

МИНОБРНАУКИ РОССИИ  
Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра материаловедения, технологии  
и управления качеством

**РАЗРАБОТКА НА ОСНОВЕ АЛГОРИТМА МАШИННОГО ОБУЧЕНИЯ  
«СЛУЧАЙНЫЙ ЛЕС» ИДЕНТИФИКАТОРА ОГНЕСТРЕЛЬНОГО  
ОРУЖИЯ ПО ВТОРИЧНЫМ СЛЕДАМ НА ПУЛЯХ, ПОКРЫТЫХ  
ТОМПАКОМ**

**АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ**

студента магистратуры 2 курса 2293 группы  
направления 22.04.01 «Материаловедение и технологии материалов»,  
профиль «Криминалистическое материаловедение»  
института физики

Мясникова Никиты Николаевича

Научный руководитель,  
доцент, к.ф.-м.н., доцент  
\_\_\_\_\_

должность, уч. степень, уч. звание

\_\_\_\_\_

подпись, дата

В.А. Федоренко  
\_\_\_\_\_

инициалы, фамилия

Зав. кафедрой,  
д.ф.-м.н., профессор  
\_\_\_\_\_

должность, уч. степень, уч. звание

\_\_\_\_\_

подпись, дата

С.Б. Вениг  
\_\_\_\_\_

инициалы, фамилия

Саратов 2023

**Введение.** Актуальность данной работы заключается в повышении эффективности расследования преступлений связанных с использованием огнестрельного оружия за счет разработки более совершенного метода сравнения изображений вторичных следов на выстреленных пулях, основанного на применении методов машинного обучения.

Целью работы является определение оптимальных критериев отнесения трасс в совмещенных динамических следах к категории «совпадающие», основанной на нахождении наименьшей энтропии Шеннона, а также применение метода «Случайный лес» для бинарной классификации следов на совпадающие и несовпадающие.

Задачами работы являются:

1. Проведение обзора научно-технической литературы и анализа современных методов сравнительного исследования вторичных следов.
2. Определение оптимального критерия совмещения совпадающих трасс в сравниваемых вторичных следах.
3. Исследование влияния механических характеристик оболочки пули на выраженность вторичных следов.
4. Анализ методов машинного обучения. Определение оптимального метода бинарной классификации пар совмещенных следов по 11 признакам.

Научная новизна работы заключается в разработке новой методики сравнения совмещенных динамических следов и проведения бинарной классификации, основанной на методе машинного обучения «Случайный лес».

Положения, выносимые на защиту:

1. Разработана и апробирована методика оценки количественного критерия отнесения трасс к категории «совпадающие» в совмещенных вторичных следах.
2. Показано, что метод «Случайный лес» может быть успешно применен для бинарной классификации в многомерном пространстве признаков

совмещенных вторичных следов по категориям «Следы совпадающие» и «Следы несовпадающие».

Предпосылками индивидуальной идентификации оружия, являются индивидуальность каждого экземпляра оружия, устойчивость его комплекса индивидуальных признаков и их стабильная отображаемость.

Одной из первых работ в которых были предприняты попытки оценки пригодности динамического следа к идентификации при проведении предварительного исследования была работа Г. Л. Грановского [1], в которой удачно использовались идеи комбинаторики.

Идея количественного обоснования внутреннего убеждения эксперта о совпадении совмещенных динамических следов также была представлена в работах В. А. Федоренко и Е. В. Сидак [2, 3]. В статье [3] описан метод позволяющий определить уникальность комплекса трасс в исследуемом бинаризованном следе с помощью автокорреляционной функции. Если функция по виду близка к дельта-функции, то комплекс трасс в следе образует уникальную совокупность признаков.

В работе А. Биазотти [4] продемонстрирован метод оценки схожести сравниваемых вторичных следов по наиболее длинным сериям последовательно совпадающих трасс. Чем длиннее серия, тем выше вероятность, что совмещенные следы сформированы одним полем нарезки.

В работе [5] представлена вероятностная модель вторичных следов, которая сводит проблему оценки уникальности следов на пулях к решению задачи по теории вероятности. В работе [6] она получает своё дальнейшее развитие. В статье [7] предложена методика оценки вероятности наилучшего совмещения всех пар вторичных следов на исследуемых пулях, а затем определения среднего значения от этих величин.

Однако, во всех перечисленных работах не отображена возможность работы с сильно деформированными пулями. В работах [8] и [9] описана методика, которая благодаря разбиению одного из сравниваемых следов на сегменты и

определению функций взаимной корреляции сегментов и второго целого следа позволяет определить, являются ли исследуемые следы парными (совпадающими) или нет.

Рассмотрев все предыдущие работы, видно, что в большей их части анализ сравниваемых следов проводится по одному-двум признакам. При проведении исследований с несколькими признаками будет наиболее удобным применить методы машинного обучения, как это представлено в статье [10], где используется алгоритм машинного обучения K-ближайших соседей.

Дипломная работа занимает 52 страницы, имеет 32 рисунка и 3 таблицы.

Обзор составлен по 27 информационным источникам.

Во введение рассматривается актуальность работы, устанавливается цель и выдвигаются задачи для достижения поставленной цели.

Первый раздел представляет собой анализ научно-методических основ идентификации огнестрельного оружия по следам на выстреленных пулях.

Во втором разделе работы рассмотрена проблема определения критерия относимости трасс к категории «совпадающие трассы» и «несовпадающие трассы» с помощью метода, основанного на расчете наименьшей энтропии Шеннона.

В третьем разделе проводится тестирование метода «Случайный лес», используемого для отнесения пар вторичных следов к парным или непарным множествам и оценивается точность прогнозирования данным методом.

### **Основное содержание работы**

Классическая криминалистическая методика сравнительного исследования довольно проста на первый взгляд. Эксперт сопоставляет два исследуемых следа друг с другом и ищет на них совпадающие трассы, основываясь на своих внутренних убеждениях.

Проблемой данной методики является то, что она требует больших трудовых затрат и большое количество времени. Также её недостатком

является, что выводы основываются только на внутренних убеждениях эксперта.

Попытки отойти от данной методики и прийти к количественному обоснованию внутренних убеждений эксперта видны в работах [2-7].

Но все перечисленные выше работы не предназначены для использования с сильно деформированными пулями. Деформированные пули составляют около 50% от пуль поступивших с мест происшествий на исследование. Данная проблема решается методом конгруэнтно совпадающих сегментов следов, представленном в работах [8] и [9].

К сожалению, в данной методике используется только 2 признака – отношение конгруэнтно совпавших максимумов в ФВК к количеству пригодных для исследования сегментов и общая корреляция двух исследуемых следов. В некоторых случаях этого бывает недостаточно.

Для возможности лучшего обоснования категорических выводов необходимо повысить количество признаков по которым проходит сравнение исследуемых следов, для чего стоит отойти от функции корреляции, обратив внимание на такие признаки как: серии совпадающих подряд трасс, плотность трасс, число совпадающих широких трасс в серии и т. д.

Для начала стоит определиться с тем какие трассы считать совпадающими.

Чтобы обеспечить автоматическое сравнение следов необходимо ответить на вопрос: «Какие трассы считать совпадающими?». Ответ на него не очевиден. Для ответа на этот вопрос необходимо определить оптимальную степень перекрытия трасс на совмещённых следах. Предлагается определять процент перекрытия по ширине совпадающих трасс в совмещённых следах, при котором массивы заведомо совпадающих и заведомо несовпадающих совмещённых следов наилучшим образом разделятся. Для этого можно применить метод, основанный на минимизации информационной энтропии Шеннона.

Формула информационной энтропии Шеннона для двух классов выглядит следующим образом:

$$S = - \sum_{i=1}^n p_i * \log_2 p_i \quad (1)$$

где  $p_i$  вероятность обнаружения объекта  $i$  – го класса.

Если упростить, то можно сказать, что энтропия отображает количество хаоса в системе. То есть, чем меньше энтропия, тем меньше хаоса в системе, тем лучше будут разделены следы по классам «совпадающие» и «несовпадающие».

В качестве объясняющей переменной  $x$  будет использоваться отношение числа совпадающих трасс  $n$ , к  $N_1$ , где  $N_1$  – число трасс в следе с наименьшим количеством трасс.

Гистограмма представленная на рисунке 1 была построена по полученным статистическим данным собранным по двум сотням валидаций, включающих по 40 совпадающих и 40 несовпадающих пар следов. Для каждой совмещенной пары следов подсчитывается отношение совпадающих трасс к их среднему числу. Данное действие проводится для всех степеней перекрытия трасс по ширине при каждой валидации.

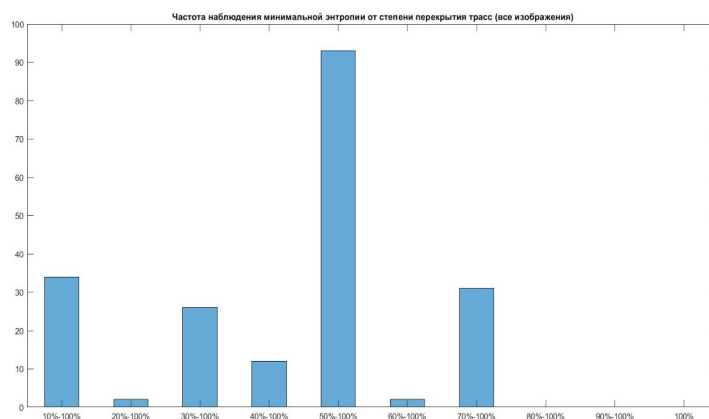


Рисунок 1 – Частота наблюдения минимальной энтропии от степени перекрытия трасс для всей исходной выборки следов

Проведенные расчеты показали, что оптимальным критерием, при котором трассы можно считать совпадающими является их перекрытие по ширине не

менее, чем на 50%. Данное значение было получено на основе статистических данных найденных из двух сотен валидационных выборок.

Чтобы уменьшить проблему потерь совпадающих трасс шириной 1 пиксель, было принято решение расширить все изображения следов в два раза. Работа проводилась с базой объемом 339 парных и 339 непарных бинаризованных следов. Валидационная выборка представлена двумя сотнями парных и непарных следов, на основе которых были рассчитаны статистические характеристики.

Так на рисунке 2 представлен один результат работы программы из серии в 20 экспериментов.

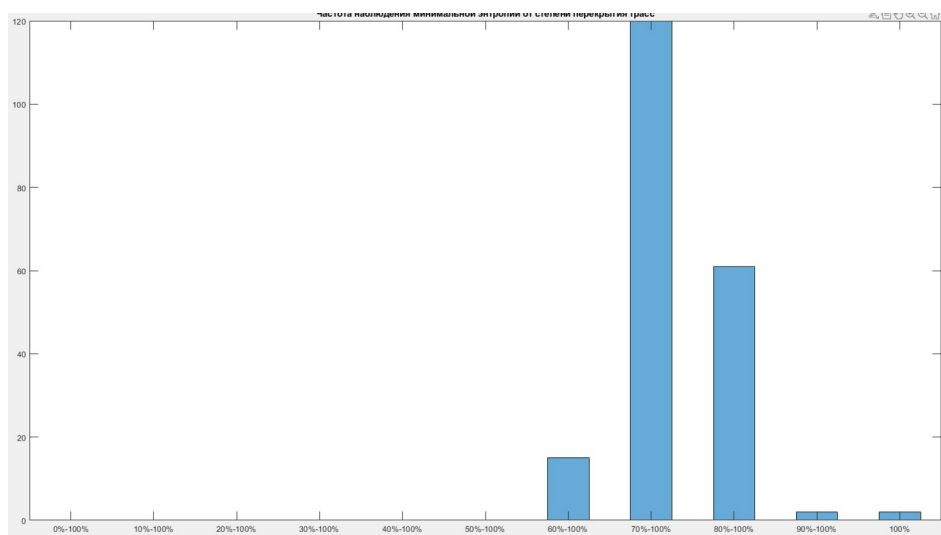


Рисунок 2 – График для одного эксперимента из серии

В данном варианте исследования с утолщением трасс можно подвести итог по полученным данным, что процент перекрытия трасс по ширине от 70 до 100 % является оптимальным.

Таким образом, при увеличении изображений, оптимальный диапазон перекрытия трасс по ширине изменился с 50-100% на 70-100%. Усреднив полученные в ходе двух экспериментов значения примем, что оптимальный диапазон перекрытия трасс по ширине можно принять равным от 60 до 100%.

С целью повысить точность и надёжность работы программы было принято решение провести работу с 11 признаками, такими как: количество серий

широких трасс, плотность трасс в следе, отношение числа совпавших трасс к среднему числу трасс в двух следах, метод оценки уникальности комплексов совпадающих трасс, коэффициент корреляции и т.д.

Результаты работы алгоритма «Случайный лес» представлены на рисунке 3, где по оси X отложено число деревьев в ансамбле, а по оси Y ошибка классификации.

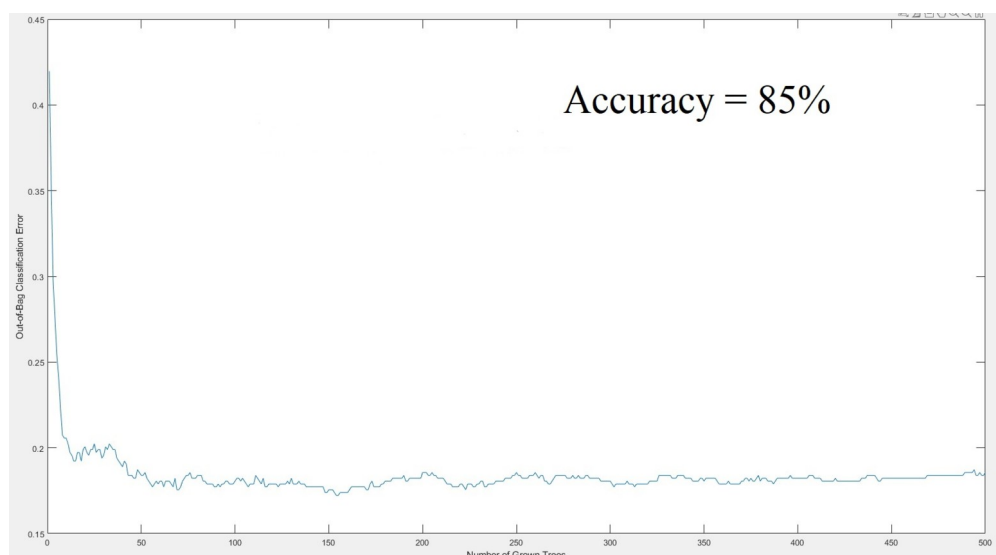


Рисунок 3 – Зависимость величины ошибки от числа деревьев в ансамбле, конечная точность классификации равна 85%

Подводя итог работы адаптированного алгоритма «Случайный лес», можно отметить, что идея его реализации для проведения классификации следов на парные и непарные множества по 11 признакам была удачной. Хорошую работу алгоритма подчеркивает точность классификации, равная 85%, вычисленная по серии из 20 экспериментальных расчетов.

### **Заключение.**

1. Проведён обзор научно-технической литературы и анализ современных методов сравнительного исследования вторичных следов.
2. Определен оптимальный критерий относимости трасс к категории «совпадающие» в совмещенных вторичных следах.



4. Проведен анализ метода машинного обучения «Случайный лес». Метод «Случайный лес» был успешно применён на практике, показав результат в 85% правильных предсказаний, что позволяет признать данную методику достойной дальнейшего развития.

#### **Список использованных источников**

1 Грановский, Г. Л. Вероятностная оценка пригодности линейных (динамических) следов для идентификации: метод. реком. для экспертов / Г. Л. Грановский. – М. : Б. и., 1985. – 20 с.

2 Федоренко, В. А. Метод бинаризации изображений следов на выстреленных пулях для автоматической оценки их пригодности к идентификации оружия / В. А. Федоренко, Е. В. Сидак // Известия Саратовского университета. Новая серия. Серия Экономика. Управление. Право. – 2016. – Т. 16, № 2. – С. 214-218.

3 Федоренко, В. А. Обработка цифровых изображений следов на пулях для автоматической идентификации оружия / В. А. Федоренко, Е. В. Сидак // Известия Саратовского университета. Новая серия. Серия Экономика. Управление. Право. – 2014. – Т.14, № 1. – С. 200-205.

4 Biasotti, A. A. Statistical Study of the Individual Characteristics of Fired Bullets / A. A. Biasotti // Journal Forensic Sciences. – 1959. – V. 4, № 1. – P. 34-50.

5 Федоренко, В. А. Концепция математической модели оценки уникальности наборов совпадающих трасс во вторичных следах на выстреленных пулях / В. А. Федоренко, О. А. Мыльцина // Известия Саратовского государственного университета. Новая серия. Экономика. Управление. Право. – 2016. – Вып. 2. – С. 209-213.

6 Федоренко, В. А. Методика оценки уникальности комплексов трасс, совмещенных во вторичных следах / В. А. Федоренко, Е. В. Сидак, О. А. Мыльцина // Известия Саратовского государственного университета. Новая серия. Экономика. Управление. Право. – 2018. – Т. 18, № 2. – С. 217-221.

7 Хэйр Э. Р. Статистические методы сопоставления пуль : дис. ... канд. юр. Наук : 12.00.12 : защищена 01.01.2017 : утв. 11.08.2018 / Эрик Реймер Хэйр ; США, Университет науки и техники штата Айова. – Эймс, 2017. – 107 с. : табл. – Библиогр.: с. 103-107.

8 Chen, Z. Fired bullet signature correlation using the Congruent Matching Profile Segments (CMPS) method / Z. Chen, W. Chu, J. A. Soons, R. M. Thompson, J. Song, X. Zhao // Forensic Science International. – 2019. – V. 305, № 13. – P. 10-19.

9 Chen, Z. Pilot study on deformed bullet correlation / Z. Chen, W. Chu, J. A. Soons, R. M. Thompson, J. Song, X. Zhao // Forensic Science International. – 2020. – V. 306, № 12. – P. 1-11.

10 Федоренко, В. А. Анализ следов на выстреленных пулях методами конгруэнтно совпадающих сегментов профилей и k-ближайших соседей / В. А. Федоренко, К. О. Сорокина, П. В. Гиверц // Информационные технологии и вычислительные системы. – 2021. – Вып. 1. – С. 70-82.

11 Сорокина, К. О. Определение оптимального критерия совмещения совпадающих трасс в сравниваемых следах полей нарезов на выстреленных пулях / К. О. Сорокина, В. А. Федоренко // Известия Саратовского государственного университета. Новая серия. Экономика. Управление. Право. – 2018. – Т. 18, № 2. – С. 217-221.