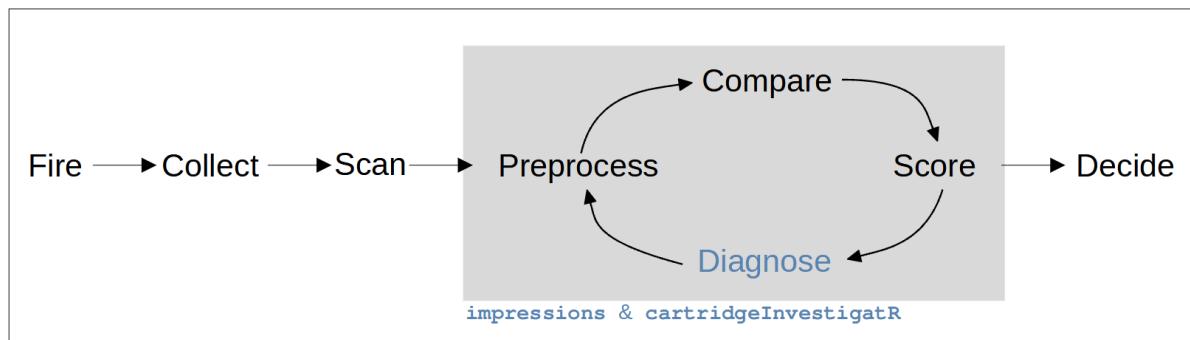
# Diagnostic Tools for Cartridge Case Comparison Algorithms

*To be submitted to The Journal of Computational and Graphical Statistics*



# Motivation

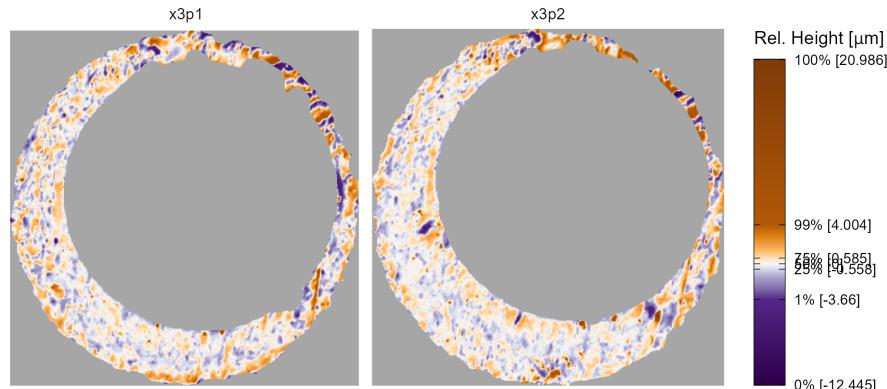
- Create static and interactive diagnostic tools to explain each step of the comparison pipeline.
  - Implement diagnostic functions in the **impressions** package to be compatible with **cmcR**.
  - Combine **cmcR** pipeline and **impressions** diagnostics in a web app.
- Useful for both researchers **and** practitioners.
  - Researchers will identify when the process goes awry.
  - Practitioners will understand and be able to explain the process.
- This is symbiotic: diagnostics inform pipeline improvements (e.g., feature engineering), which in-turn require more sophisticated diagnostics.



# x3pListPlot

- Relevant pipeline step: Preprocess
- Use to:
  - Assess isolation of identifiable impressions
  - Identify large, non-BF markings that may affect comparison
- Computational details:
  - Surface value quantiles are mapped to a divergent color scale.

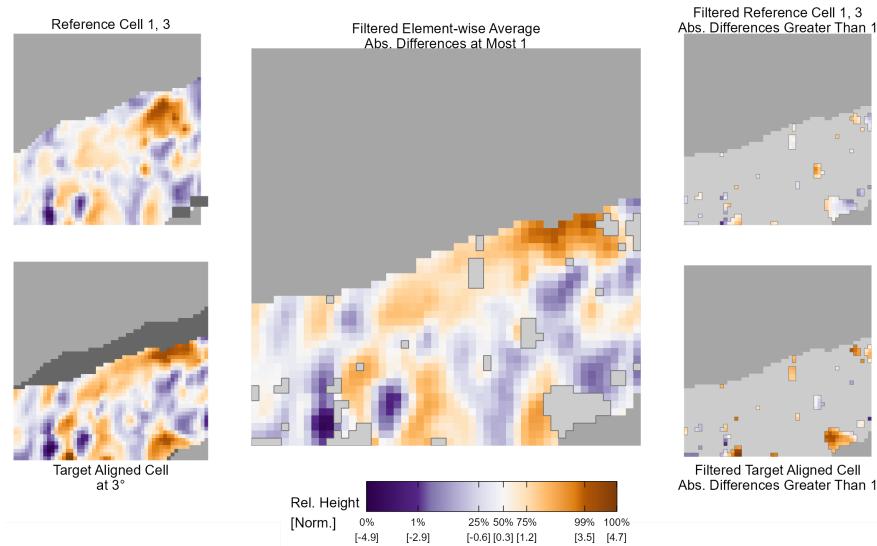
```
x3pListPlot(list("reference" = x3p1, "target" = x3p2))
```



# x3pComparisonPlot

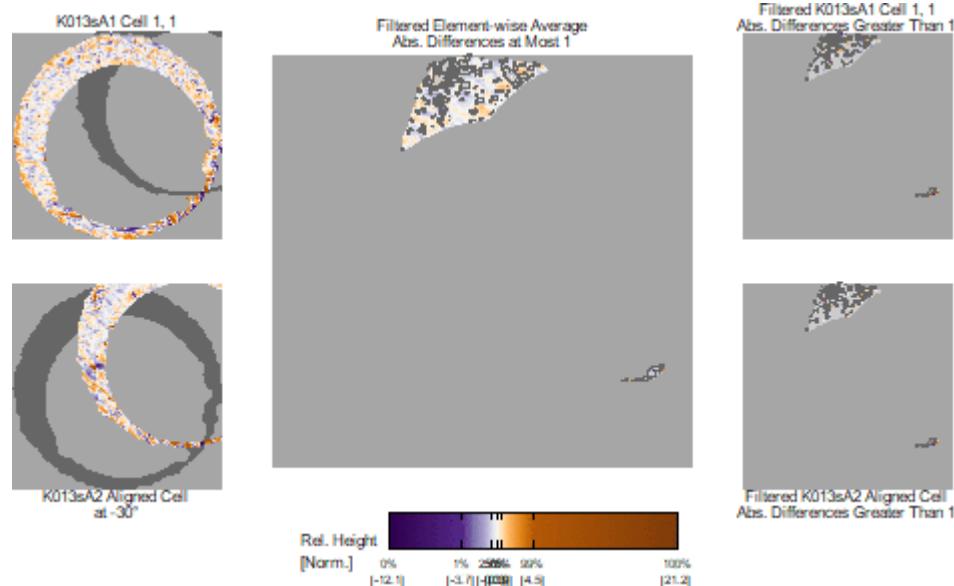
- Relevant pipeline step: Compare
- Left column shows aligned cells (dark gray represents non-overlap)
- Middle column shows the absolute difference-filtered element-wise average.
- Right column shows surface values for which the absolute difference between the two scans is larger than 1 s.d. (in this example)

```
x3pComparisonPlot(reference = refCell,target = targCell,cutoffThresh = 1)
```



# x3pComparisonPlot

- Also useful for comparing full scans.
- Note the similarity at a  $3^\circ$  rotation of the second scan.

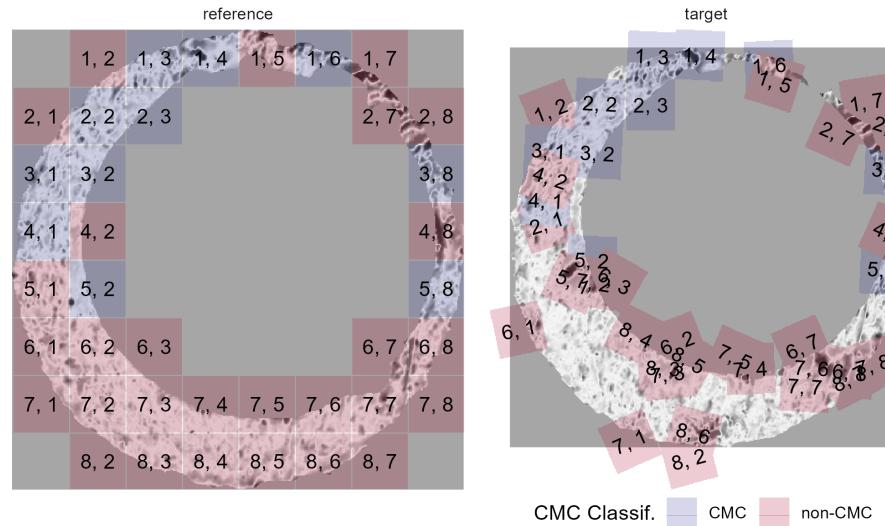


- The cell-wise vs. scan-wise plots provide different insights.
  - Cell-wise plot assesses the rotation/translation estimation procedure.
  - Scan-wise plot measures the similarity between the full scans.

# cmcPlot

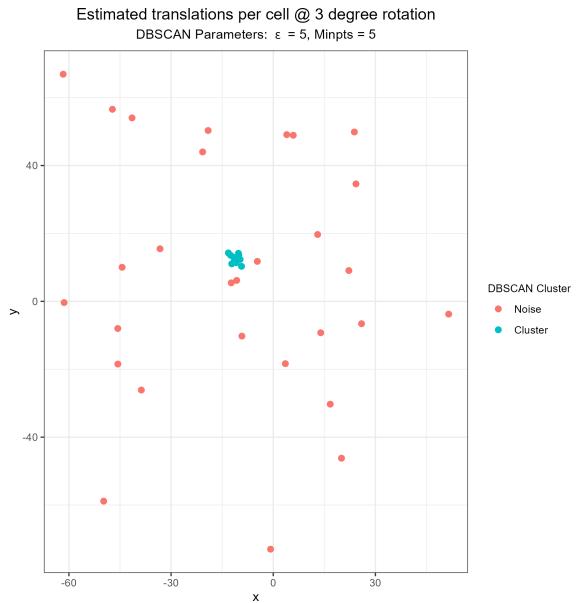
- Relevant pipeline step: Score
- Use to identify the spatial relationship between CMCs.
- Computational details:
  - Use estimated rotation and translation of each cell to calculate alignment in the target scan.

```
cmcPlot(reference = x3p1,target = x3p2,cmcClassifs = compData)
```



# DBSCAN Clusters

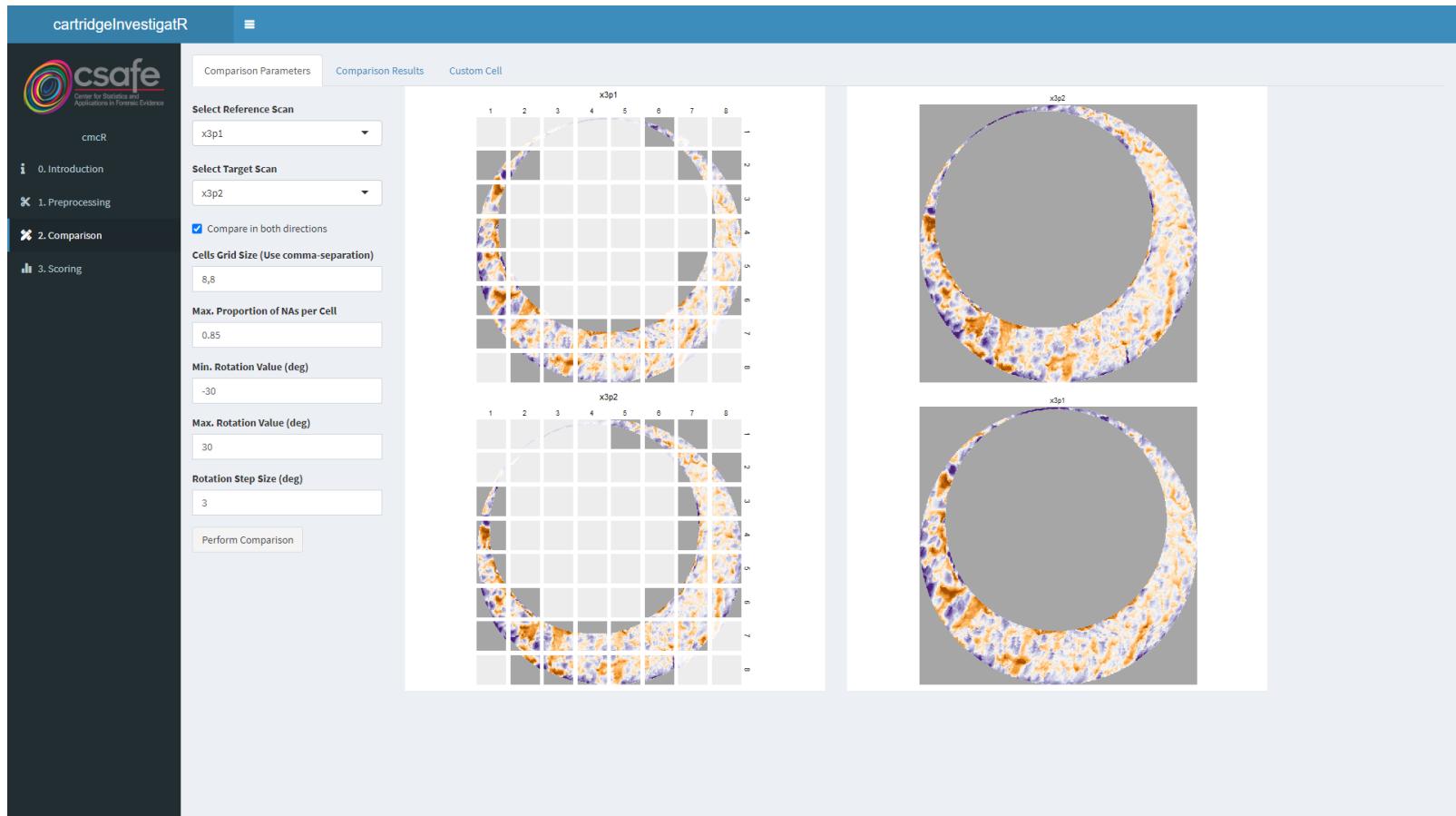
- Relevant pipeline step: Compare, Score
- Use to identify the number of cells that "agree" on a rotation/translation
- Computational details:
  - Density-Based Spatial Clustering of Applications with Noise algorithm identifies clusters vs. noise based on point density [Est+96].<sup>[1]</sup>



[1] See appendix for more details on DBSCAN algorithm.

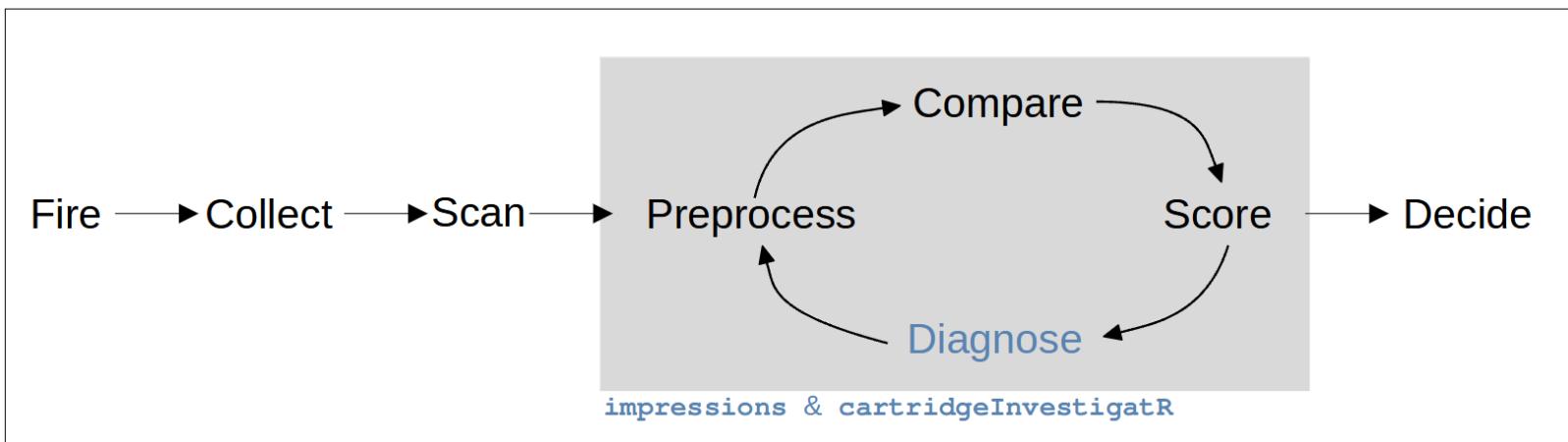
# cartridgeInvestigatR

- shiny web application [Cha+21] to combine cmcR pipeline and impressions diagnostics.



# Takeaways from Paper #2

- Diagnostics help both researchers and practitioners.
  - Researchers study and correct the behavior the pipeline.
  - Practitioners understand and can explain the pipeline.
- Visual diagnostics inform pipeline improvements (e.g., new preprocessing steps, feature engineering).
- Interactive diagnostics allow users to formulate and answer their own research questions about the pipeline.

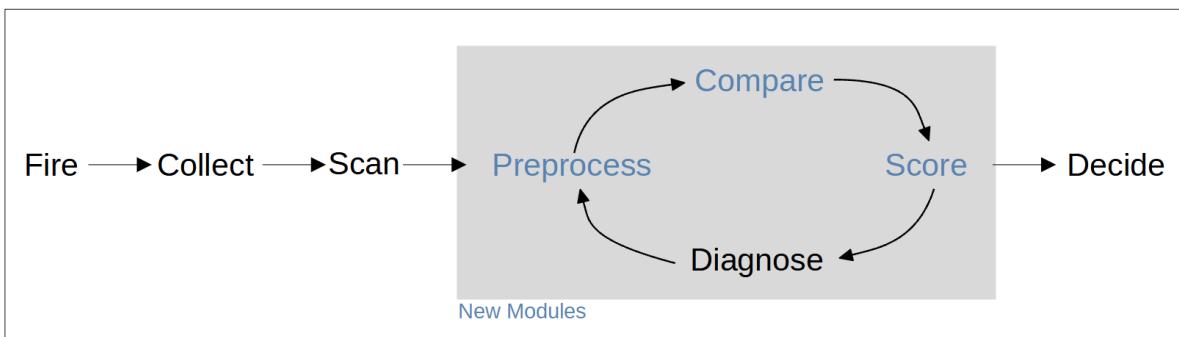


# Paper #2 To-do

- `impressions` package:
  - Tidy-up the functionality (e.g., break into simple, `%>%`-able functions).
  - Publish to CRAN (Comprehensive R Archive Network).
- `cartridgeInvestigatR` app:
  - Incorporate diagnostic functions and new `cmcR` modules.
  - Implement manual annotation of scans for preprocessing.
  - Automatically calculate most similar regions based on annotation.
  - Bug fixes

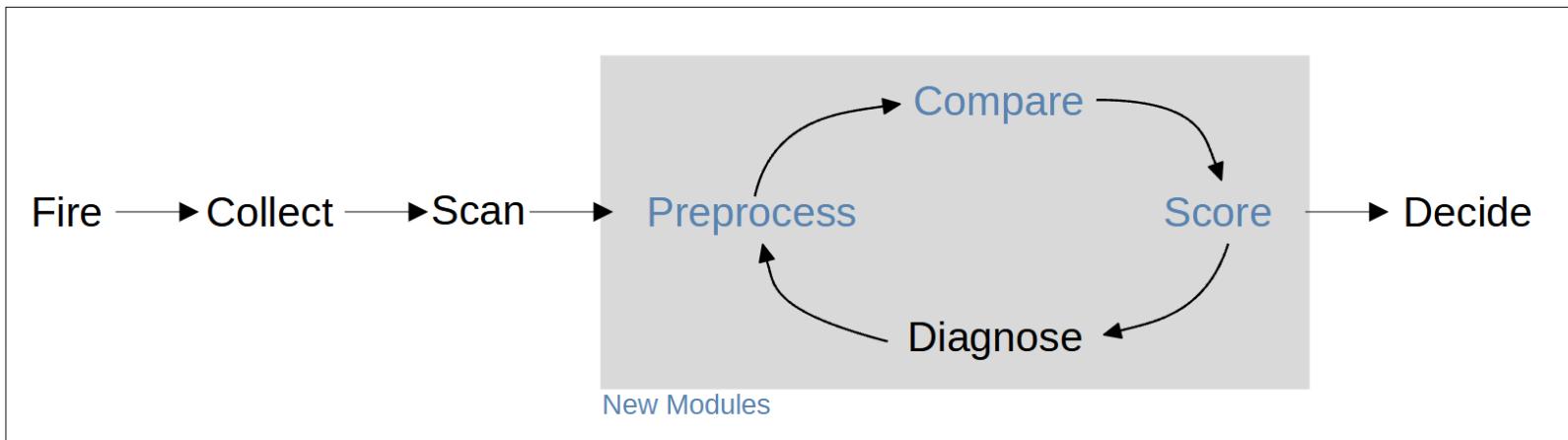
# Automatic Matching of Cartridge Case Impressions

*To be submitted to The Journal of Data Science,  
Statistics, and Visualizations*



# Motivation

- Use diagnostic tools to inform improvements to the pipeline.
- Validate these improvements on a data set of 510 scans.
- Train & test a classifier
  - If possible, use resulting score (e.g., random forest probability) to calculate score-based likelihood ratios (SLRs).<sup>[1]</sup>

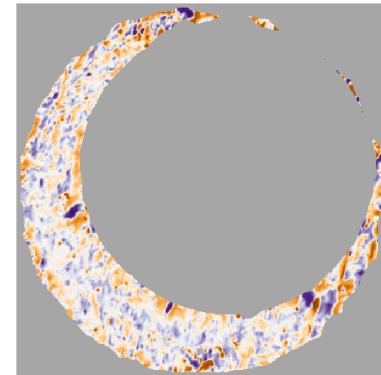
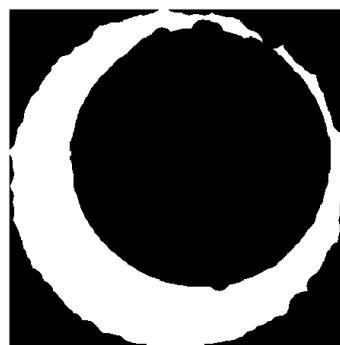
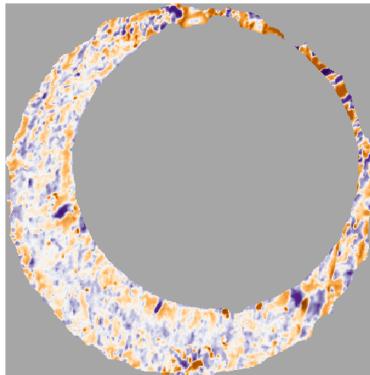


[1] See appendix for SLR details.

# Morphological Operations

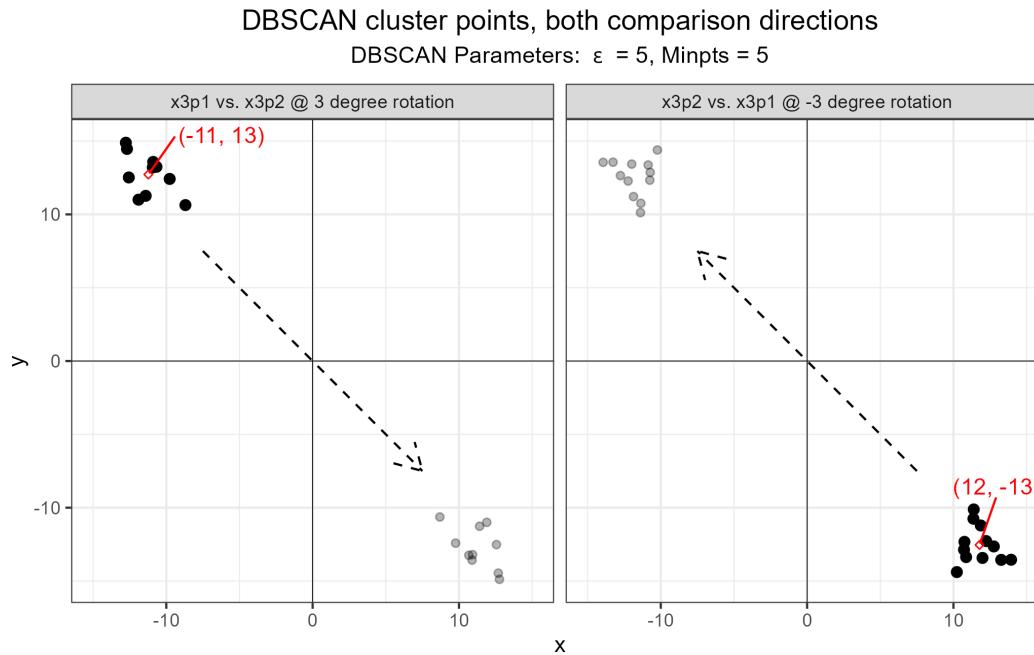
- "Some regions on the sample, such as the primer roll-off, are not informative, but may significantly affect similarity results if they are compared." [Soo22]
- Erosion and dilation "shave" off primer roll-off.
- `cmcR::preProcess_erosion` function

```
x3p2 %>%
  preProcess_erosion(region = "interior",morphRadius = 5) %>%
  preProcess_erosion(region = "exterior",morphRadius = 5)
```



# DBSCAN Algorithm Features

1. Perform comparison procedure in *both* directions.
2. Calculate rotation at which the points reach highest density,  $\hat{\theta}_1, \hat{\theta}_2$ .
3. Use cluster centroids (means) as estimated translations,  
 $(\widehat{\Delta x}_1, \widehat{\Delta y}_1), (\widehat{\Delta x}_2, \widehat{\Delta y}_2)$ .
4. Calculate cluster sizes in both directions,  $C_1, C_2$ .



# DBSCAN Algorithm Features

For a truly matching cartridge case pair, we assume:

- Rotations and translations in both directions are opposites of each other.
- Cluster sizes are large in both directions.

Features:

1) Absolute sum of estimated rotations:  $\Delta R \equiv |\hat{\theta}_1 + \hat{\theta}_2| = |3 + (-3)| = 0$

2) Sum of squared sums of estimated translations:

$$\Delta T \equiv (\widehat{\Delta x}_1 + \widehat{\Delta x}_2)^2 + (\widehat{\Delta y}_1 + \widehat{\Delta y}_2)^2 = (-11 + 12)^2 + (13 + (-12))^2 = 2$$

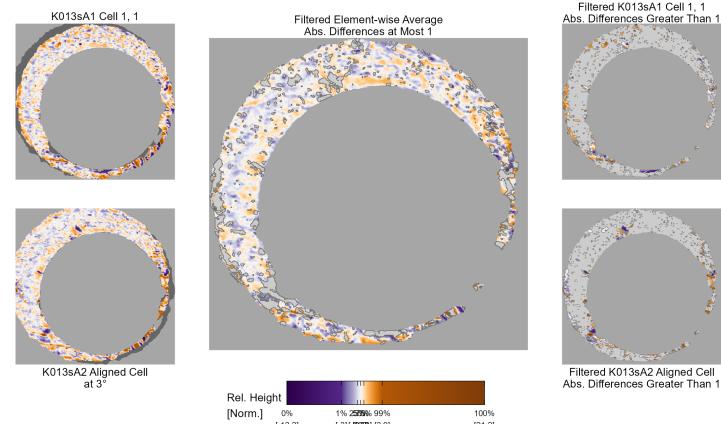
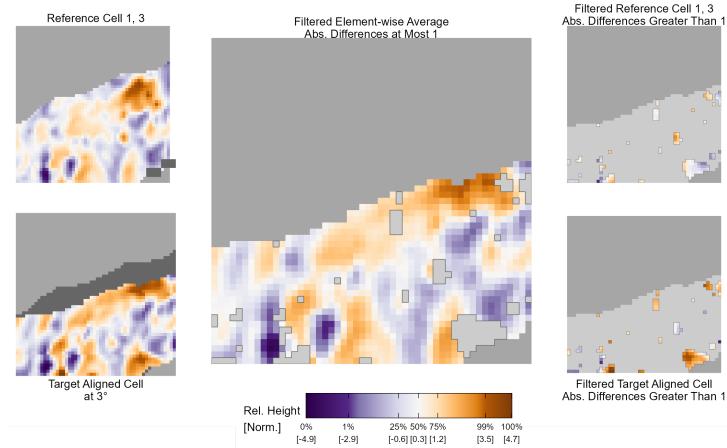
3) Average cluster size across both directions:

$$\bar{C} = \frac{1}{2}(C_1 + C_2) = \frac{1}{2}(10 + 12) = 11$$

- For matching scans, features 1 and 2 should be small and 3 should be large.

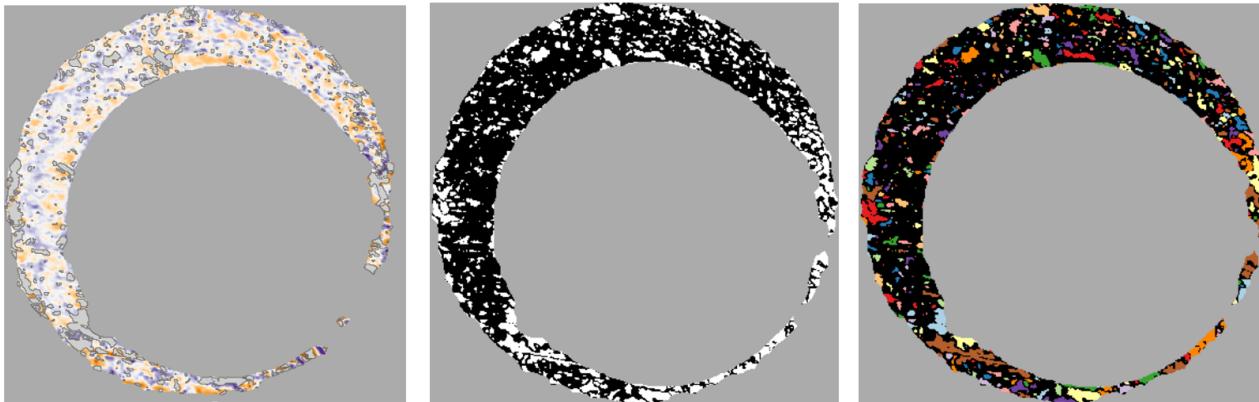
# Comparison Plot Features

For a truly matching cartridge case pair, we assume dissimilar regions will be small at the cell and full scan level, on average.



# Comparison Plot Features

- Connected Components algorithm [HS92] to identify *blobs* of filtered indices; say  $\mathcal{B} = \{B_1, B_2, \dots, B_L\}$  where  $B_l$  is  $l$ th set of connected indices.

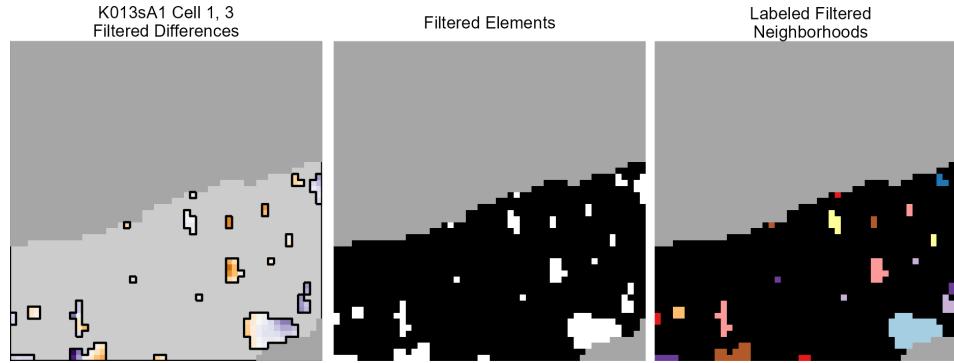


- Noting that we obtain a set of connected components  $\mathcal{B}_{d,full} = \{B_{full,1}, B_{full,2}, \dots, B_{full,L_d}\}$  per direction  $d = 1, 2$ , we average over both directions to obtain fourth feature:
  - 4) Average size of blobs for full scans:

$$\bar{B}_{full} = \frac{1}{2} \sum_{d=1}^2 \frac{1}{L_d} \sum_{l=1}^{L_d} |B_{d,l}|$$

# Comparison Plot Features

- Also compute average size of blobs across each cell:



- Noting that we obtain a set of connected components  $\mathcal{B}_{d,i} = \{B_{d,i,1}, B_{d,i,2}, \dots, B_{d,i,L_{d,i}}\}$  per cell  $i = 1, \dots, N$  and per direction  $d = 1, 2$ , we average over all cells and both directions to obtain fifth feature:

- 5) Average size of blobs across all cells:

$$\bar{B}_{cell} = \frac{1}{2} \sum_{d=1}^2 \frac{1}{N_d} \sum_{i=1}^{N_d} \frac{1}{L_{d,i}} \sum_{l=1}^{L_{d,i}} |B_{d,i,l}|$$

- Features 4 and 5 should be small for truly matching scans.

# Paper #3 Proposed Validation

For each pair of cartridge cases, we obtain five features:

- Three DBSCAN algorithm features
  - 1) Absolute sum of estimated rotations,  $\Delta R$
  - 2) Sum of squared sums of estimated translations,  $\Delta T$
  - 3) Average cluster size across both directions,  $\bar{C}$
- Two comparison plot features:
  - 4) Average size of blobs for full scans,  $\bar{B}_{full}$
  - 5) Average size of blobs across all cells,  $\bar{B}_{cell}$
- Data set: 510 cartridge cases fired from 25 barrels
  - C.C.s fired from the same barrel are "matches", otherwise "non-matches."
  - Randomly select 10 barrels (210 scans,  $\binom{210}{2} = 21945$  comparisons) for training. The remaining  $\binom{300}{2} = 44850$  comparisons are for testing.

# Training Procedure

1) For each training pair,  $p = 1, \dots, 21945$ :

1.1) Direction  $d = 1$ , one scan is "reference" and other is "target":

1.1.1) Using a rotation grid  $\Theta = \{-30, -27, \dots, 27, 30\}$ ,

calculate rotation & translation,  $(\hat{\theta}_1, \widehat{\Delta x}_1, \widehat{\Delta y}_1)$  that maximizes the CCF between the full scans.

1.1.2) After aligning the scans by  $(\hat{\theta}_1, \widehat{\Delta x}_1, \widehat{\Delta y}_1)$ , perform the cell-based comparison procedure on a small rotation grid  $\Theta^* = \{\hat{\theta} - 2, \dots, \hat{\theta} + 2\}$ .

1.1.3) Return estimated translations/rotations and blobs:

$\{(\hat{\theta}_{1,i}, \widehat{\Delta x}_{1,i}, \widehat{\Delta y}_{1,i}, \mathcal{B}_{d,i}, \mathcal{B}_{d,full}) : i = 1, \dots, N_1\}$ .

1.2) Repeat 1.1.1 - 1.1.3 in direction  $d = 2$ .

1.3) Return  $(\Delta R, \Delta T, \bar{C}, \bar{B}_{full}, \bar{B}_{cell}, M_p)_p$  where  $M_p \in \{0, 1\}$  is match/non-match ground truth for pair  $p$ .

2) Fit classifier to training data, optimizing classification accuracy.

# Testing Procedure

3) For each testing pair,  $q = 1, \dots, 44850$ :

    3.1) Repeat steps 1.1 through 1.3 of previous slide to calculate features  $(\Delta R, \Delta T, \bar{C}, \bar{B}_{full}, \bar{B}_{cell}, M_q)_q$ .

    3.2) Predict match/non-match class  $\hat{M}_q$  using trained classifier.

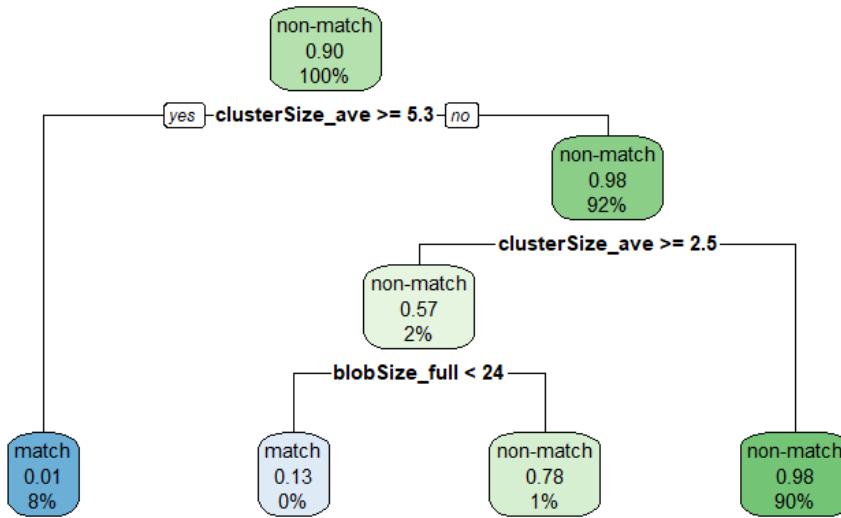
4) Compare predicted class  $\hat{M}_q$  to ground-truth  $M_q$  using 0-1 loss.

Possible classifiers:

- Decision tree (easily interpretable)
- Random forest (yields class probabilities)
- Others?

# Preliminary Training Results

- `clusterSize_ave` = average cluster size,  $\bar{C}$
- `blobSize_full` = average blob size for full scans,  $\bar{B}_{full}$ .



- Nodes: (1) classification, (2) probability of being a non-match given the node, (3) percent of observations falling into the node

# Preliminary Training Results

The training data confusion matrix for this method is given below:

	<b>Predicted Match</b>	<b>Predicted Non-Match</b>
<b>Actual Match</b>	1848	351
<b>Actual Non-Match</b>	39	19707

Training error rate is 1.8%; 0.20% False Positive and 16.0% False Negative

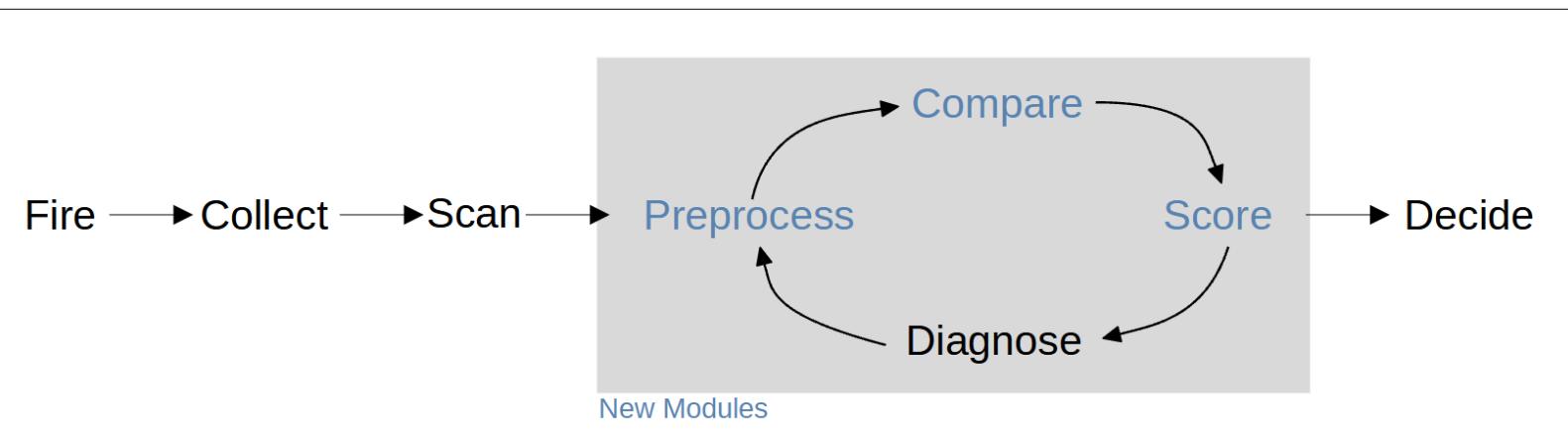
Compare to our implementation of the CMC method with optimized parameters on the same data set:

	<b>Predicted Match</b>	<b>Predicted Non-Match</b>
<b>Actual Match</b>	1419	780
<b>Actual Non-Match</b>	257	19489

Error rate is 4.7%; 1.3% False Positive and 35.5% False Negative

# Takeaways from Paper #3

- Diagnostics are used to inform improvements to the pipeline
  - Morphological transformations help remove primer roll-off effects.
  - DBSCAN algorithm provides a cogent alternative to the CMC scoring method
  - Visual diagnostic features are intuitive to explain
- Comparison pipelines are validated using the proposed train/test procedure.



# Paper #3 To-do

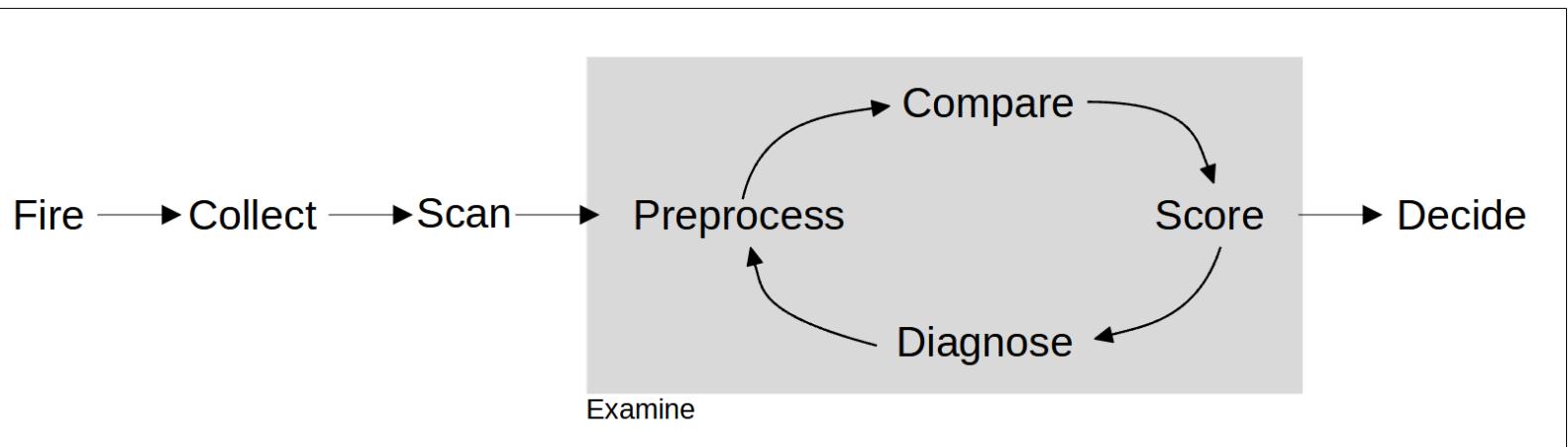
Pipeline validation

- Calculate features for test data
- Consider additional features:
  - Measures of spread for blob features or estimated rotation/translation
  - Correlation measures
- Fit random forest or other classifier
  - Compare SLR results to those from CMC method

Create new package with improved pipeline functions

- Use same "tidy" philosophy to implement
- Use `parsnip` R package from `tidymodels` to implement train/test procedure [KW20]

# CONCLUSIONS



- Paper #1 to *The R Journal*, Re-submit by 6/30/22
  - A modularized, tidy implementation of the cartridge case comparison pipeline.
  - Eases experimentation and improvement.
  - Encourages reproducibility.
- Paper #2 to JCGS, Submit by 11/1/2022
  - Tools to diagnose, explain, and suggest improvements to the pipeline.
  - Useful to researchers to improve the pipeline.
  - Useful to practitioners to understand and explain the pipeline.
- Paper #3 to JDSSV, Submit by 4/1/2022
  - Improvements to the pipeline.
  - Uses diagnostics to inform new modules and features.
  - Will be the largest (published) study of a C.C. comparison algorithm.
- Dissertation defense by 4/21/22 (last day for Spring 2023 semester)



# REFERENCES

AFTE Criteria for Identification Committee (1992). "Theory of identification, range striae comparison reports and modified glossary definitions". In: *AFTE Journal* 24.3, pp. 336-340.

Baldwin, D. P. et al. (2014). *A Study of False-Positive and False-Negative Error Rates in Cartridge Case Comparisons*. URL: <https://doi.org/10.21236/ada611807>. DOI: [10.21236/ada611807](https://doi.org/10.21236/ada611807).

Chang, W. et al. (2021). *shiny: Web Application Framework for R*. R package version 1.7.1.

Ester, M. et al. (1996). "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise". In: *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. KDD'96. . Portland, Oregon: AAAI Press, p. 226–231.

Haralick, R. M. et al. (1992). *Computer and Robot Vision*. Vol. 1. Addison-Wesley, pp. 28-48.

Hare, E. et al. (2017). "Automatic Matching of Bullet Land Impressions". In: *The Annals of Applied Statistics* 11.4, pp. 2332-2356.





# REFERENCES

Krishnan, G. et al. (2018). "Adapting the Chumbley Score to Match Striae on Land Engraved Areas (LEAs) of Bullets, ". In: *Journal of Forensic Sciences* 64.3. URL: <https://doi.org/10.1111/1556-4029.13950>, pp. 728-740.

Kuhn, M. et al. (2020). *Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles*.

National Research Council (2009). *Strengthening Forensic Science in the United States: A Path Forward*. Washington, DC: The National Academies Press. ISBN: 978-0-309-13130-8.

President's Council of Advisors on Sci. & Tech. (2016). "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods".

Riva, F. et al. (2014). "Automatic Comparison and Evaluation of Impressions Left by a Firearm on Fired Cartridge Cases". In: *Journal of Forensic Sciences* 59.3. URL: <https://doi.org/10.1111/1556-4029.12382>, pp. 637-647.





# REFERENCES

## Warning: Riva2020: unknown macro '\textendash'

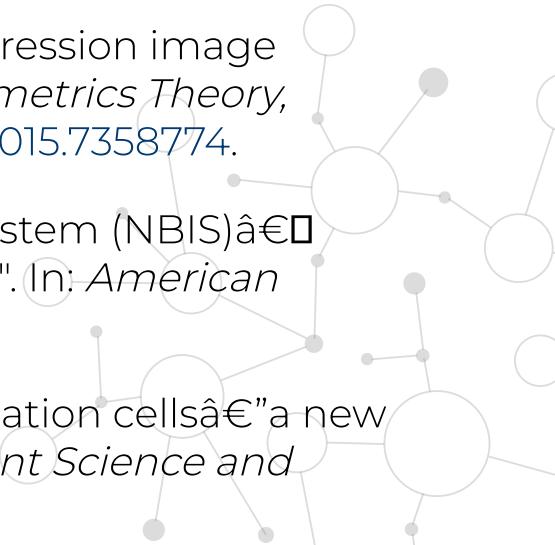
Riva, F. et al. (2014). "Automatic Comparison and Evaluation of Impressions Left by a Firearm on Fired Cartridge Cases". In: *Journal of Forensic Sciences* 59.3. URL: <https://doi.org/10.1111/1556-4029.12382>, pp. 637-647.

Riva, F. et al. (2020). "Comparison and interpretation of impressed marks left by a firearm on cartridge cases \textendash Towards an operational implementation of a likelihood ratio based technique". In: *Forensic Science International* 313. URL: <https://doi.org/10.1016/j.forsciint.2020.110363>, p. 110363.

Roth, J. et al. (2015). "Learning-based ballistic breech face impression image matching". In: *2015 IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. , pp. 1-8. DOI: [10.1109/BTAS.2015.7358774](https://doi.org/10.1109/BTAS.2015.7358774).

Song, J. (2013). "Proposed â€œNIST Ballistics Identification System (NBIS)â€ Based on 3D Topography Measurements on Correlation Cells". In: *American Firearm and Tool Mark Examiners Journal* 45.2, p. 11.

Song, J. et al. (2014). "3D topography measurements on correlation cellsâ€a new approach to forensic ballistics identifications". In: *Measurement Science and Technology* 25.6, p. 064005.





# REFERENCES

Soons, J. (2022). *Computer-Aided Firearm and Toolmark Analysis*. Forensic Technology Center of Excellence.

Tai, X. H. et al. (2018). "A Fully Automatic Method for Comparing Cartridge Case Images". In: *Journal of Forensic Sciences* 63.2, pp. 440-448.

Thompson, R. (2017). *Firearm Identification in the Forensic Science Laboratory*. National District Attorneys Association, p. 32.

Tong, M. et al. (2015). "An Improved Algorithm of Congruent Matching Cells (CMC) Method for Firearm Evidence Identifications". In: *Journal of Research of the National Institute of Standards and Technology* 120, p. 102. DOI: [10.6028/jres.120.008](https://doi.org/10.6028/jres.120.008).

Wickham, H. et al. (2019). "Welcome to the tidyverse". In: *Journal of Open Source Software* 4.43, p. 1686. DOI: [10.21105/joss.01686](https://doi.org/10.21105/joss.01686).

Wilkinson, M. D. et al. (2016). "The FAIR Guiding Principles for scientific data management and stewardship". In: *Scientific Data* 3.1. DOI: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18).

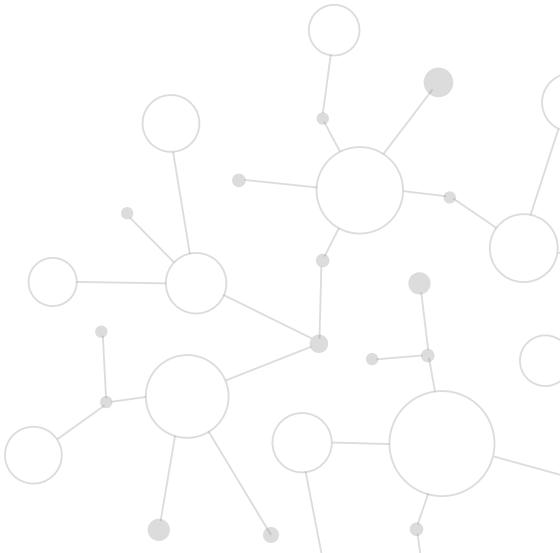




# REFERENCES

```
## Warning in `[[.BibEntry`(x, ind): subscript out of bounds
```

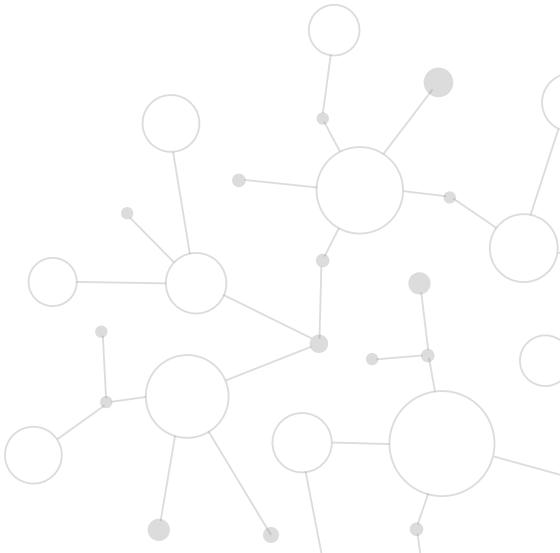
Zheng, X. et al. (2014). "Applications of surface metrology in firearm identification". In: *Surface Topography: Metrology and Properties* 2.1. , p. 014012.





# REFERENCES

```
## Warning in `[[.BibEntry`(x, ind): subscript out of bounds
```



# **APPENDIX**

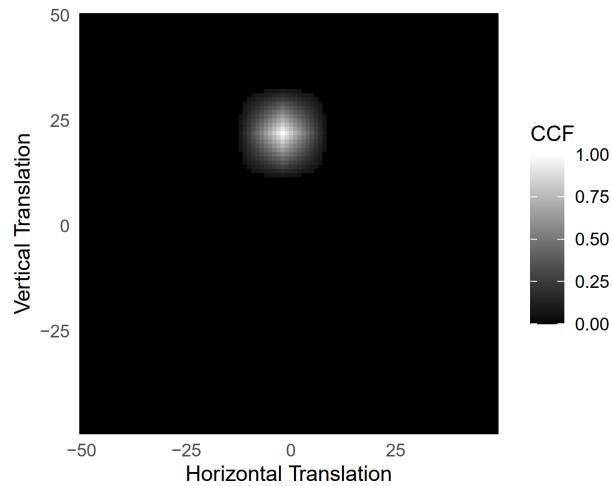
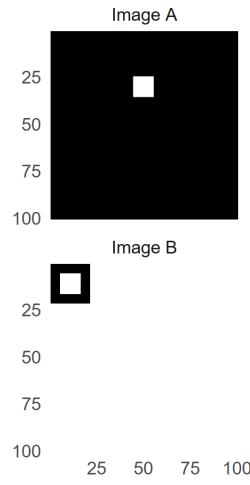
# Cross-Correlation Function

- Let  $A$  be a  $P \times Q$  "reference" image and  $B$  a  $M \times N$  "target" image.
- The *cross-correlation function* (CCF), denoted  $(A \star B)$ , is given by:

$$(A \star B)[m, n] = \sum_{i=1}^M \sum_{j=1}^N A[i, j]B[(i + m), (j + n)]$$

where  $1 \leq m \leq M + P - 1$  and  $1 \leq n \leq N + Q - 1$ .

- Measures similarity between  $A$  and  $B$  translated  $m$  units horizontally and  $n$  units vertically.



# *DBSCAN Algorithm*

- In the example below,  $Minpts = 4$  and  $\epsilon$  is not specified
- Clusters consist of:
  - Red *core* points have at least 4 points in their  $\epsilon$ -neighborhood
  - Yellow points are neighbors of a core point
- Blue *noise* points do not belong to a cluster

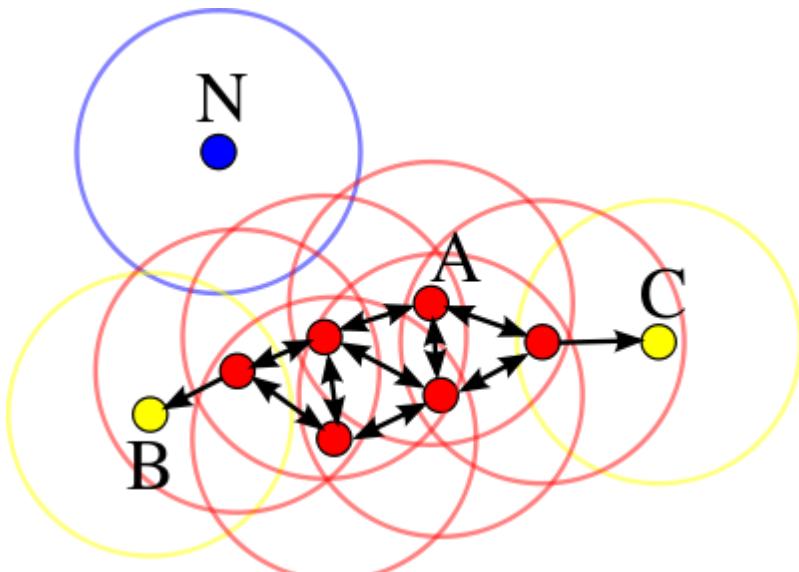


Image source: [Chi11]

# *Common Source Score-Based Likelihood Ratios*

Given recovered evidence of unknown source,  $\mathbf{u}_x$  and  $\mathbf{u}_y$ , consider the following hypotheses:

- $H_p$  (prosecution): The recovered evidence comes from the same source
- $H_d$  (defense): The recovered evidence comes from two different sources

$\mathbf{u}_x, \mathbf{u}_y$  are commonly modeled as random variables governed by an unknown distribution. In-practice, evidence is often high-dimensional (e.g., a cartridge case scan) making it intractable to assume or estimate its distribution. Instead, the evidence is mapped to some low-dimensional score whose distribution is easier assume or estimate [OS18].

Let  $C(\mathbf{u}_x, \mathbf{u}_y)$  denote a score function of the evidence and  $g(C(\mathbf{u}_x, \mathbf{u}_y)|H)$  its pdf under hypothesis  $H$ .

The *common source score-based likelihood ratio* (SLR) is:

$$\delta_{CS}(\mathbf{u}_x, \mathbf{u}_y) = \frac{g(C(\mathbf{u}_x, \mathbf{u}_y)|H_p)}{g(C(\mathbf{u}_x, \mathbf{u}_y)|H_d)}$$

Commonly,  $C$  is a complex function (e.g., mapping two cartridge case scans to a match class probability) and  $g$  is estimated with data (e.g., KDE).