

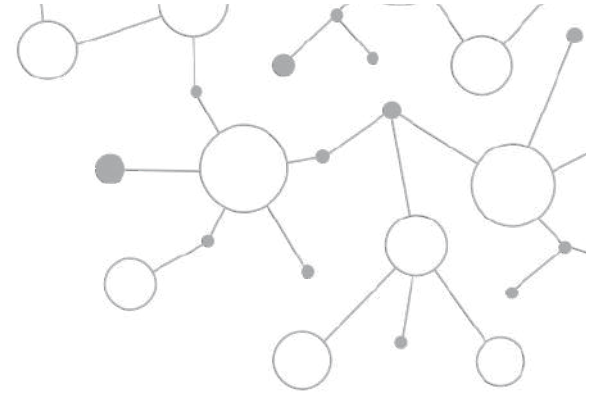
Similarity between outsole impressions using SURF

Soyoung Park, Ph.D.

Postdoctoral Research Associate
CSAFE/Iowa State University

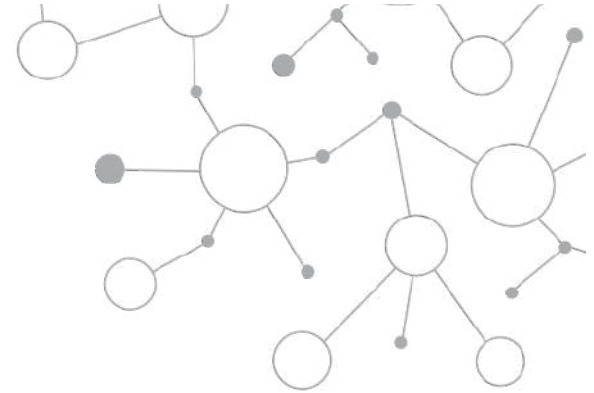
Feb. 2020

Publication



Park, S. and Carriquiry, A., 2020, **An algorithm to compare 2D footwear outsole images using maximum clique and SURF**, *Statistical Analysis and Data Mining*, In press.

It takes a village

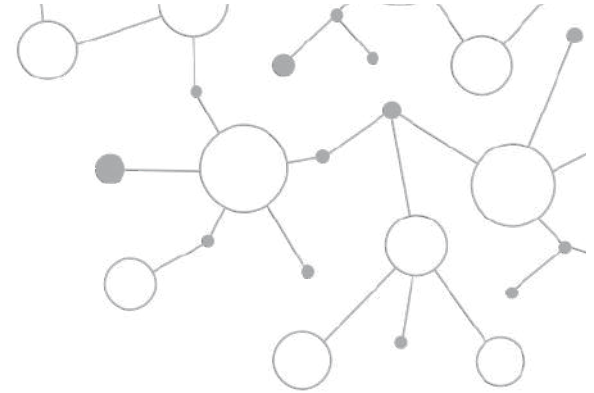


Collaborators:

- Dr. Hari Iyer, NIST
- Sarena Wiesner, Yaron Shor, Israeli Police
- Dr. Guillermo Basulto-Elias, Mr. James Kruse, CSAFE
- Ms. Lesley Hammer, Hammer Forensics
- Dr. Eric Hare, Omni Analytics
- A small army of super smart undergraduates, Iowa State University.



Goals of presentation

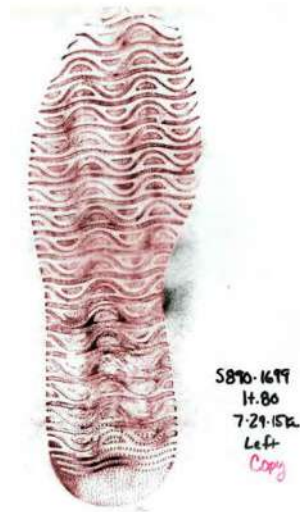


- Introduce an objective method to quantify the similarity between two outsole impressions.
- Show that algorithm is accurate and reliable even when outsoles share class characteristics and degree of wear.
- Show that algorithm is robust even when one image is degraded and partially observed.

The forensic question

- Could the shoe on the left be the source of the impression on the right?
- Q = questioned
- K = known

Challenging!

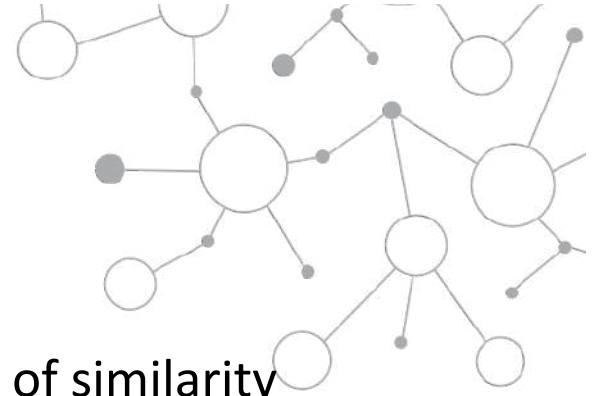


The state of the art

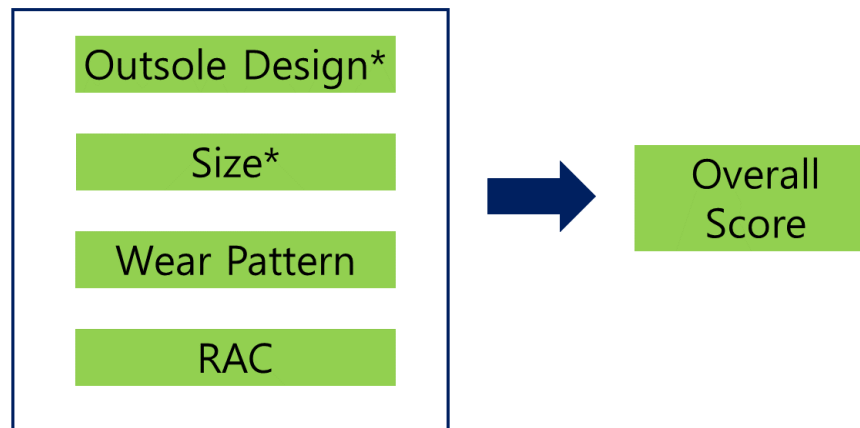


- At present, practitioners rely on training and experience to:
 - Identify features of interest that can be used to compare outsoles.
 - **Subjectively** determine whether outsoles are “similar enough” to suggest same source.
 - Use 7-scale decisions (SWGTHREAD)
- Two implicit assessments:
 - How similar are the outsoles?
 - How probative is observed degree of similarity.

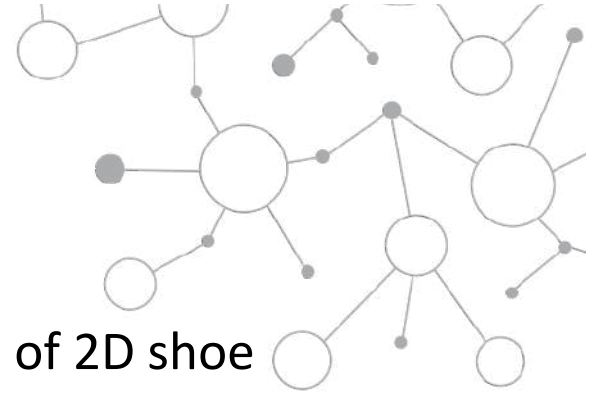
Objective of this project



- Develop a *score* that quantifies the degree of similarity between two outsole 2D images.
- Contributors to the score



Data collection 1



- CSAFE constructed a longitudinal database of 2D shoe outsole impressions.
- 160 participants were recruited and received a pair of new shoes.
- Participants were asked to use the shoes and return to CSAFE every eight weeks, for six months.



NIKE Winflow 4

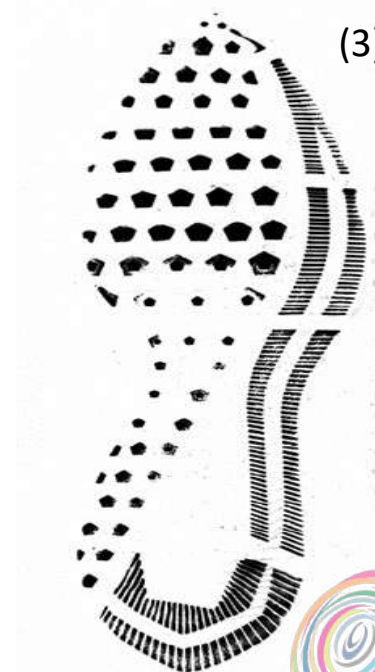
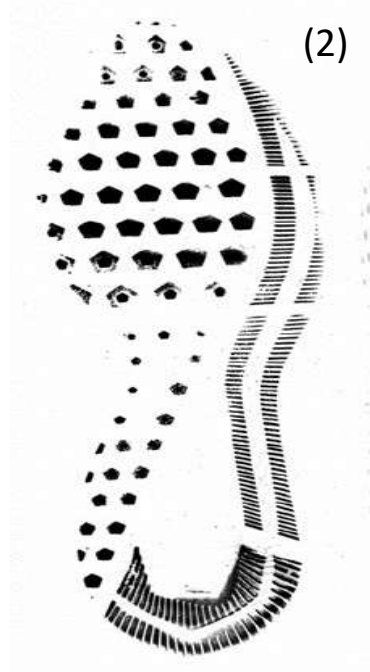
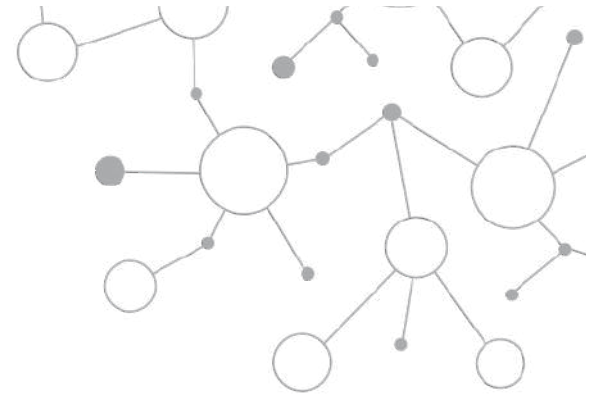


Adidas Seeley Skateboard

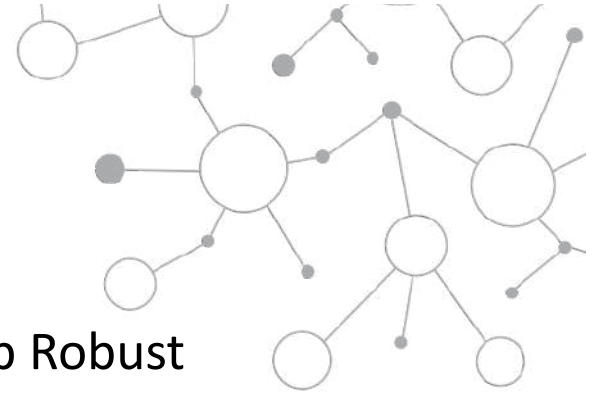


Data collection 2

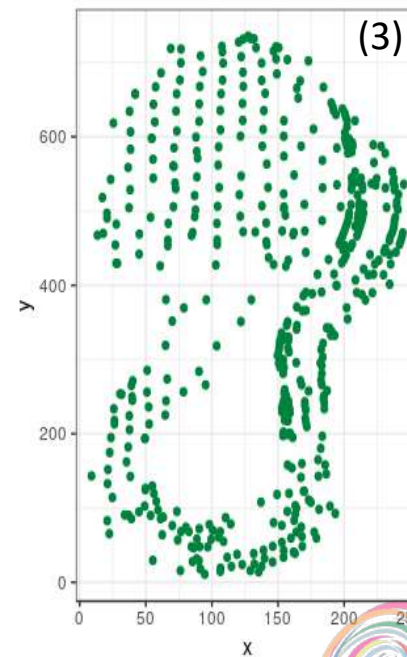
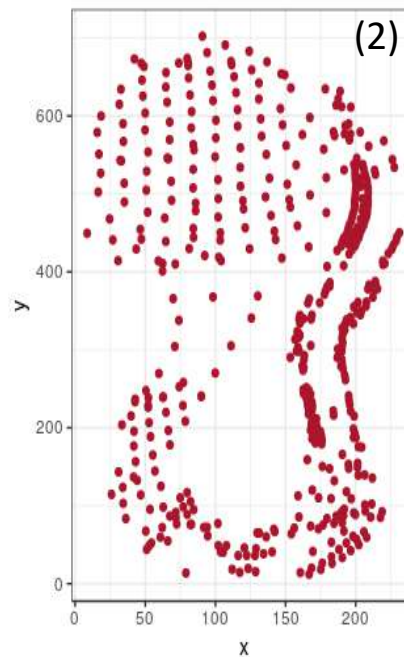
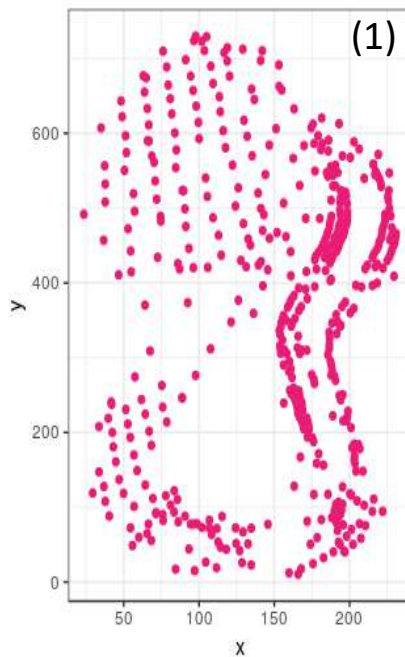
- Can you visually find the differences?
- Mates: (1) and (2) & Non-mates: (1) and (3)



Algorithm I - SURF

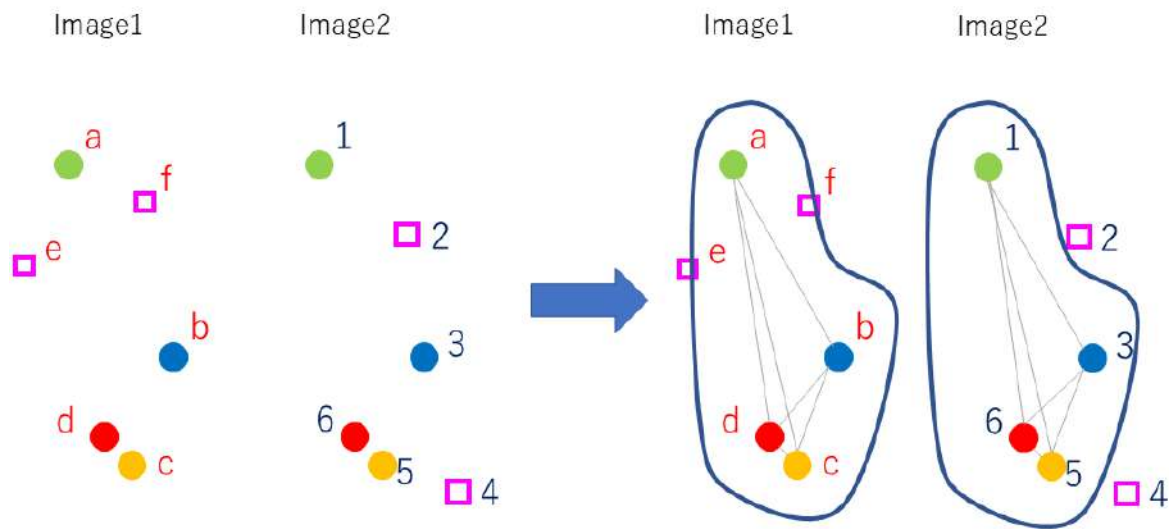
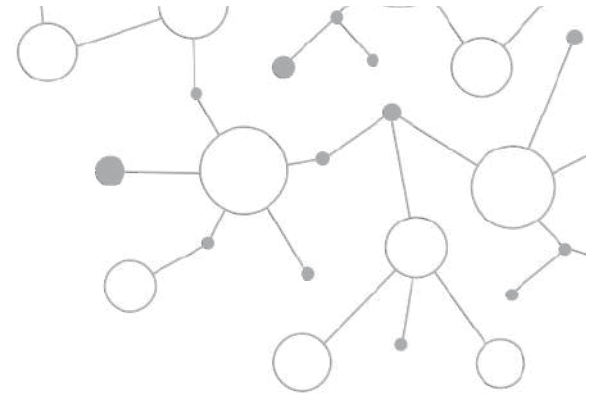


- We used the 500 strong SURF (Speeded-Up Robust Features (Bay et al. 2006)) as the points of interest.



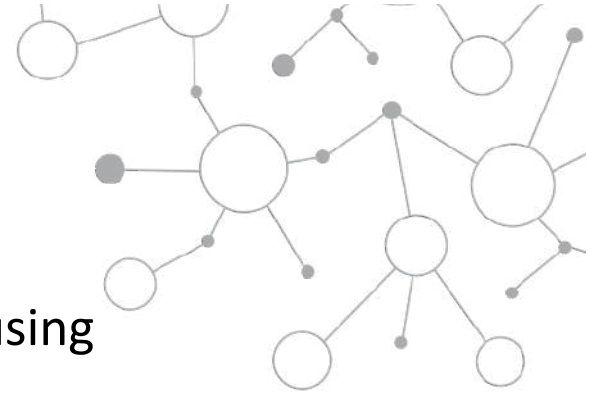
Algorithm II - Alignment

- Maximum clique from the graph theory



R-package: Shoeprintr

- Delicately developed for image matching using maximum clique
- Parallelized maximum clique with smart sampling
- Visit <https://github.com/CSAFE-ISU/shoeprintr>



```
README.md

Shoeprint Boosted

How-To

1. Make sure you have the devtools package installed
2. From R, run devtools::install_github("CSAFE-ISU/shoeprintr")
3. (If running on the CSAFE server or another machine that restricts executing binary files) From a Terminal, run chmod +x /home/username/R/x86_64-pc-linux-gnu-library/3.4/shoeprintr/bin/linux64/pac, replacing the path as necessary
4. From R, Load shoeprintr by typing library(shoeprintr)
5. Run the following sample code, modifying the input files and parameters as necessary:

## Load in the input and the reference images
data(input_example)
data(reference_example)

## Transform all the points to have (0,0) at lower left corner.
print_in <- leftcorner_cent(input_example)
print_ref <- leftcorner_cent(reference_example)

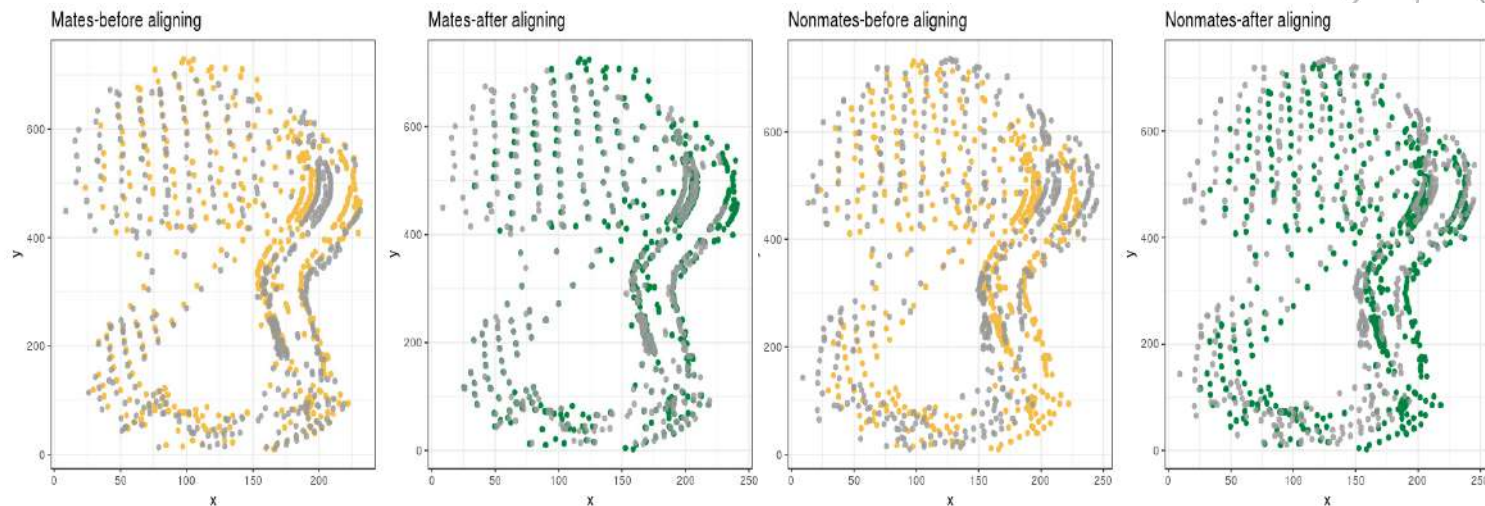
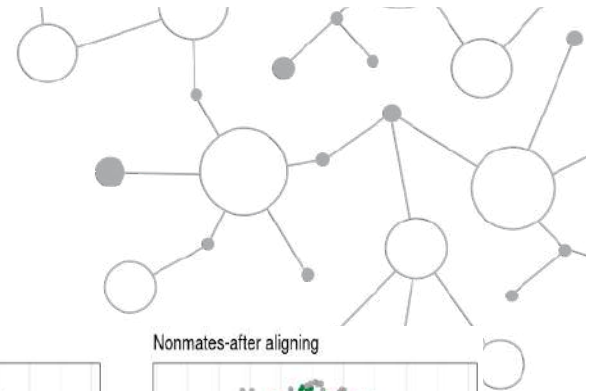
## Perform Print Match
print_stats <- match_print(print_in, print_ref,
  ncross_in_bins = 30, xbins_in = 20, ncross_in_bin_size = 1,
  ncross_ref_bins = NULL, xbins_ref = 30, ncross_ref_bin_size = NULL,
  eps = .75, seed = 1, num_cores = parallel::detectCores(),
  plot = TRUE, verbose = TRUE)

print_stats

6. (Optional) We provide the pac binary with the R package. If you would like to update this binary to use the latest development version, follow the instructions at the bottom of this README file.
```



Aligned images

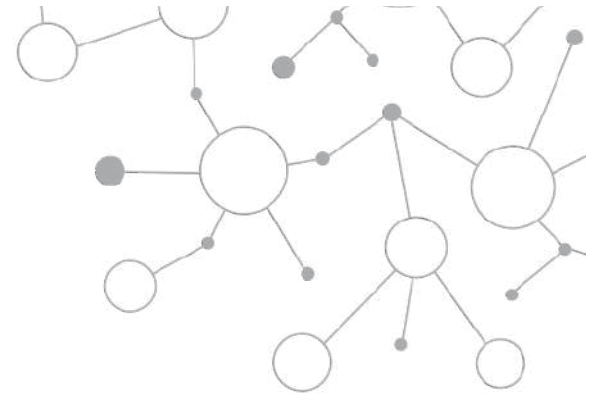


Mates; 56.5% Overlap

Non-mates; 12.1% Overlap

Class	Clique size	Rot. Angle	% Overlap	Med. distance
Mates	18	2.11	56.46%	0.78
Non-mates	9	6.43	12.08%	1.39

Source prediction performance

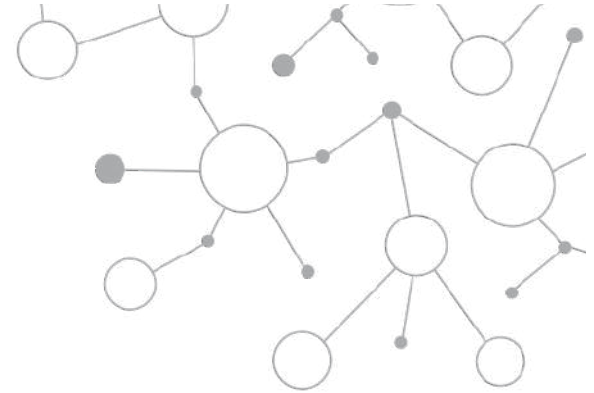


- Prediction result by the RF method

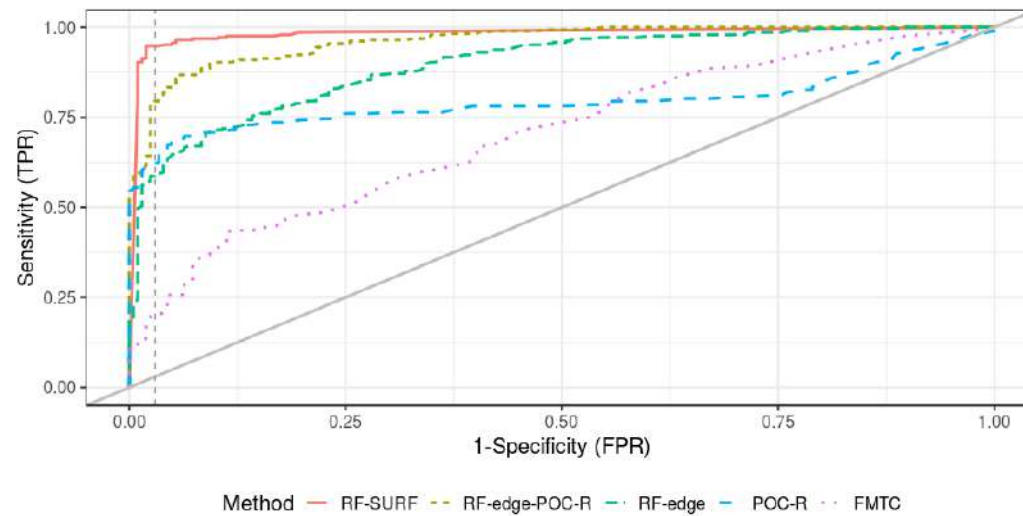
Prediction	Truth	
	Mates	Non-mates
Predicted same shoe	277	8
Predicted different shoe	11	196
Total	288	204

- 96% Sensitivity and 96% Specificity
- 4% Error rate

Source prediction performance

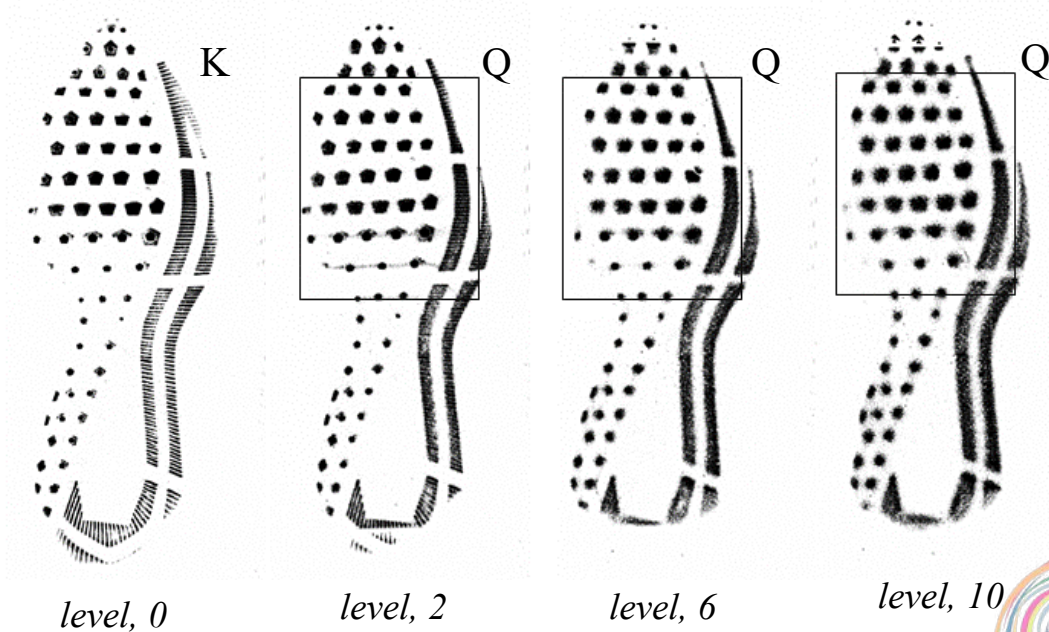
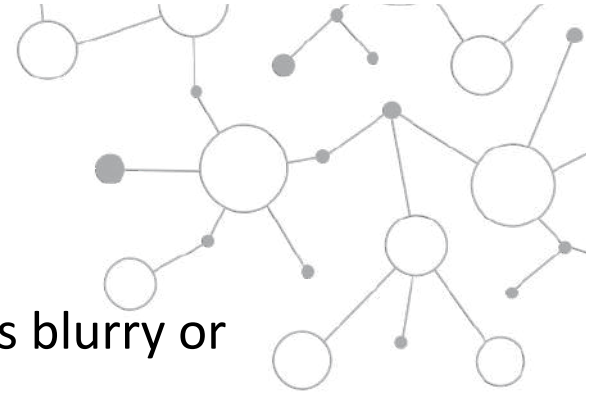


- ROC curves

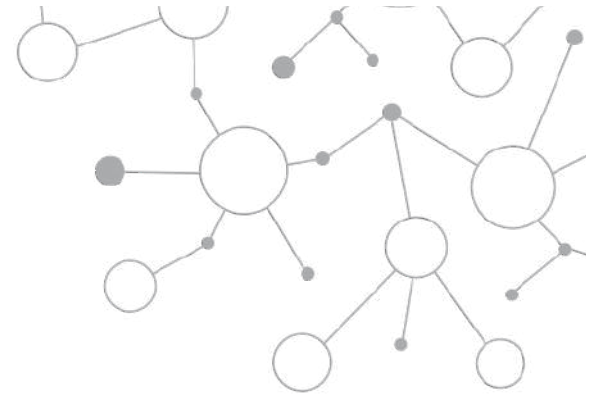


Degraded impressions

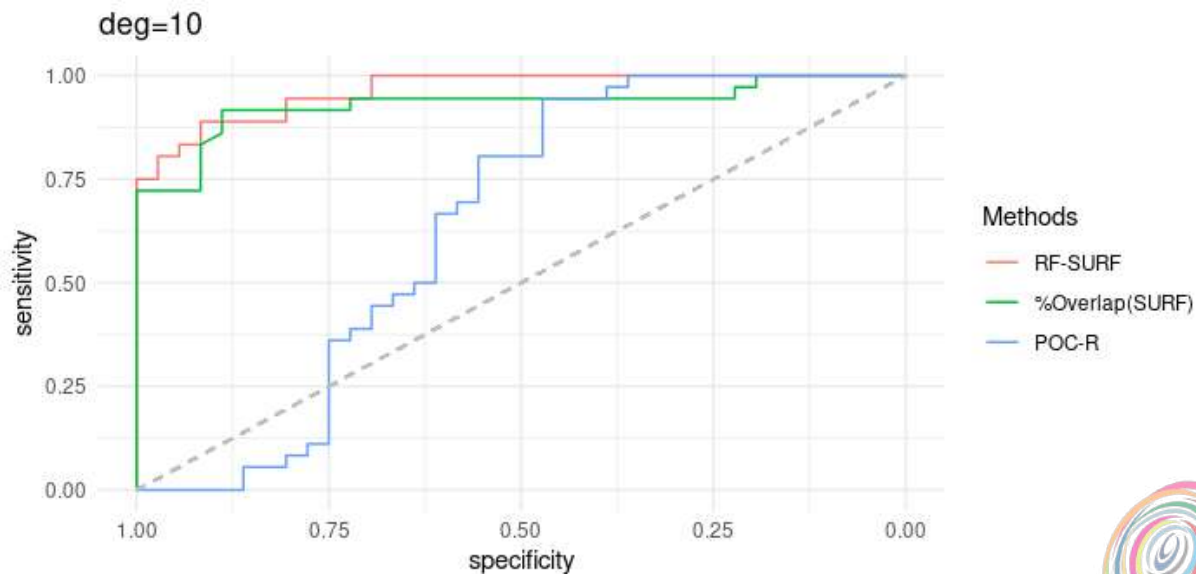
- Often, the latent print at the crime scene is blurry or partially observed.
- Q (partial and blurry) vs. K (full and clean)



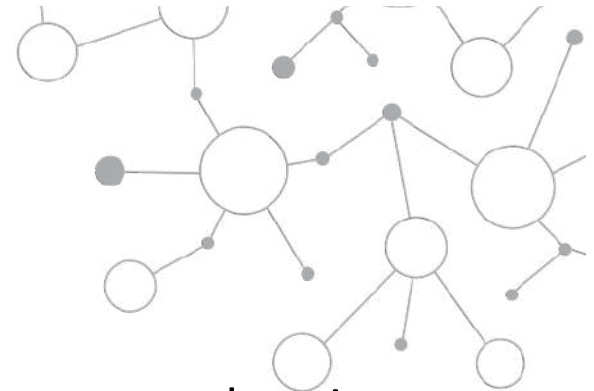
Performance on degraded Q



- ROC curves
- *Q (toe area and level 10 degraded) vs. K (full and clean)*
 - False positive rate: 3% -- Incorrectly conclude same shoe
 - False negative rate: 9% -- Incorrectly conclude diff shoe

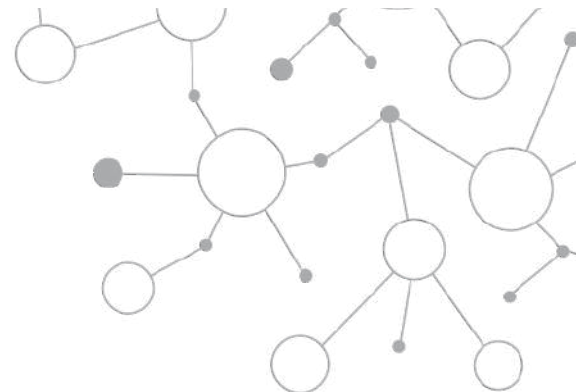


Summary



- Automatically finding interesting points in an outsole using SURF appears promising.
- When both impressions are of reasonably good quality, an algorithm based on three features has excellent performance, at least for Nike Airflows and Adidas Steeley.
- When Q is degraded, features no longer serve to predict whether Q and K were made by same shoe.
- **Exception is % points that overlap; this feature is robust on degraded and partially observed images.**

Open questions



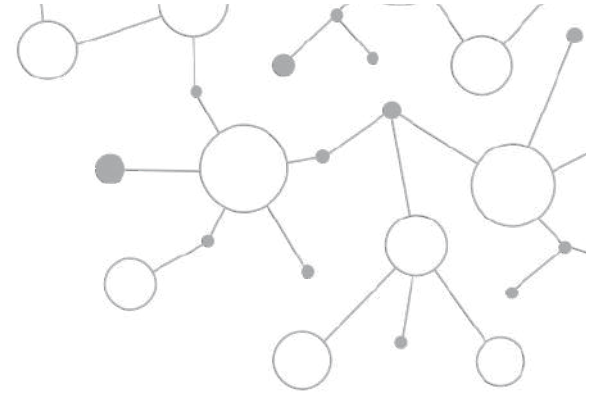
- Will algorithms continue to perform well when:
 - We include other brands of shoes (we think yes).
 - Q is degraded in some other way (we need more research).
- Black-box study: compare the outcome of the automated method to the scores produced by a trained examiner.
- Similarity algorithm using convolutional neural network.

References



- Park, S and Carriquiry, A, 2019: Similarity of two-dimensional images: An application to the forensic comparison of shoe outsole impressions, manuscript under the revision requested by *Journal of Applied Statistics*.
- Park, S., 2018. Learning algorithms for forensic science applications.
- Richetelli, N., Lee, M.C., Lasky, C.A., Gump, M.E. and Speir, J.A., 2017. Classification of footwear outsole patterns using Fourier transform and local interest points. *Forensic science international*, 275, pp.102-109.
- Bay, H., Tuytelaars, T. and Van Gool, L., 2006, May. Surf: Speeded up robust features. In *European conference on computer vision* (pp. 404-417). Springer, Berlin, Heidelberg.
- Breiman, L., 2001. Random forests. *Machine learning*, 45(1), pp.5-32.

Thanks!



- Please find me at:

sypark@iastate.edu

- Our website:

www.forensicstats.org

