



Opinion

The Daily Campus

EDITORIAL BOARD:

Karla Perez EDITOR-IN-CHIEF

Tomas Hinckley..... OPINION EDITOR

Patrick Minnerly.. ASSOCIATE OPINION EDITOR

Colin Hamilton..... WEEKLY COLUMNIST

Gianna Cassino..... WEEKLY COLUMNIST

Disclaimer: The views and opinions expressed by individual writers in the opinion section do not reflect the views and opinions of The Daily Campus or other staff members. Only articles labeled “Editorial” are the official opinions of The Daily Campus.

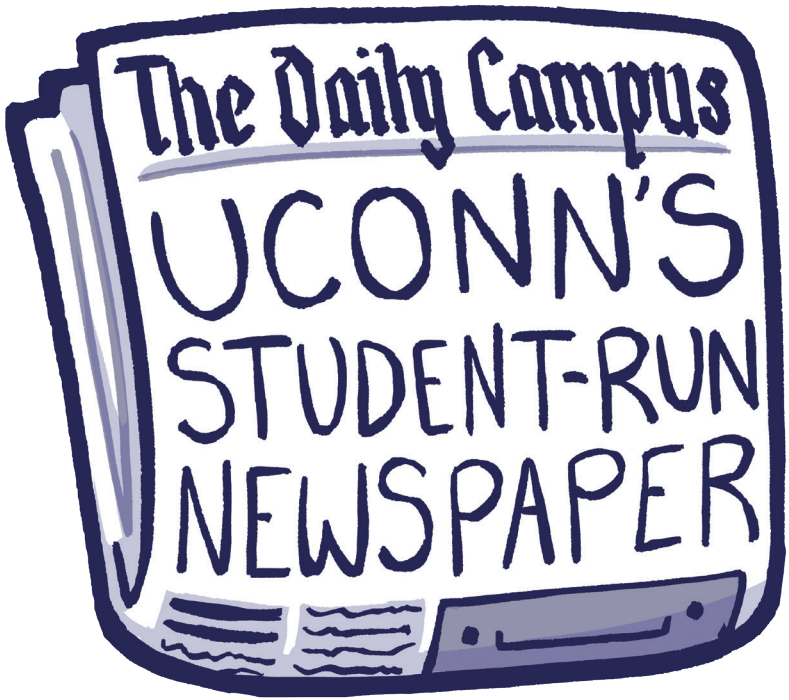


ILLUSTRATION BY COLBI LORANGER, ASSOCIATE MANAGING EDITOR/THE DAILY CAMPUS

op-ed

Every Shoe Sole Is Unique:

A new statistical model can now prove it in court

by Alokesh Manna,
UConn Department of Statistics
aalokesh.manna@uconn.edu
by Neil Spencer
UConn Department of Statistics
bnell.spencer@uconn.edu
and by Dipak K. Dey
UConn Department of Statistics
cdipak.dey@uconn.edu

When investigators arrive at a crime scene, one of the most overlooked yet powerful clues they may find is a shoe print. But ‘matching’ a shoe to a print with ‘class characteristics’ such as brand, type and size is considered the first step; it is more complicated than it sounds and a team of University of Connecticut statisticians is working to make that process more rigorous and reliable. A new study by Alokesh Manna, Professor Neil Spencer and Distinguished Professor Dipak K. Dey, all from UConn’s Department of Statistics, introduces a novel Bayesian statistical model that substantially improves how forensic scientists evaluate shoe print evidence. The paper, recently posted to arXiv (arXiv:2602.07006, Manna, Spencer and Dey (2026)), answers the question: how do you prove a shoe is truly unique?

The Problem with “Matching” a Shoe

Identifying a shoe’s make and model from a crime scene print is a useful first step, but it is far from definitive. Thousands of identical shoes roll off production lines each year, like the Nike Air Force 1, which sells 10 million pairs annually. What actually distinguishes one individual shoe from another is called “accidentals”: the unique cuts, scrapes, holes and embedded debris on a given shoe. These accumulate over time through everyday wear, walking pattern, pronation, etc. No two shoes develop the exact same pattern of accidentals.

Theoretically, if class characteristics and accidentals match 100%, this will indicate the suspect’s shoe is almost certainly the source from the crime scene, but practical challenges are common. Crime scene prints are sometimes of lower quality and the accidentals may vary slightly during the evidence processing steps, leading to uncertainty deviating from the theoretical perspective. The key forensic question becomes how rare this particular pattern of accidentals is. If a suspect’s shoe matches a crime

scene print not just in size and model but also in its unique pattern of wear damage, investigators need a way to quantify how unlikely that match is by coincidence. That probability is called the ‘random match prob-

Previous approaches to estimating random match probabilities were either too simple — assuming accidentals are equally likely anywhere on a shoe sole, which is known to be false — or too computationally

cess — a statistical framework for describing where random events occur in images and links those locations to the specific tread pattern of each shoe. Areas with heavy tread contact, and especially areas with

ferent regions of the shoe, the team achieves a 60 times more accurate picture of where accidentals are likely to appear.

Crucially, the team employed a technique called Integrated Nested Laplace Approximation, which makes Bayesian inference computationally feasible even for the large West Virginia University shoe database of 1,300 shoes — more than three times the size of the dataset used in the best previous method. In 10-fold cross-validation, the UConn proposed model achieved superior performance over all competing methods across every data split. Most strikingly, models that ignored the individual shoe’s tread pattern performed dramatically worse — a finding that underscores just how important it is to account for each shoe’s unique contact surface when estimating match probabilities. Getting these estimates wrong can mean dramatically overstating or understating the strength of forensic evidence in a courtroom.

“Forensic examination of shoe prints can provide compelling physical evidence linking a suspect to a crime scene. When unique characteristics such as tread patterns, wear marks, cuts, and individualizing features correspond between a recovered print and a suspect’s footwear, the likelihood of coincidence is significantly reduced. Such matching shoe print evidence, when properly documented and analyzed, can play a critical role in reconstructing events and establishing presence at the scene,” Dr. Dey said.

Why does this matter?

Forensic shoe print analysis is used in criminal investigations worldwide, yet the statistical foundations underpinning courtroom testimony have lagged behind modern data science. This research is part of a broader effort to bring rigorous, data-driven methods to forensic science in a call echoed by the National Academy of Sciences and the President’s Council of Advisors on Science and Technology. By making the model more accurate and scalable, the UConn team lays the groundwork for methods that could eventually be deployed in real forensic casework, helping courts make more reliable and defensible judgments about shoe print evidence.



Shoe prints left at a crime scene. New research from the Statistics department at UConn has helped create a more efficient way to identify the type of shoe worn.

PHOTO COURTESY OF WIKIMEDIA COMMONS

ability’, and calculating it accurately is critical to the weight of evidence presented in court; otherwise, this evidence will be vastly overstated.

A Smarter, Scalable Model

expensive to apply to large datasets. The UConn team’s new model solves both problems at once. Their approach treats the locations of accidentals on a shoe sole as a spatial point pro-

cess — sharp tread edges (captured by an “image gradient” feature), are more prone to accumulating accidentals. By modeling this relationship explicitly and allowing it to vary across dif-

op-ed