UC Irvine
UC Irvine Previously Published Works

Title

Towards a likelihood ratio approach for bloodstain pattern analysis

Permalink

https://escholarship.org/uc/item/493686zj

Authors

Zou, Tong

Stern, Hal S

Publication Date

2022-12-01

DOI

10.1016/j.forsciint.2022.111512

Towards a likelihood ratio approach for bloodstain pattern analysis

Tong Zou *,[1], Hal S. Stern

Department of Statistics, University of California, Irvine, Irvine, CA 92617, USA

ARTICLE INFO

ABSTRACT

In this work, we explore the application of likelihood ratio as a forensic evidence assessment tool to evaluate the causal mechanism of a bloodstain pattern. It is assumed that there are two competing hypotheses regarding the cause of a bloodstain pattern. The bloodstain patterns are represented as a collection of ellipses with each ellipse characterized by its location, size and orientation. Quantitative measures and features are derived to summarize key aspects of the patterns. A bivariate Gaussian model is chosen to estimate the distribution of features under a given hypothesis and thus approximate the likelihood of a pattern. Published data with 59 impact patterns and 55 gunshot patterns is used to train and evaluate the model. Results demonstrate the feasibility of the likelihood ratio approach for bloodstain pattern analysis. The results also hint at some of the challenges that need to be addressed for future use of the likelihood ratio approach for bloodstain pattern analysis.

1. Introduction

Bloodstain pattern analysis (BPA) has been employed as a means of crime scene reconstruction and testimonial evidence for more than one hundred years [21]. Frequently, bloodstain patterns encountered at a crime scene are evaluated to determine the mechanism behind the bloodletting event [10]. By analyzing the shapes, sizes, orientations and locations of bloodstains along with other information, BPA experts develop hypotheses about on how the event happened. Another task of BPA is to determine the spatial location of the bloodletting event by inferring the trajectories of blood droplets [6,13], but this is not addressed by the work reported here. Although BPA can play a critical role as a forensic tool, its accuracy and reliability have been questioned during the last decade. A comprehensive study of forensic science by the National Research Council of the National Academy of Sciences expressed concerns. The report notes that "*the opinions of bloodstain pattern analysts are more subjective than scientific*" and "*The uncertainties associated with bloodstain pattern analysis are enormous*" [30]. Current BPA approaches rely heavily on techniques taught during workshops, and on experience and experimentation. The recently published black box study [18] finds several examples of bloodstain patterns for which BPA analysts do not agree on the mechanism, and others where the majority of analysts conclude an incorrect mechanism.

There has been recent interest in developing quantitative approaches to BPA. The hope is that such approaches may allow for more scientific analyses that appropriately address uncertainty. The most common approach to date has attempted to build models to classify bloodstain patterns as being due to one of a few possible causal mechanisms [4,5,25]. In these studies, a number of quantitative features are devised based on an ellipse representation of bloodstain patterns. Then a classifier is trained using those features to distinguish between two causal mechanisms, e.g. impact vs cast-off [4] and gunshot vs impact [25]. A significant limitation of the classification framework is that a pattern will always be classified into one of the trained categories, even if it is actually caused by a different mechanism. It is not generally appropriate to say with certainty that a pattern is produced by one of only two mechanisms.

An alternative to the classification approach is the likelihood ratio (LR) approach. The LR is a concept in statistics designed as a measure of degree to which observed data is more likely under one hypothesis as against an alternative hypothesis. In the context of forensics, the LR can be used to quantitatively estimate the relative support provided by the evidence for two competing propositions of forensic interest [33]. A number of authors have advocated for the LR approach [2,34]. Attempts have been made to apply the LR approach to many kinds of forensic evidence like DNA [32], latent prints [31], digital [16,17], firearms [12] and handwriting [11,15,27]. We briefly explain the approach here. Suppose evidence is to be evaluated under two competing hypotheses H_1 and H_2. By definition LR is the ratio of the probabilities of the evidence conditioning on H_1 and H_2, in formula:

---

* Corresponding author.
E-mail addresses: zout3@uci.edu (T. Zou), sternh@uci.edu (H.S. Stern).
[1] ORCID: 0000-0003-1302-8363.

$$LR = \frac{P(E|H_1)}{P(E|H_2)} \qquad (1)$$

A value of the LR greater than one implies that the evidence is more likely under H_1 than under H_2 and therefore provides more support for H_1 than H_2, and vice versa for a value of the LR smaller than one. Larger (or smaller) value of the LR provides stronger support for H_1 (or H_2). Various authors have provided interpretations for different values of the LR (see, e.g., [23]). It is important to note that the value of the LR greater than one does not indicate that H_1 is more probable that H_2. Such a statement would be referring to posterior probabilities $P(H_1|E)$ and $P(H_2|E)$, which require specifying prior probabilities. For more on this point see [2]. The likelihood evaluation process for many evidence types models the similarity between a crime scene item and a reference item. The corresponding hypotheses concern whether they are from the same source or not. This is where bloodstain pattern analysis diverges from other forms of forensic evidence. The BPA setting does not involve assessing between the evidence and a reference item. Recently, [7] reviewed the potential for applying the LR approach in BPA. They note the complexities, describe needed development and demonstrate via a hypothetical example regarding the time of an event. This article provides a proof of concept for the LR approach in the context of mechanism determination.

The hypotheses H_1 and H_2 are assumed to represent hypotheses a bloodstain pattern is produced under two different mechanisms. For example, in the David Camm case [18], H_1 and H_2 can be taken to represent the hypotheses that bloodstain pattern is caused by gunshot or by transfer. The example we focus on, based on data availability, is discrimination of gunshot patterns and impact patterns. As noted in [25], this remains a challenging and meaningful problem in forensics. Based on this scenario, we show how the LR framework can be used to analyze bloodstain evidence.

In order to obtain the LR, we need to be able to estimate $P(E|H_1)$ and $P(E|H_2)$ in a quantitative way given a bloodstain pattern E. These are certainly cases where one can be confident that a pattern is produced by one of two mechanisms, but this need not generally be the case. In the Appendix A we show that even when one considers a wide range of potential mechanisms, it is still relevant to model $P(E|H)$ for each of the potential mechanisms. In terms of modeling the likelihood of a bloodstain pattern, our approach consists of three main steps: first, the bloodstain pattern image is processed and represented by a collection of ellipses approximating the constituent stains. We applied the technique from the work of [36] to infer the number of ellipses and the numerical characteristics of the ellipses. Second, a few predefined quantitative features are extracted from the ellipse representation of each pattern. The features are all dimensionless quantities such that they are highly robust and can be easily adapted for other data sets. Third, the distribution of the features is estimated for a set of bloodstain patterns produced by the same mechanism and used as the likelihood model. By demonstrating the potential use of the LR framework for bloodstain patterns, we aim to provide insights for future researchers and legal professionals as they work to incorporate statistical tools into reasoning about bloodstain evidence.

Section 2 describes the data we use, characteristics of the images and the image processing methods. Section 3 describes the statistical approach to evaluating the LR in detail. This includes the definition of features and the development of probability models that estimate the likelihood under different mechanisms. Application of the approach to data from impact and gunshot patterns [8,9] is described in Section 4. Discussion including limitations and challenges is presented in Section 5.

2. Data preparation

2.1. Bloodstain images

Two collections of bloodstain images are published as open source data sets, with one containing impact patterns [8] and the other gunshot spatters [9]. An impact pattern is created when an object strikes a blood source. A gunshot pattern is created when the blood source is shot by a bullet and the stains generated from high-velocity droplet projectiles. In this study we use the 59 single-source impact patterns and the 55 gunshot patterns on vertical targets from these data sets.

Here we briefly describe the experiments designed to generate the patterns. All patterns are produced indoors using fresh swine blood at room temperature. A vertical target wall covered with cardstock or butcher paper collects the flying blood droplets to form a pattern. Since the scanner window used to create the images is smaller than the target area, patterns are scanned piece-wise at a resolution of 600dpi and then stitched together using a graphics editor.
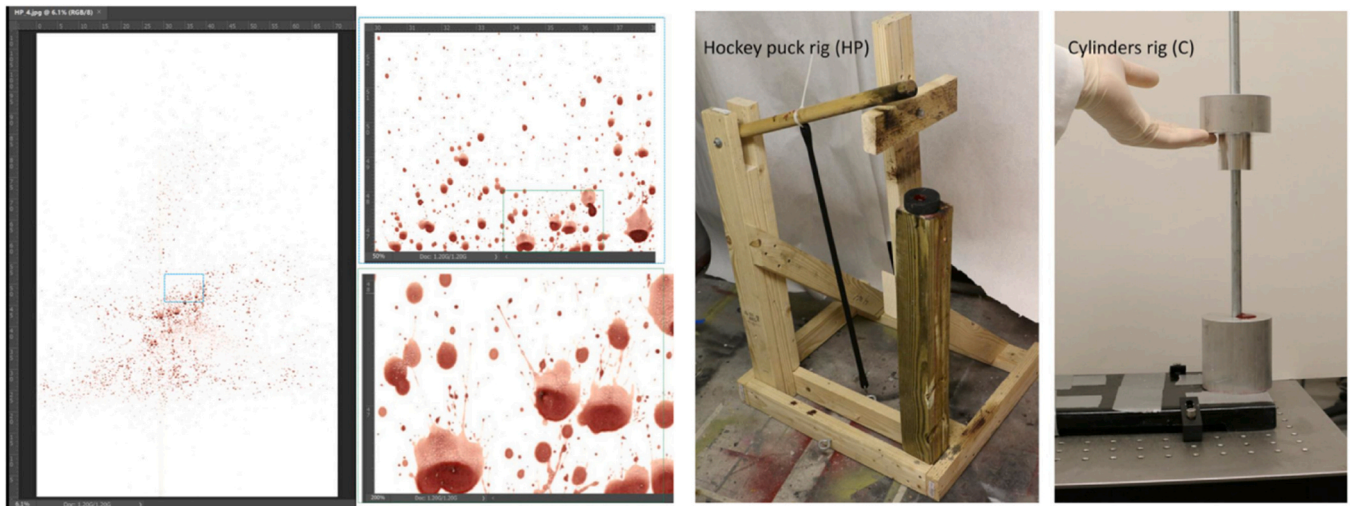
Two apparatuses are used to create the impact patterns. In the first, a blood pool around 1 mL is hit by a metal cylinder weight released from different heights above the blood. This simulates stepping into puddles of blood. In the second apparatus, the blood pool is placed on a hockey puck and hit by a thin wooden rod with different velocities, imitating beating incidents with a blunt weapon. The horizontal distance between the blood source and target wall is varied between 30 and 200 cm. Fig. 1 shows a sample impact pattern and pictures of the two apparatuses.

The gunshot patterns are all back spatters where bloodstains are splashed in the opposite direction to that of the bullet. The source is a piece of foam filled with blood. It is then shot by a bullet traveling horizontally, creating a pattern on the target wall between the source and the gun. To add variability, handguns and rifles with different bullet shapes (pointy, round, and flat) are used. The horizontal distance between the target wall and blood source is varied between 10 and 120 cm. Fig. 2 shows a sample gunshot pattern and a picture of the laboratory setting.
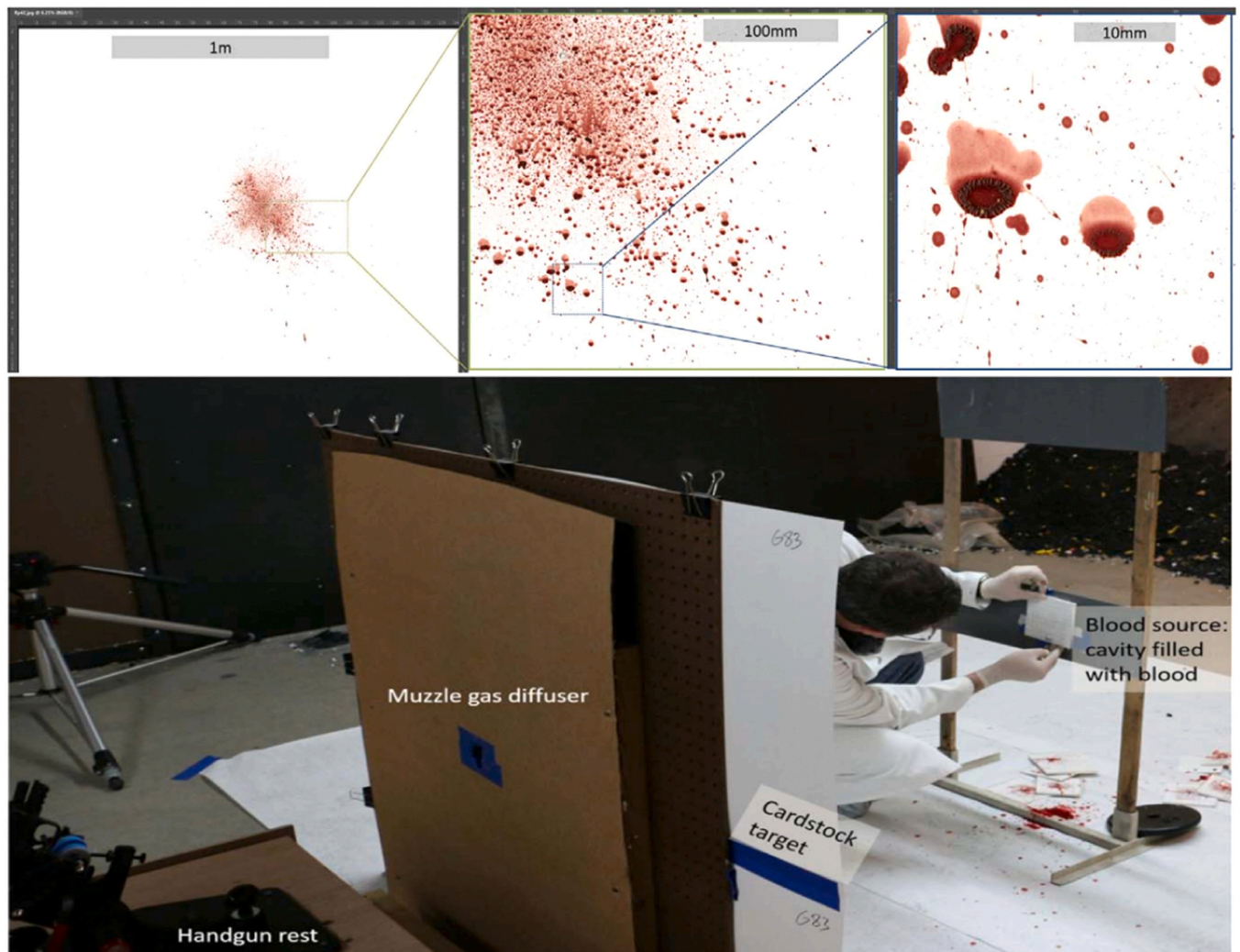
2.2. Image processing

We assume the evidence E in formula (1) in a bloodstain pattern analysis is an image of the crime scene. Directly modeling the distribution of a bloodstain pattern with a limited number of samples is beyond the bounds of possibility due to the high-dimensionality of image data. For impact and gunshot patterns, many of the bloodstains in a pattern are in the shape of ellipses. Considering a pattern as a collection of ellipses can greatly simplify the data structure and retain most of the information in the pattern. Thus, the first step in our analysis is processing the images to obtain an ellipse representation for each bloodstain pattern.

The JPEG images are imported in MATLAB R2021 [28] and pre-processed using the DIPImage Toolbox v2.9 [14]. We follow the procedures described in the work of [5] to transform the original RGB image to a smoothed binary image. The transformation is carried out through a series of steps: *background subtraction*, *element segmentation* and *morphological operations*. First, the image is converted to grayscale and an inferred background is subtracted out. The background image is estimated through downsampling of the original image, median filtering, and upsampling to the original size. Then, the image is converted to black and white (the segmentation step) using the Triangle thresholding algorithm [35]. Lastly, binary erosion and dilation each with four iterations are applied to eliminate image noise and remove any tail pixels associated with individual bloodstains.

Fig. 1. Images taken from [8]. On the left is an impact pattern on a 70 × 110 cm board with different scales of insets. On the right are the two apparatuses used to create impact patterns.

Fig. 2. Images taken from [9]. On the top is a gunshot pattern on a 140 × 110 cm cardstock board with different scales of insets. On the bottom is the experiment preparation of a backspatter with a handgun.
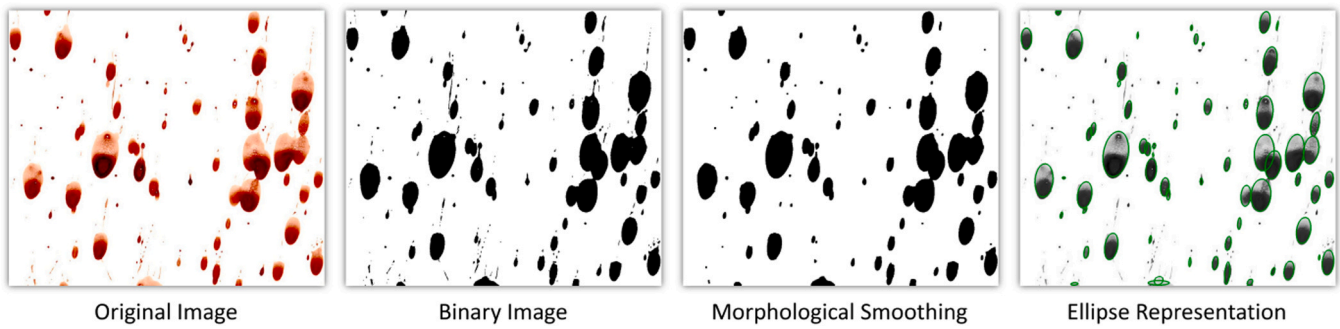
Fig. 3. An outline of image processing. The binary image is obtained after thresholding. The morphological smoothed image is obtained after binary erosion and dilation. The ellipse representation is produced by the ellipse recognition algorithm.

After those steps, Arthur et al. use the *regionprops* function (in MATLAB's Image Processing Toolbox) to label and measure every region of the binary image. In particular, each region is fitted by an ellipse to approximate the shape of a blood droplet, and parameters of the ellipse are later used to define features. A limitation of this method is that *regionprops* only fits each region with one ellipse while regions that may be composed of multiple droplets are discarded, which inevitably results in a loss of information. Instead, we use the algorithm developed in [36] that can approximate a non-elliptical region with a set of overlapping ellipses. The idea of the algorithm is first to generate a pool of candidate ellipses through a series of distance transforms of the region, then partition the region contour into several segments defined by its concave points, and finally match each segment with an ellipse from the candidate pool. Regions that are purely noise or poorly approximated by ellipses are excluded using the following approach. Two metrics, the Jaccard Index [20] and Hausdorff distance [19], are computed to assess the fit between each region and its approximating ellipses (or a single ellipse). A region is discarded if either of the metrics exceeds a specified threshold (0.9 for Jaccard Index and 5 for Hausdorff distance). An example of image processing is provided in Fig. 3. As a final step, any bloodstain pattern that is represented by five ellipses or fewer is discarded as not having sufficient data. The processing results in 59 impact patterns and 52 gunshot patterns (i.e., 3 gunshot pattern images are removed).

Now that bloodstains are approximated by a collection of ellipses, we use the parameters of these ellipses as a quantitative representation of a pattern. An ellipse can be characterized by five parameters (x, y, a, b, ϕ), where (x, y) denote the center coordinate, a and b the radii of major and minor axes, and ϕ the angle between the x-axis and the major axis of the ellipse (see Fig. 4). A pattern is represented by an n by 5 table where n is the number of approximating ellipses and each row contains the ellipse parameters.

3. Method

Given the available data introduced in Section 2.1, we specify the two competing hypotheses regarding a bloodstain pattern to match the two mechanisms of the datasets. H_1 and H_2 in formula (1) now propose that the bloodstain pattern is produced by gunshot and impact respectively as follows:

H_1: *The pattern is produced by gunshot*
H_2: *The pattern is produced by impact*

In Section 2.2 we transform a bloodstain pattern image into a collection of ellipses, or equivalently, an n by 5 table filled with values of ellipse parameters. Despite the great reduction in data complexity, the data table is still in a high-dimensional space and modeling the likelihood is not straightforward with a limited number of samples. We further simplify the problem by modeling the likelihood of a number of features $\boldsymbol{f}$ that are summaries of the ellipse representation. As a result, the likelihood ratio of a given pattern is defined by the following:

$$LR_{BPA} = \frac{P(\boldsymbol{f}|H_1)}{P(\boldsymbol{f}|H_2)} \qquad (2)$$

It is worth mentioning that our goal in this study is to illustrate how the likelihood ratio approach can be used to assess the relative support for competing hypotheses about the mechanism that produced a given bloodstain pattern. The limited data available and simplified model indicate that this is a demonstration of the approach rather than a definitive LR calculation.

3.1. Feature design

Previous studies [4,25] have proposed a number of quantitative summary features useful for classifying patterns from different mechanisms. Most of the features rely on the ellipse representation. Examples include the average area of ellipses, the average ratio of the lengths of the major and minor axes and the number of ellipses. Both those studies identify the orientation of bloodstains as an important characteristic of a pattern and therefore proposed descriptive statistics sensitive to orientation as features. For instance, bloodstains in cast-off patterns tend to have more consistent orientation than those in impact patterns. This suggests that the variance of the slopes of the ellipses in an impact pattern is likely to be larger than that of a cast-off pattern. However, as discussed in the following section, because of the inherent property of periodicity of orientation variables, it is not appropriate to use the traditional definition of variance in statistics. Instead, we applied directional statistics to solve the problem of periodicity and designed new
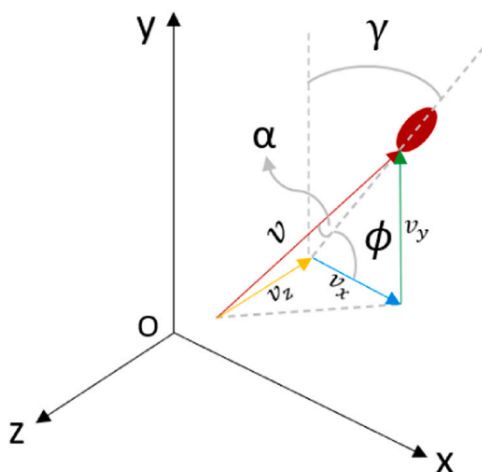

Fig. 4. Coordinate system and orientation angles associated with a blood droplet.

features that exploit the expected importance of orientation of bloodstains.

3.1.1. Circular features

Two commonly used measures to gauge the orientation of an individual bloodstain are the gamma angle and the impact angle. As shown in Fig. 4, the gamma angle, γ, is the angle between the blood droplet path projected on the target surface ($Z = 0$) and the plumb (vertical) line. The impact angle α is the angle between the blood droplet path and the target surface, With γ and α, one can deduce the direction of velocity of a blood droplet when it hits the surface. In terms of the ellipse representation, one can estimate the impact angle using the formula:

$$\alpha = \arcsin \frac{b}{a} \qquad (3)$$

As for the gamma angle, it seems intuitive to express γ via ellipse slope ϕ since they are complementary angles. However, due to the fact that an ellipse is symmetric in both of its major and minor axes, ϕ cannot determine the head and tail of a blood droplet as it ranges from 0 to 180 degrees whereas γ can range from 0 to 360 degrees. Arthur et al. [5] proposed comparing the centroids of a bloodstain before and after the morphological operations to determine its directionality (head and tail). This method is applicable only to single-droplet stains with a connected tail, while the directionality of stains without a noticeable tail or overlapped with multiple droplets are difficult to infer. In this study, we choose to use ellipse slope ϕ in place of the gamma angle in order to include more bloodstains into the analysis.

Summary statistics like the variance of the distribution of ϕ and α have previously been used to characterize the orientation of ellipses [4,25]. However, traditional statistics of angular variables actually fail to properly describe the population. For example, the simple average of a 2 degree angle and a 358 degree angle is $(2 + 358)/2 = 180$ degrees, which is far from the more sensible answer 0 degree. In principle, statistics should be invariant and consistent to coordinates change. But the average value of angles depends on the reference line that defines 0 degree and whether the angle ranges from 0 to 360 degrees or -180–180 degrees. This is because conventional statistics assumes all variables belong to the Euclidean space where a straight line never forms into a loop, but angular variables lie in a circular space. The subject that studies the statistical behavior of such variables is called directional statistics [26]. A common approach is to map an angular variable β ranging from 0 to 360 degrees into 2-dimensional space via polar coordinates:

$$\beta \in [0, 360) \mapsto \boldsymbol{u} = (\cos \beta, \sin \beta) \in \mathbb{R}^2 \qquad (4)$$

Now the angular variables are distributed on a 2-dimensional unit circle. To obtain the average of a group of angles, we can instead average the corresponding vectors and compute the angle of the centroid vector as the angle average:

$$\bar{\boldsymbol{u}} = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{u}_i = \left(\frac{1}{n} \sum_{i=1}^{n} \cos \beta_i \quad, \quad \frac{1}{n} \sum_{i=1}^{n} \sin \beta_i \right) \qquad (5)$$

$$\bar{\beta} = \text{angle between } \bar{\boldsymbol{u}} \text{ and vector } (1, 0) \qquad (6)$$

The length of $\bar{\boldsymbol{u}}$ is called the *mean resultant length* and denoted as $\bar{R}$. The variance of β is then defined as one minus the mean resultant length:

$$Var[\beta] = 1 - \|\bar{\boldsymbol{u}}\| \qquad (7)$$

$$= 1 - \bar{R} \qquad (8)$$

$$= 1 - \sqrt{\left(\frac{1}{n} \sum_{i=1}^{n} \cos \beta_i \right)^2 + \left(\frac{1}{n} \sum_{i=1}^{n} \sin \beta_i \right)^2} \qquad (9)$$

It is obvious that $\bar{R}$ is always less than one, so the variance of an angular variable ranges from zero to one. Considering the extreme cases makes it easier to understand why the variance is defined in this manner. If all β_i's have the same value, which means there is no variation, $\bar{R}$ will be exactly one and $Var[\beta]$ will be zero. On the other hand, if β_i's are uniformly distributed between 0 and 360, which means maximum variation, $\bar{\boldsymbol{u}}$ will be averaged into zero vector, leading $Var[\beta]$ to its maximum one. $\bar{R}$ and $Var[\beta]$ reflects the degree of concentration and dispersion of angles respectively.

The variance of the impact angle α can be computed using formula (9). In terms of the slope angle ϕ for which 0 degree is the same as 180 degrees, it is a standard procedure to first multiply ϕ by a factor of two so that it ranges from 0 to 360 degrees to conform with the expected range for arguments of trigonometric functions. Formulas used to calculating the variances of the two angles are given in Table 2.

3.1.2. Spherical features

Because the direction of incident velocity of a blood droplet can be deduced given α and γ (or its complement ϕ), features defined in terms of those variables (e.g. $Var[\phi]$ and $Var[\alpha]$) can summarize the distribution of incident direction of blood droplets to some extent. A limitation of this approach is that information about the correlation of these two angles is not incorporated.

Here we describe an approach based on viewing the data through vectors of incident direction. This will allow us to incorporate information about the correlation. Note that we can derive the vector of incident direction in three dimensions from γ and α:

$$\boldsymbol{m} = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} = (-\cos \alpha \cos \gamma, -\cos \alpha \sin \gamma, -\sin \alpha) \qquad (10)$$

where $\boldsymbol{v}$ is the incident velocity (Fig. 4) and $\boldsymbol{m}$ is a unit vector in the same direction as $\boldsymbol{v}$. Since the gamma angle has been replaced by ϕ as described in the previous section, we use the following instead:

$$\tilde{\boldsymbol{m}} = (-\cos \alpha \cos 2\phi, -\cos \alpha \sin 2\phi, -\sin \alpha) \qquad (11)$$

where the factor of 2 before ϕ is to maintain its circular structure as explained above. The vector $\tilde{\boldsymbol{m}}$ does not necessarily equal $\boldsymbol{m}$ due to the use of ϕ. But the distribution of $\tilde{\boldsymbol{m}}$ still precisely characterizes the variation of incident directions of blood droplets in the case that their head and tail directions are unknown.

Although $\tilde{\boldsymbol{m}}$ is a 3-dimensional vector, by its construction from the two angles, it is constrained to lie on the surface of a hemisphere with unit radius centered at the origin. Due to the non-Euclidean nature of the spherical surface, conventional statistical methods fail to function properly. The strategies we used in the previous section to handle angular variables are applicable to spherical variables. For example, the mean resultant length of $\tilde{\boldsymbol{m}}$, defined as:

$$\tilde{R} = \left\| \frac{1}{n} \sum_{i=1}^{n} \tilde{\boldsymbol{m}}_i \right\| \qquad (12)$$

is still a good measure of concentration. But since $\tilde{\boldsymbol{m}}$ is 3-dimensional, a more informative and accurate alternative is to compute the scatter matrix $\boldsymbol{T}$:

$$\boldsymbol{T} = \sum_{i=1}^{n} \tilde{\boldsymbol{m}}_i \tilde{\boldsymbol{m}}_i^T \qquad (13)$$

In mechanics, $\boldsymbol{T}$ is interpreted as the inertia tensor of a rigid body about the origin, where the body is composed of equal weight particles at each of the locations $\tilde{\boldsymbol{m}}_1, ..., \tilde{\boldsymbol{m}}_n$. In statistics, $\boldsymbol{T}$ is analogous to the covariance matrix that provides information regarding variances and correlations of data. One possible summary to capture the dispersion of $\boldsymbol{T}$ is its determinant $\det(\boldsymbol{T})$, which in fact relates to the data entropy assuming the distribution is multivariate Gaussian [1]. Another way to more specifically characterize the shape of the distribution of incident directions is via an eigendecomposition of $\boldsymbol{T}$.

Table 1

Descriptive interpretation of shapes of spherical distribution based on the eigendecomposition of $\boldsymbol{T}$ [26].

Relative magnitudes of eigenvalues	Shape of distribution	Other features
$t^1 \approx t^2 \approx t^3$	uniform	
t^1 large; t^2, t^3 small		if $t^2 \approx t^3$, rotational symmetry about t^1 otherwise, concentrated at t^1 (and $-t^1$ for bimodal)
(i) $\bar{R} \approx 1$	unimodal	
(ii) $\bar{R} < 1$	bimodal	
t^1, t^2 large; t^3 small	girdle	if $t^2 \approx t^3$, rotational symmetry about t^3 otherwise, concentrated about circle in plane of t^1, t^2
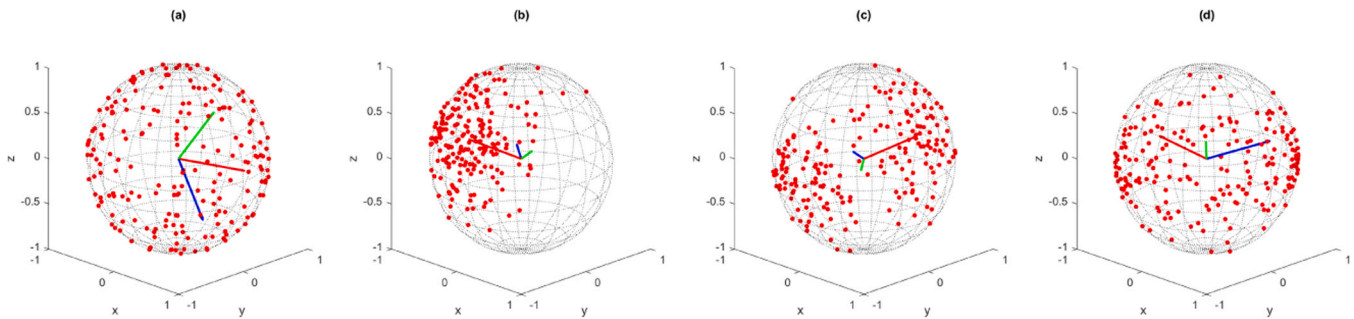
Fig. 5. Examples of spherical distributions described in Table 1. (a) Uniform. (b) Unimodal. (c) Bimodal. (d) Girdle. Data points (red dots) are scattered on the sphere surface. In each sphere there are three lines of red, blue and green indicating the directions of eigenvectors $\boldsymbol{t}^1$, $\boldsymbol{t}^2$, $\boldsymbol{t}^3$ respectively. Their lengths are proportional to t^1, t^2, t^3 and scaled so that the red line has the same length as radius.

The eigendecomposition of the covariance matrix provides a set of orthogonal principal components that are often used to extract features. Indeed, as pointed out in [26], the eigenvalues t^1, t^2, t^3 (in decreasing order) and eigenvectors $\boldsymbol{t}^1$, $\boldsymbol{t}^2$, $\boldsymbol{t}^3$ of $\boldsymbol{T}$ indicate the general shape of the data as described in Table 1 and illustrated in Fig. 5. The eigenvalues of the scatter matrix are great candidates to serve as features for bloodstain patterns due to their mutual independence and summarization of the distribution of incident directions. The distribution of $\tilde{\boldsymbol{m}}$ will most likely fall into the unimodal category in Table 1, because $\tilde{\boldsymbol{m}}$ is constrained on only half of the sphere and categories with other shapes of distribution involve the whole sphere. Thus, t^1 is expected to be much greater than t^2 and t^3. The ratio between t^2 and t^3 indicates the symmetry of the distribution and can serve as an appropriate summary.

3.2. Likelihood evaluation

The LR framework involves computing the likelihoods of a bloodstain pattern under different assumptions about the mechanism. As described earlier, directly modeling the likelihood of a bloodstain pattern is extremely difficult. To demonstrate the approach, we instead model the joint likelihood of features extracted from the pattern. Theoretically, patterns that are caused by different mechanisms become more distinguishable if more features are used in the model. But this makes the result more difficult to interpret and harder to visualize. As a compromise, we model the joint distribution of two selected features f_1 and f_2 with a bivariate Gaussian distribution under H_1 and under H_2:

$$\begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \Big| H_1 \sim N_2(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1), \quad \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \Big| H_2 \sim N_2(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2) \tag{14}$$

where the parameters $\boldsymbol{\mu}_1$, $\boldsymbol{\Sigma}_1$ are the mean and covariance of the distribution of features extracted from gunshot patterns, and $\boldsymbol{\mu}_2$, $\boldsymbol{\Sigma}_2$ are those parameters for features extracted from impact patterns. The choice of the Gaussian distribution may seem like a strong assumption, but works well for transformations of the features in our example as described in Section 4.1. Mixture models and kernel density estimation are alternative approaches when the distribution of features is complex and hard to transform into a Gaussian

distribution [22,24]. The LR of a bloodstain pattern is evaluated by first extracting its pair of features $\boldsymbol{f} = (f_1, f_2)^T$ and then computing the Gaussian density functions under the competing hypotheses:

$$LR_{BPA} \approx \frac{p_{N_2}(\boldsymbol{f}|\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)}{p_{N_2}(\boldsymbol{f}|\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)} \tag{15}$$

where $p_{N_2}(\boldsymbol{f}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the density function of a bivariate Gaussian distribution. Parameters of the density functions are estimated by maximum likelihood on a training data set. For example, parameters of the model under H_1 are estimated as follows:

$$\hat{\boldsymbol{\mu}}_1 = \frac{1}{N} \sum_{j=1}^{N} \boldsymbol{f}_j \tag{16}$$

$$\hat{\boldsymbol{\Sigma}}_1 = \frac{1}{N} \sum_{j=1}^{N} (\boldsymbol{f}_j - \hat{\boldsymbol{\mu}}_p)(\boldsymbol{f}_j - \hat{\boldsymbol{\mu}}_p)^T \tag{17}$$

where $\boldsymbol{f}_j$'s are features extracted from patterns that are produced by gunshot in our data set, and N is the number of gunshot patterns. Parameters of the model under H_2 are estimated similarly using impact patterns. Note here we use pattern index j and number of patterns N to differentiate from the ellipse index i and number of ellipses n within a pattern.

The likelihood ratio that is developed here is very similar to that introduced in the context of handwriting by [15]. They identify two features as summaries of handwritten signatures and apply a bivariate normal probability distribution to demonstrate the use of likelihood ratios to assess the strength of evidence regarding potential authors. There is a relationship with the likelihood ratio for assessing trace evidence presented by [3] but there is an important distinction. The data in that context comprise two separate evidence samples, one from the crime scene and one from the suspect, and the LR is assessing the strength of the evidence for same source versus different source. The LR in that case must account for the fact that the means characterizing the samples are drawn from a population of possible sources and thus the LR formula in [3] includes integration over that population distribution. The BPA LR and the handwriting LR of [15] concern a single sample (bloodstain pattern or signature) and assess the evidence contained in that single sample with respect to the two hypotheses of interest.

In practice, the model requires a much larger data set with a representative distribution of bloodstain patterns from each

Table 2
Characteristic features designed to represent bloodstain patterns.

Feature	Formula	Description	Range	Transform
Circular Features	$Var[\phi] = 1 - \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}\cos 2\phi_i\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n}\sin 2\phi_i\right)^2}$	variance of gamma angle	$(0, 1)$	logit
	$Var[\alpha] = 1 - \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}\cos \alpha_i\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n}\sin \alpha_i\right)^2}$	variance of impact angle	$(0, 1)$	logit
Spherical features	t^3/t^2	symmetry of incident directions	$(0, 1)$	logit
	$\det(\boldsymbol{T}) = t^1 \times t^2 \times t^3$	dispersion of incident directions	$(0, \infty)$	log

hypothesized mechanism to accurately estimate the parameters. Despite the limitation of the simplified Gaussian model and the small data set used here, our work still shows promising results in the following section. Consequently, the simplified approach helps to demonstrate the essence of the LR framework and the potential for its implementation in BPA.

4. Experiments

The objective of this section is to illustrate the LR approach presented in Section 3 using the data introduced in Section 2.

4.1. Feature preprocessing

Features derived in Section 3.1 are grouped into two pairs: (1) the circular features are $Var[\phi]$ and $Var[\alpha]$, the variances of two orientation angles; (2) the spherical features are the ratio between the third and second largest eigenvalues of the scatter matrix $\boldsymbol{T}$ (which measures symmetry) and the determinant of $\boldsymbol{T}$ (which measures dispersion). Descriptions and formulas to compute these features are give in Table 2.

Our bivariate Gaussian model assumes the data distribution for the features is unbounded and not skewed. Therefore, instead of directly using these features in the model, we first map them into the real line $\mathbb{R}$ via transformations that should better approximate the assumed Gaussian distribution (see Table 2). For example, $Var[\phi]$ ranging from 0 to 1 can be transformed to $\mathbb{R}$ by the logit function:

$$\mathrm{logit}(x) = \log\left(\frac{x}{1-x}\right) \tag{18}$$

The transformed features extracted from all bloodstain patterns are plotted in Fig. 6. The confidence ellipse defines the smallest region that contains 95% of the points for each cluster. Due to transformations of the features, the shape of red and blue clusters in both

plots roughly resemble bivariate Gaussian distributions. It is worth noting that the cluster of spherical features of gunshot patterns seems to consist of two subclusters, which means a mixture model might fit the data better. But here we maintain the Gaussian model for simplicity. Gunshot patterns and impact patterns are not perfectly separated in either plot, though the degree of overlap between the two types of pattern (in red and blue colors) is obviously less for the spherical features than for the circular features.

On the left plot showing the circular features, we observe that gunshot patterns on average have smaller $Var[\alpha]$ and larger $Var[\phi]$. This indicates that in gunshot patterns, the impact angles of bloodstains are more converged and the slopes of ellipses are more dispersed than those in impact patterns. On the right plot showing the spherical features, we observe that incident directions are less disperse and more symmetric in gunshot patterns. Those observations may be explained by the fact that the average velocity of blood droplets produced by gunshot is higher than for those produced by impact.

4.2. Model evaluation

To illustrate the LR calculation in a way that provides for a realistic evaluation of the results, we perform a leave-one-out cross validation. First, we select one bloodstain pattern from the data set as the test pattern. The rest of the patterns are used to estimate the parameters of the models using equations (16) and (17). Finally the selected test pattern is evaluated via the LR defined by formula (15) with the estimated parameters to produce a LR value. This procedure is iterated until all patterns have been selected as the test pattern, thereby yielding $N = 111$ LR values. We expect gunshot patterns to produce LR values greater than 1, and impact patterns to have LR values less than one.

Fig. 7 presents the LR values of patterns in color (actually the logarithm of the LR values) and highlights the cases for which the LR
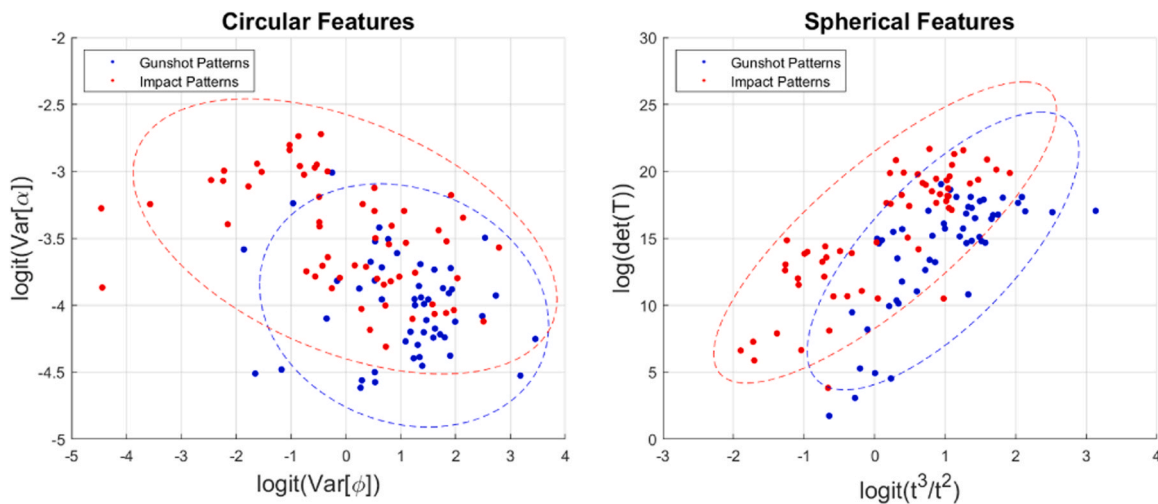
Fig. 6. Scatter plot of the transformed features. On the left are the circular pair of features for the patterns. On the right are the spherical pair of features for the patterns. A 95% confidence ellipse is computed for each cluster.
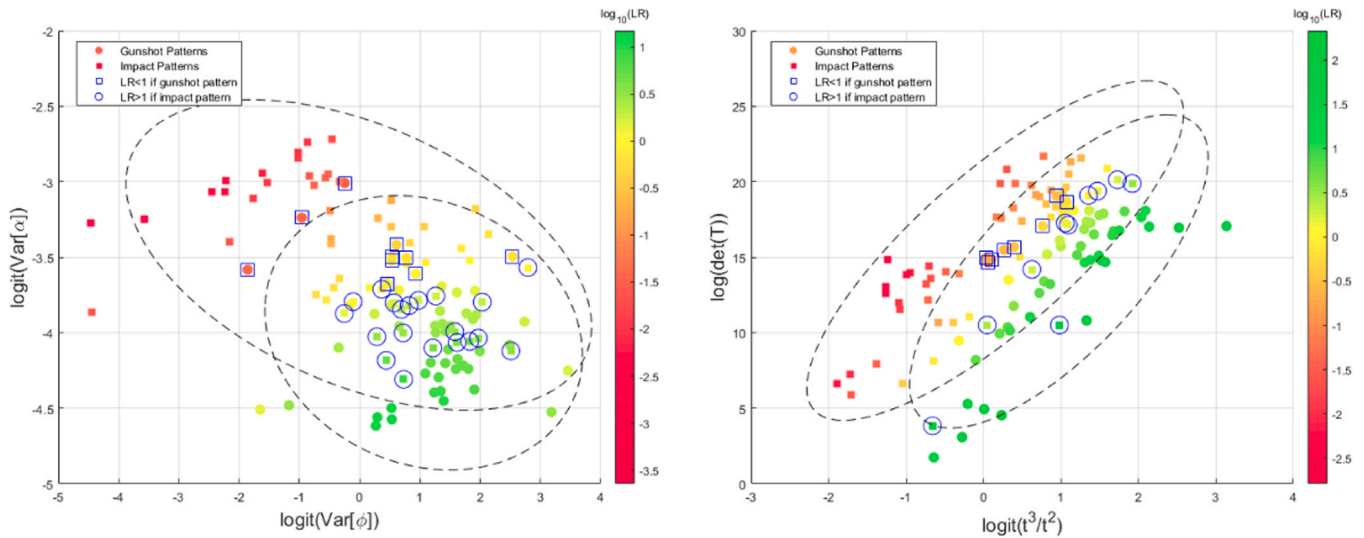
Fig. 7. Scatter plot of the circular features (left) and spherical features (right). The 95% of confidence ellipses for impact and gunshot patterns are depicted in dash lines. Solid circles and squares represent gunshot and impact patterns respectively. The values of the log of the LR obtained from cross validation are expressed in color. Hollow circles and squares are superimposed on patterns where the LR supports the wrong hypothesis.

Table 3
Confusion matrices with $LR = 1$ as threshold.

Ground Truth	Circular Features		Spherical Features	
	$LR > 1$	$LR < 1$	$LR > 1$	$LR < 1$
H_1	42	10	44	8
H_2	20	39	10	49

supports the wrong hypothesis. Points distant from the boundary between squares and circles have large absolute value of the log of LR. The LR of some patterns is greater than 100 or less than 0.01. Table 3 reports the classification for all patterns using 1 as a threshold for assessing the LR. For both circular and spherical features, impact patterns being misclassified as gunshot patterns occurs more frequently than the reverse. Error rates for circular features and spherical features are respectively 27% and 16%. LR values near one indicate that the data are equally consistent with both hypotheses. Patterns with such values may contributed to these high error rates when we classify using a threshold of one. To examine whether the results are sensitive to the choice of the LR threshold, Table 4 gives the numbers of patterns with LR inside and outside the intermediate range from 0.5 to 2. There are more patterns assigned to the intermediate range for the model using circular features than for the model using spherical features. This indicates that the model of circular features yields LRs more concentrated around one and therefore more sensitive to the choice of LR threshold. As expected, the cases with $0.5 < LR < 2$ are evenly divided among the two mechanisms. These cases account for a number of errors in each analysis.

Morrison et al. [29] suggest the use of a Tippett plot to diagnose bias of the LR model. The Tippett plot portrays empirical cumulative probability distributions of LR under H_1 and H_2. In general, greater separation and gentler slopes of the two curves indicate better performance. We created the Tippett plot for both circular and

spherical features, see Fig. 8. The blue line gives the proportion of gunshot patterns (true H_1) with $log_{10}(LR)$ less than or equal to the corresponding value on the x axis. The red line gives the proportion of impact patterns (true H_2) with $log_{10}(LR)$ greater than or equal to the corresponding value on the x axis. The Tippett plot of circular features exhibits a slight shift to the right where the two curves intersect, meaning that impact patterns are more prone to misclassification than gunshot patterns with threshold $LR = 1$. The plot of spherical features has a flatter curve slope and almost no shift. Alternatively, the results can be visualized using the receiver operating characteristic (ROC) curve that treats the likelihood ratios as scores from a classifier and considers all possible decision thresholds. Fig. 9 shows the ROC curve for models based on the circular features and spherical features. The areas under the curve for the two models are 0.795 (circular) and 0.893 (spherical). All of these approaches indicate that the spherical features demonstrate better performance than the circular features, indicating that the distribution of incident directions is more discriminating than the distribution of orientation angles.

4.3. Misclassified patterns

Analyzing patterns with LR values that are not consistent with their true mechanism can help us better understand and improve the model. There are two factors that were systematically controlled when the data were generated [8,9]: (a) the distance from the blood source to the target board; (b) the initial velocity of the blood droplets. The spherical features show better performance than the circular features in the sense of having fewer misclassified patterns, so we choose to examine those patterns with respect to the two factors mentioned above. Fig. 10 shows the factor level for each pattern and identifies the misclassified ones with a red cross according to LR values assigned by the model of spherical features. We observe that the majority of the misclassified impact patterns (8 out of 10)
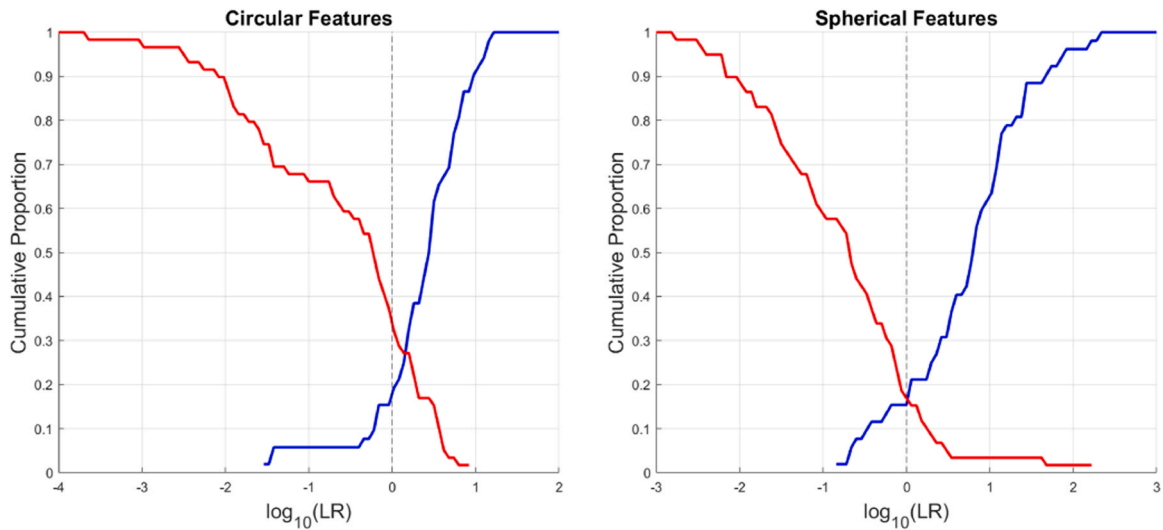
Table 4
Confusion matrices with $0.5 < LR < 2$ as an intermediate zone.

Ground Truth	Circular Features			Spherical Features		
	$LR > 2$	$0.5 < LR < 2$	$LR < 0.5$	$LR > 2$	$0.5 < LR < 2$	$LR < 0.5$
H_1	32	16	4	39	7	6
H_2	10	22	27	5	15	39

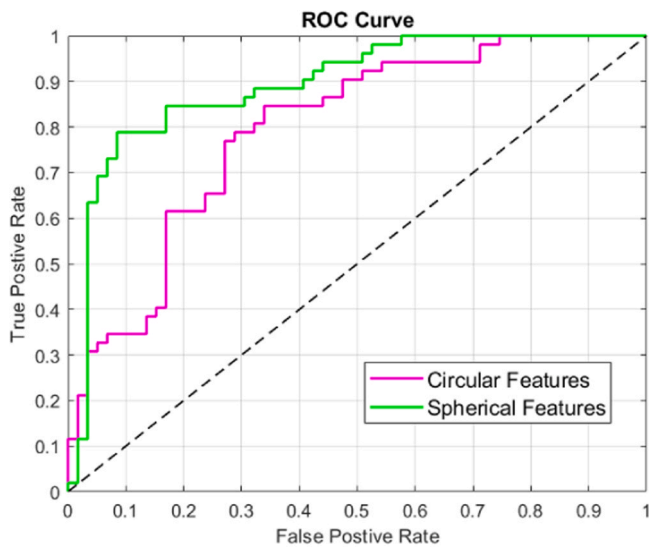**Fig. 8.** Tippett plots for circular and spherical features.

Fig. 9. ROC curves for the likelihood ratios obtained from models with circular features (purple) and spherical features (green).

occurred at the higher velocity levels, while the majority of the misclassified gunshot patterns (5 out of 8) occurred at the lower velocity level. This result supports our conjecture in Section 4.1 that the observation of less dispersion of the impact angle and incident direction for the gunshot patterns is due to the greater velocity of

blood droplets. Gunshot patterns with low velocity and impact patterns with high velocity may possess similar characteristics in terms of bloodstain orientation and hence are hard to distinguish for the model.

5. Discussion

A likelihood ratio approach to evaluating bloodstain pattern evidence is presented as a demonstration of the potential for this approach. Building on the work of [7], we address the problem of identifying the bloodletting mechanism. Based on two published data sets [8,9], statistical models of novel features are built to estimate the likelihood ratio of a bloodstain pattern to assess the relative support for the two possible mechanisms. Our work serves as a demonstration of the approach rather than a definitive likelihood ratio calculation as there are multiple ways in which it can be improved.

One limitation of the work is the size and variation of the data set we used. All the patterns were created under similar laboratory conditions. Although some experimental settings like the velocity of the blood droplets were varied, under each setting only a few patterns were generated. Therefore, we believe the data is not very representative of crime scenes in the real world.

Another limitation is that we used a relatively simple model to focus on the application of likelihood ratio framework. Using features to represent bloodstain patterns inevitably yields a loss of information. Incorporating more features to help separate patterns of different mechanisms is one approach to improvement. But this
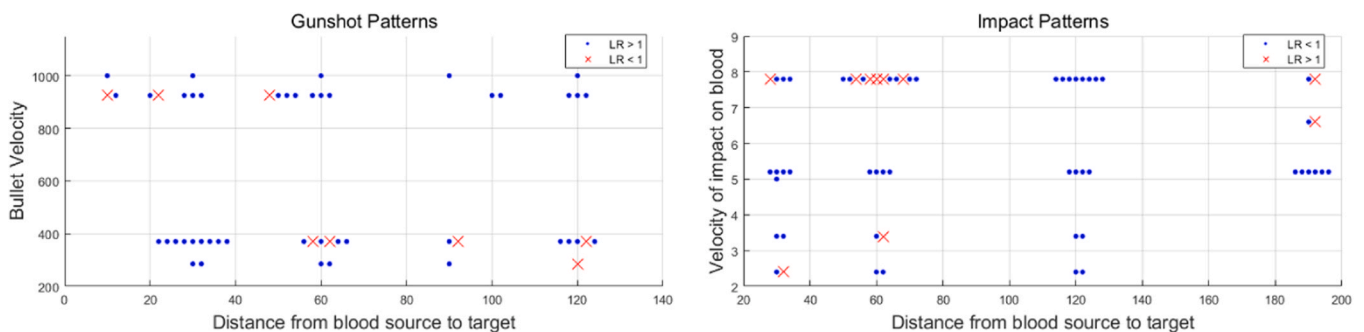

Fig. 10. Scatter plot of gunshot (left) and impact (right) patterns in terms of their conditions under which they were generated. The horizontal axis denotes the distance from the blood source to the target board. The vertical axis denotes different velocity levels of impact on blood when creating the pattern. Patterns where the LR obtained from the model of spherical features supports the wrong hypothesis are marked by red cross.

requires clever design of characteristic features and it becomes more difficult as the number of possible mechanisms goes up. In addition, the assumption of a Gaussian distribution was driven primarily by convenience for this conceptual demonstration. A mixture model or kernel-based model [22,24] can be used to more flexibly describe complex data distributions. Of course, a more flexible model with high-dimensional input space can easily lead to overfitting and produce extreme values of the likelihood ratio. Even with our relatively simple model, the absolute value of the logarithm of the likelihood ratio can be large. Possible solutions to reduce overfitting include parameter regularization, hierarchical modeling, and collecting more data.

Despite the limitations, the results show the potential for application of the likelihood ratio approach in bloodstain pattern analysis. In future study we will focus on collecting more representative data and building more flexible statistical models to better estimate the likelihood of bloodstain patterns.

CRediT authorship contribution statement

Tong Zou: Methodology, Software, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Hal Stern**: Conceptualization, Methodology, Formal analysis, Writing – review & editing, Supervision, Funding acquisition.

Conflict of interest

None.

Acknowledgements

Appendix A

For cases where there are more than two possible hypotheses regarding the bloodstain pattern evidence E, formula (1) can be generalized to evaluate the strength of the evidence supporting one hypothesis against the others. Assume there are K mutually exclusive hypotheses $H_1, H_2, \ldots, H_K$ proposing different possible mechanisms for E. Without loss of generality, let H_1 be the main hypothesis to be considered against the general alternative "some other mechanisms", then the LR can be written as:

$$LR = \frac{P(E|H_1)}{P(E|H_2 \text{ or } M_3 \text{ or} \ldots \text{or } H_K)} = \frac{P(E|H_1)}{P(E| \cup_{i=2}^K H_i)} \tag{19}$$

According to the definition of conditional probability, the denominator can be organized as:

$$P(E| \cup_{i=2}^K H_i) = \frac{P(E \cap (\cup_{i=2}^K H_i))}{P(\cup_{i=2}^K H_i)} \tag{20}$$

$$= \frac{\sum_{i=2}^K P(E \cap H_i)}{\sum_{i=2}^K P(H_i)} \tag{21}$$

$$= \frac{\sum_{i=2}^K P(E|H_i)P(H_i)}{\sum_{i=2}^K P(H_i)} \tag{22}$$

$$= \sum_{i=2}^K w_i P(E|H_i) \tag{23}$$

where $w_i = P(H_i) / \sum_{i=2}^K P(H_i)$ are prior probability weights for the other hypotheses. This result shows that evaluating $P(E|H_i)$, the likelihood of a pattern under each hypotheses H_i, remains the key calculation that is required.

References

[1] N.A. Ahmed, D.V. Gokhale, Entropy expressions and their estimators for multivariate distributions, IEEE Trans. Inf. Theory 35 (1989) 688–692.
[2] C. Aitken, F. Taroni, Statistics and the Evaluation of Evidence for Forensic Scientists, John Wiley & Sons, 2004.
[3] C.G.G. Aitken, D. Lucy, Evaluation of trace evidence in the form of multivariate data, J. R. Stat. Soc. Ser. C Appl. Stat. 53 (2004) 109–122.
[4] R.M. Arthur, J. Hoogenboom, M. Baiker, M.C. Taylor, K.G. De Bruin, An automated approach to the classification of impact spatter and cast-off bloodstain patterns, Forensic Sci. Int. 289 (2018) 310–319.
[5] R.M. Arthur, P.J. Humburg, J. Hoogenboom, M. Baiker, M.C. Taylor, K.G. DeBruin, An image-processing methodology for extracting bloodstain pattern features, Forensic Sci. Int. 277 (2017) 122–132.
[6] D. Attinger, P.M. Comiskey, A.L. Yarin, K. De Brabanter, Determining the region of origin of blood spatter patterns considering fluid dynamics and statistical uncertainties, Forensic Sci. Int. 298 (2019) 323–331.
[7] D. Attinger, K. De Brabanter, C. Champod, Using the likelihood ratio in bloodstain pattern analysis, J. Forensic Sci. 67 (2022) 33–43.
[8] D. Attinger, Y. Liu, T. Bybee, K. De Brabanter, A data set of bloodstain patterns for teaching and research in bloodstain pattern analysis: impact beating spatters, Data Brief 18 (2018) 648–654.
[9] D. Attinger, Y. Liu, R. Faflak, Y. Rao, B.A. Struttman, K. De Brabanter, P.M. Comiskey, A.L. Yarin, A data set of bloodstain patterns for teaching and research in bloodstain pattern analysis: gunshot backspatters, Data Brief 22 (2019) 269–278.
[10] T. Bevel, R.M. Gardner, Bloodstain Pattern Analysis with an Introduction to Crime Scene Reconstruction, CRC Press, Boca Raton, FL, 2008.
[11] S. Bozza, F. Taroni, R. Marquis, M. Schmittbuhl, Probabilistic evaluation of handwriting evidence: likelihood ratio for authorship, J. R. Stat. Soc. Ser. C Appl. Stat. 57 (2008) 329–341.
[12] S. Bunch, G. Wevers, Application of likelihood ratios for firearm and toolmark analysis, Sci. Justice 53 (2013) 223–229.
[13] F. Camana, Determining the area of convergence in bloodstain pattern analysis: a probabilistic approach, Forensic Sci. Int. 231 (2013) 131–136.
[14] DIPimage, version 2.9, Delft University of Technology. ⟨https://www.diplib.org/⟩.
[15] L. Gaborini, A. Biedermann, F. Taroni, Towards a Bayesian evaluation of features in questioned handwriting signatures, Sci. Justice 57 (2017) 209–220.
[16] C. Galbraith, P. Smyth, Analyzing user-event data using score-based likelihood ratios with marked point processes, Digit. Investig. 22 (2017) S106–S114.
[17] C. Galbraith, P. Smyth, H.S. Stern, Quantifying the association between discrete event time series with applications to digital forensics, J. R. Stat. Soc. Ser. A Stat. Soc. 183 (2020) 1005–1027.
[18] R.A. Hicklin, K.R. Winer, P.E. Kish, C.L. Parks, W. Chapman, K. Dunagan, N. Richetelli, E.G. Epstein, M.A. Ausdemore, T.A. Busey, Accuracy and

reproducibility of conclusions by forensic bloodstain pattern analysts, Forensic Sci. Int. (2021).

[19] D.P. Huttenlocher, G.A. Klanderman, W.J. Rucklidge, Comparing images using the Hausdorff distance, IEEE Trans. Pattern Anal. Mach. Intell. 15 (1993) 850–863.

[20] P. Jaccard, The distribution of the flora in the alpine zone. 1, New Phytol. 11 (1912) 37–50.

[21] S.H. James, P.E. Kish, T.P. Sutton, Principles of Bloodstain Pattern Analysis: Theory and Practice, CRC Press, Boca Raton, FL, 2005.

[22] M.C. Jones, J.S. Marron, S.J. Sheather, A brief survey of bandwidth selection for density estimation, J. Am. Stat. Assoc. 91 (1996) 401–407.

[23] R.E. Kass, A.E. Raftery, Bayes factors, J. Am. Stat. Assoc. 90 (1995) 773–795.

[24] B.G. Lindsay, Mixture Models: Theory, Geometry, and Applications, Institute of Mathematical Statistics, 1995.

[25] Y. Liu, D. Attinger, K. De Brabanter, Automatic classification of bloodstain patterns caused by gunshot and blunt impact at various distances, J. Forensic Sci. 65 (2020) 729–743.

[26] K.V. Mardia, P.E. Jupp, Directional Statistics, John Wiley & Sons, 2009.

[27] R. Marquis, S. Bozza, M. Schmittbuhl, F. Taroni, Handwriting evidence evaluation based on the shape of characters: application of multivariate likelihood ratios, J. Forensic Sci. 56 (2011) S238–S242.

[28] MATLAB version 9.11.0.1769968 (R2021b), The MathWorks Inc., Natick, Massachusetts.

[29] G.S. Morrison, E. Enzinger, V. Hughes, M. Jessen, D. Meuwly, C. Neumann, S. Planting, W.C. Thompson, D. van der Vloed, R.J.F. Ypma, C. Zhang, et al., Consensus on validation of forensic voice comparison, Sci. Justice 61 (2021) 299–309.

[30] National Research Council, et al., Strengthening Forensic Science in the United States: A Path Forward, National Academies Press, 2009.

[31] C. Neumann, I.W. Evett, J. Skerrett, Quantifying the weight of evidence from a forensic fingerprint comparison: a new paradigm, J. R. Stat. Soc. Ser. A Stat. Soc. 175 (2012) 371–415.

[32] C.D. Steele, D.J. Balding, Statistical evaluation of forensic DNA profile evidence, Annu. Rev. Stat. Appl. 1 (2014) 361–384.

[33] H.S. Stern, Statistical issues in forensic science, Annu. Rev. Stat. Appl. 4 (2017) 225–244.

[34] S. Willis, L. McKenna, S. McDermott, G. O'Donell, A. Barrett, B. Rasmusson, A. Nordgaard, C. Berger, M. Sjerps, J. Lucena-Molina, et al., Strengthening the evaluation of forensic results across europe (steofrae), ENFSI Guideline for Evaluative Reporting in Forensic Science, 2015.

[35] G.W. Zack, W.E. Rogers, S.A. Latt, Automatic measurement of sister chromatid exchange frequency, J. Histochem. Cytochem. 25 (1977) 741–753.

[36] T. Zou, T. Pan, M. Taylor, H. Stern, Recognition of overlapping elliptical objects in a binary image, Pattern Anal. Appl. 24 (2021) 1193–1206.