

АВТОМАТИЗИРОВАННЫЙ АНАЛИЗ СОСТОЯНИЯ ЛЕДОВОГО ПОКРЫТИЯ НА ОСНОВЕ ДАННЫХ АКСЕЛЕРОМЕТРА И ГИРОСКОПА

AUTOMATED ANALYSIS OF THE ICE SURFACE CONDITION BASED ON ACCELEROMETER AND GYROSCOPE DATA

V. Sorokin

Summary. This study proposes a method for automated ice surface quality assessment using accelerometer and gyroscope data from a smartphone. Signals collected during user movement on an ice rink were preprocessed through filtering, segmentation, and feature extraction. Machine learning models—including logistic regression, support vector machine, and random forest—were employed for binary classification of surface condition (smooth/rough). Logistic regression with an optimized classification threshold demonstrated the best performance. The results confirm the feasibility of the proposed method for ice surface monitoring and safety assurance.

Keywords: machine learning, inertial sensors, accelerometer, gyroscope, signal processing, ice surface, classification.

Сорокин Владимир Борисович

ФГБОУ ВО Волгоградский государственный
технический университет
vovochka.sorokin.2001@mail.ru

Аннотация. В работе представлена методика автоматизированного анализа состояния ледового покрытия на основе данных акселерометра и гироскопа, встроенных в мобильное устройство. Сигналы инерциальных датчиков, полученные при движении пользователя по катку, подвергались предварительной обработке, включая фильтрацию, сегментацию и извлечение признаков. Для решения задачи бинарной классификации покрытия (ровное/неровное) использовались модели машинного обучения: логистическая регрессия, метод опорных векторов и случайный лес. Наилучшие результаты показала логистическая регрессия с оптимизированным порогом классификации. Результаты подтверждают применимость разработанного подхода для мониторинга и обеспечения безопасности ледовых арен.

Ключевые слова: машинное обучение, инерциальные датчики, акселерометр, гироскоп, обработка сигналов, ледовое покрытие, классификация.

Обеспечение безопасности при эксплуатации ледовых арен и катков требует регулярного и объективного контроля состояния ледового покрытия. Повреждения, трещины и другие неровности на льду могут привести к травмам как у профессиональных спортсменов, так и у рядовых посетителей. Традиционные методы оценки состояния покрытия чаще всего основаны на визуальном осмотре и ручных измерениях, что делает их трудоёмкими, субъективными и не всегда своевременными.

Современные исследования демонстрируют высокую эффективность и перспективность применения мобильных сенсорных систем для автоматизированного мониторинга состояния различных типов покрытий. В частности, в работе [1] представлена информационная система, предназначенная для оценки состояния дорожного полотна, в которой используются акселерометры и GPS-модули, установленные на транспортных средствах. В [2] описан аналогичный подход, реализованный с использованием мобильных устройств, регистрирующих вибрационные и инерциальные характеристики при движении по автомобильным дорогам. В исследовании [3] предложены алгоритмы обработки сигналов инер-

циальных датчиков, направленные на анализ качества транспортной инфраструктуры, включая выявление дефектов покрытия и других аномалий. Приведённые работы демонстрируют потенциал мобильных устройств как доступного и масштабируемого инструмента для задач мониторинга, что обосновывает целесообразность их применения в смежных областях, таких как контроль состояния ледовых поверхностей.

В настоящей работе рассматривается возможность адаптации этих методов для оценки состояния ледового покрытия. В отличие от указанных исследований, в данном случае данные собирались с инерциальных датчиков смартфона, находящегося в кармане пользователя, перемещающегося по катку. Такой подход позволяет не только упростить сбор информации, но и сделать процесс мониторинга менее инвазивным и более доступным. Анализ полученных временных рядов даёт возможность выявлять участки льда с повышенным уровнем вибраций или неровностей, что, в перспективе, может быть использовано для автоматического предупреждения об опасных зонах и своевременного технического обслуживания покрытия.

Для сбора данных использовался смартфон Samsung Galaxy S23 FE, расположенный во внутреннем кармане одежды участника. Приложение Sensor Logger осуществляло непрерывную запись данных с акселерометра и гироскопа. Исходные данные сохранялись в формате CSV и включали временные метки, а также значения ускорений и угловых скоростей по трём осям.

Данные проходили предварительную обработку: синхронизация по времени, фильтрация шума, сегментация по временным окнам и извлечение признаков (включая применение Быстрого Преобразования Фурье — БПФ). Полученные признаки служили входными данными для алгоритма машинного обучения.

Сбор экспериментальных данных осуществлялся с использованием мобильного приложения Sensor Logger, обеспечивающего непрерывную регистрацию сигналов с инерциальных сенсоров, встроенных в смартфон. В процессе эксперимента фиксировались значения линейных ускорений (акселерометр) и угловых скоростей (гироскоп) по трём ортогональным осям (X, Y, Z). Основные технические параметры регистрации включали частоту дискретизации 100 Гц и экспорт данных в формате CSV. Структура выходных данных включала следующие поля: time (абсолютная временная метка), seconds_elapsed (время, прошедшее с начала записи), а также x, y, z — значения измеряемых величин по каждой пространственной оси.

Регистрация данных проводилась в условиях реальной эксплуатации на ледовой арене. Смартфон располагался в боковом кармане одежды испытуемого, что обеспечивало регистрацию вибрационных воздействий, передающихся через тело человека при перемещении по ледовому покрытию. Такой способ размещения сенсора имитирует практическое использование мобильного устройства для мониторинга состояния поверхности в бытовых и спортивных условиях.

Экспериментальная сессия включала траектории движения одного участника по всей поверхности катка с варьируемой интенсивностью и скоростью, включая участки с визуально идентифицированными дефектами ледового покрытия. Целью этапа являлось выявление характерных вибрационных паттернов, обусловленных различиями в микрорельефе ледяной поверхности.

До проведения анализа собранные данные прошли этап предварительной обработки, включавший:

- фильтрацию шумов и устранение выбросов с применением медианного фильтра и методов сглаживания;
- нормализацию значений по осям измерения;
- фрагментацию временного ряда на фиксированные окна;

- синхронизацию данных акселерометра и гироскопа по временным меткам.

Для каждого временного окна осуществлялось извлечение совокупности признаков, охватывающих как статистические, так и спектральные характеристики сигнала, включая:

- среднее значение, максимум и стандартное отклонение;
- энергию сигнала;
- амплитудные пики;
- спектральные параметры, вычисленные посредством быстрого преобразования Фурье (БПФ).

Полученные признаки были использованы в последующем этапе обучения моделей машинного обучения, направленных на автоматизированную классификацию состояния ледового покрытия (например, «гладкое», «неровное» и др.).

Целью данной работы является разработка метода автоматической классификации состояния ледового покрытия (ровный/неровный лёд) на основе данных инерциальных датчиков (акселерометра и гироскопа), установленных на конькобежце. Для этого необходимо решить задачу бинарной классификации с применением методов машинного обучения.

Пусть имеется множество временных рядов, полученных с акселерометра и гироскопа, установленных на спортсмене. Каждому участку записи сопоставлялась метка ($y \in \{0,1\}$ где $y = 0$ соответствует участку с ровным льдом, а $y = 1$ участку с неровным льдом).

Данные представлены в виде CSV-файлов, каждый из которых содержит трёхосные показания сенсоров (ось x, y, zx, y, y, z) и временные метки. Для совместного анализа данные акселерометра и гироскопа синхронизируются по времени и объединяются в одну выборку.

Этапы решения задачи

1. Загрузка и маркировка данных:

Для каждого CSV-файла определяется принадлежность к классу (ровный или неровный лёд) на основании его имени. Исключаются неклассифицированные или неполные записи.

2. Синхронизация данных акселерометра и гироскопа:

Используется линейная интерполяция для выравнивания временных меток и построения единого синхронизированного дата сета.

3. Оконное разбиение сигнала:

Временные ряды разбиваются на неперекрывающиеся окна фиксированной длины (например, 150 отсчётов). Для каждого окна извлекается набор признаков.

4. Извлечение признаков:

Для каждого сигнала (6 осей: 3 от акселерометра и 3 от гироскопа) извлекаются статистические и спектральные признаки: среднее, стандартное отклонение, максимум, минимум, энергия, автокорреляция, энтропия спектра, частота с наибольшей амплитудой и др. Всего формируется 120 признаков на окно (20 признаков × 6 осей).

5. Аугментация данных:

Для повышения обобщающей способности модели применяется добавление гауссовского шума к сигналам и повторное извлечение признаков из искажённого сигнала.

6. Предобработка:

Выполняется удаление постоянных признаков, нормализация (стандартизация), отбор признаков методом RFE и балансировка классов с использованием метода SMOTE.

7. Обучение моделей:

Используются несколько алгоритмов классификации: логистическая регрессия, метод опорных векторов (SVM) и случайный лес (Random Forest). Подбираются оптимальные гиперпараметры с помощью GridSearchCV и проводится кросс-валидация.

8. Оценка качества моделей:

Производится оценка точности, полноты, F1-меры, а также построение ROC-кривой и PR-кривой для анализа поведения модели при различных порогах.

Результатом работы модели является предсказание ($y \in \{0,1\}$) для каждого окна наблюдения. В дальнейшем возможно применение модели для:

- мониторинга качества ледовой поверхности в реальном времени;
- анализ тренировочных сессий и выявления опасных участков на катке;

Для решения задачи классификации дорожного покрытия на ровный и неровный лед были применены три модели машинного обучения: логистическая регрессия, метод опорных векторов (SVM) с радиально-базисной функцией (RBF) и случайный лес. Выбор этих моделей об-

условлен их эффективностью в задачах бинарной классификации, а также способностью обрабатывать высоко размерные данные с ограниченным объемом выборки.

Логистическая регрессия была выбрана в качестве базовой модели благодаря ее интерпретируемости и устойчивости при работе с небольшими наборами данных. Для оптимизации гиперпараметров использовался поиск по сетке (Grid Search) с перекрестной проверкой (Stratified K-Fold, $k=15$), что позволило подобрать оптимальные параметры: $C=100$, штраф L2, решатель liblinear. Для повышения качества предсказаний применялась балансировка классов с использованием параметра `class_weight='balanced'`. Оптимизация порога классификации проводилась на основе максимизации F1, что позволило улучшить баланс между точностью (Precision) и полнотой (Recall). Оптимальный порог составил 0.42.

Модель SVM с ядром RBF была выбрана для учета нелинейных зависимостей в данных. Параметры модели, включая веса классов (`class_weight='balanced'`), были настроены для компенсации дисбаланса классов. Вероятностные оценки, полученные с помощью метода `predict_proba`, использовались для анализа качества классификации. Порог классификации для SVM был установлен на уровне 0.45.

Случайный лес был применен как ансамблевый метод, способный эффективно обрабатывать сложные зависимости в данных и снижать риск переобучения. Для настройки модели использовались параметры: глубина дерева (`max_depth=10`), количество деревьев (`n_estimators=250`), минимальное число образцов для разделения узла (`min_samples_split=10`) и минимальное число образцов в листе (`min_samples_leaf=3`). Балансировка классов осуществлялась с помощью весов (`class_weight={0:1, 1:3}`), что позволило учесть неравномерное распределение классов. Порог классификации для случайного леса был установлен на уровне 0.50. Качество модели оценивалось с использованием перекрестной проверки (5-кратная, метрика F1).

Данные акселерометра и гироскопа подвергались предварительной обработке, включая синхронизацию временных меток, интерполяцию и нормализацию. Для извлечения признаков применялся анализ временной области (среднее, стандартное отклонение, медиана, межквартильный размах, энергия сигнала, асимметрия и др.) и частотной области (энергия высокочастотных, среднечастотных и низкочастотных компонент, доминирующая частота, спектральная энтропия). Для повышения устойчивости модели к шуму в данные добавлялся искусственный шум с фактором 0.02, что способствовало улучшению обобщающей способности моделей.

Для уменьшения размерности и устранения избыточных признаков применялись следующие методы:

- Variance Threshold: удаление признаков с нулевой дисперсией.
- Recursive Feature Elimination (RFE): отбор 30 наиболее значимых признаков на основе логистической регрессии.
- SMOTE: балансировка классов путем генерации синтетических образцов для миноритарного класса (баланс классов после SMOTE: 3567 образцов для каждого класса).

Корреляционный анализ показал, что отобранные признаки имеют умеренную корреляцию с целевой переменной (коэффициенты Пирсона в диапазоне от -0.06 до 0.14). Наиболее коррелирующими оказались признаки на позициях 16, 40 и 56 (корреляции 0.13 , 0.13 и 0.14 соответственно).

Качество моделей оценивалось с использованием метрик Accuracy, Precision, Recall и F1-score. Для анализа производительности строились ROC-кривые и матрицы ошибок. Ниже приведены результаты для тестовой выборки (30 % от общего объема данных после применения SMOTE).

По логистической регрессии:

- Accuracy: 0.75 (с исходным порогом); 0.77 (с оптимальным порогом 0.42 , вычислено на основе Precision и Recall)
- Precision: 0.75 (с исходным порогом); 0.72 (с оптимальным порогом)
- Recall: 0.75 (с исходным порогом); 0.83 (с оптимальным порогом)
- F1-score: 0.75 (с исходным порогом); 0.77 (с оптимальным порогом)

Площадь под ROC-кривой (AUC) составила 0.83 , что свидетельствует о хорошей способности модели разделять классы видно на рис. 1. Матрица ошибок с оптимальным порогом показывает, что модель корректно классифицировала 724 примера класса 0 и 891 примера класса 1, допустив 347 ложноположительных и 179 ложноотрицательных ошибок, представлено на рис. 2.

SVM с ядром RBF:

- Accuracy: 0.79
- Precision: 0.79
- Recall: 0.78
- F1-score: 0.78

SVM продемонстрировала стабильные результаты, немного превосходя логистическую регрессию по Accuracy и Precision, но уступая по Recall при использовании оптимального порога для логистической регрессии.

Случайный лес:

- Accuracy: 0.72

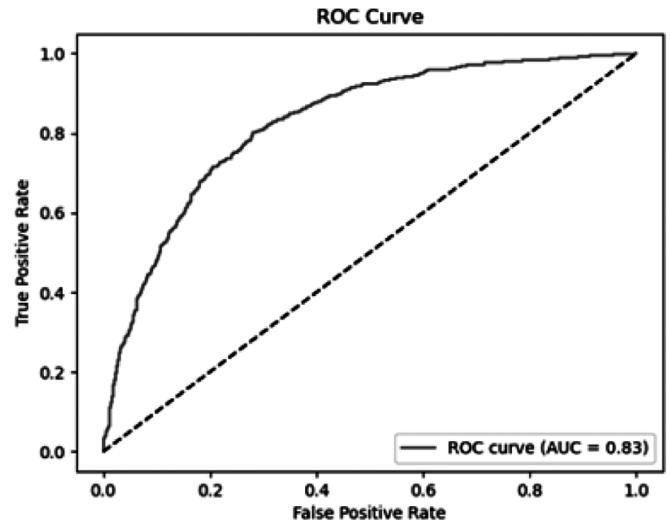


Рис. 1. ROC-кривая для логистической регрессии с $AUC = 0.83$

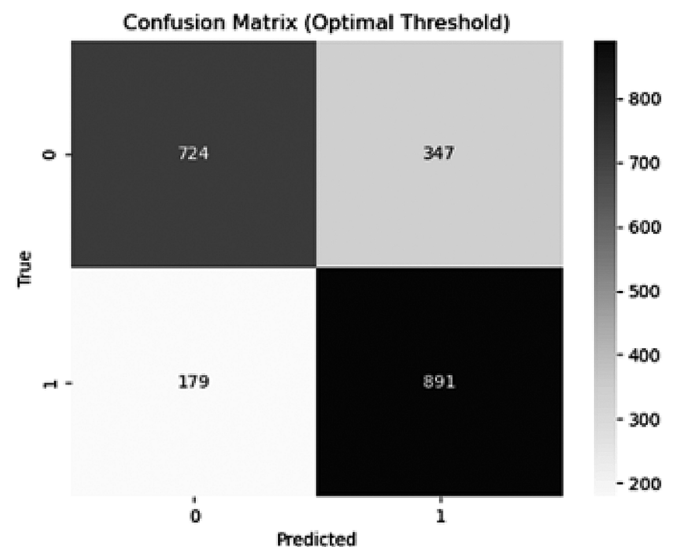


Рис. 2. Матрица ошибок для логистической регрессии с оптимальным порогом 0.42

- Precision: 0.64
- Recall: 0.99
- F1-score: 0.78 (среднее по перекрестной проверке: 0.73 ± 0.01)

Случайный лес показал высокую полноту (Recall), что делает его подходящим для задач, где важно минимизировать ложноотрицательные ошибки. Однако низкая точность (Precision) указывает на большое количество ложноположительных предсказаний.

Сравнительный анализ показал, что модель логистической регрессии с оптимизированным порогом (0.42) демонстрирует наилучший баланс метрик: F1-score 0.77 , Accuracy 0.77 , с $AUC = 0.83$. SVM с ядром RBF также показала высокие результаты (Accuracy 0.79 F1-score 0.78), но уступает по Recall (0.78 против 0.83 у логистической регрессии с оптимальным порогом). Случайный лес, не-

смотря на высокую полноту (Recall 0.99), имеет низкую точность (Precision 0.64), что делает его менее предпочтительным для данной задачи.

Для повышения надежности предсказаний применялся метод большинства голосов (majority voting) на основе предсказаний всех трех моделей для тестовых файлов с префиксами «00» (ровный лед) и «111» (неровный лед). Результаты агрегации показали, что для префикса «00» все модели единогласно предсказали класс 0 (ровный лед), а для префикса «111» логистическая регрессия и случайный лес предсказали класс 1 (неровный лед), тогда как SVM ошибочно классифицировал его как 0.

Из пяти предварительных экспериментов с различными конфигурациями моделей (варьирование гиперпараметров и порогов) лучшей оказалась текущая конфигурация логистической регрессии с оптимизированным порогом, так как она обеспечила наилучший баланс между Precision и Recall, что критически важно для задачи классификации покрытия горизонтальной поверхности.

Ограниченный объем данных требовал применения методов аугментации (добавление шума, SMOTE), что позволило улучшить обобщающую способность моделей. Однако дальнейшее увеличение объема данных и включение дополнительных признаков (например, кросс-корреляций между акселерометром и гироскопом) может повысить точность классификации. Также стоит отметить, что оптимизация порогов классификации была критически важной для достижения сбалансированных метрик Precision и Recall. В будущем рекомендуется провести дополнительные эксперименты с более крупным и разнообразным набором данных, а также исследовать влияние различных методов извлечения признаков на качество классификации.

В рамках проведенного исследования разработана и апробирована методика автоматизированного анализа состояния ледового покрытия с использованием данных инерциальных датчиков — акселерометра и гироскопа, встроенных в смартфон. Предложенный под-

ход включает сбор данных в реальных условиях эксплуатации ледовой арены, их предобработку, выделение признаков и обучение моделей машинного обучения для бинарной классификации покрытия на ровный и неровный лёд.

Сравнительный анализ трёх моделей — логистической регрессии, метода опорных векторов (SVM) с ядром RBF и случайного леса — показал, что наилучший баланс метрик качества классификации достигается при использовании логистической регрессии с оптимизированным порогом (0.42). Данная модель обеспечила Accuracy 0.77 F1-score 0.77 и AUC 0.83, что свидетельствует о её способности эффективно различать классы даже при ограниченном объёме данных. Метод большинства голосов, применённый для агрегации предсказаний, подтвердил надёжность классификации: для тестовых данных с префиксом «00» (ровный лёд) все модели единогласно предсказали класс 0, а для префикса «111» (неровный лёд) логистическая регрессия и случайный лес корректно определили класс 1.

Полученные результаты демонстрируют перспективность использования инерциальных датчиков смартфона для мониторинга состояния ледового покрытия. Такой подход может быть интегрирован в системы управления ледовыми аренами для своевременного выявления дефектов покрытия, что повысит безопасность спортсменов и посетителей. Кроме того, методика отличается доступностью, так как не требует специализированного оборудования, и может быть адаптирована для анализа других типов поверхностей.

Тем не менее, ограниченный объём данных и необходимость применения методов аугментации (SMOTE, добавление шума) указывают на потребность в дальнейшем расширении экспериментальной базы. В перспективе планируется увеличение объёма выборки, включение дополнительных признаков (например, кросс-корреляций между данными акселерометра и гироскопа) и исследование более сложных моделей для улучшения качества классификации.

ЛИТЕРАТУРА

1. Eriksson J., Girod L., Hull B., Newton R., Madden S., Balakrishnan H. Automated Sensing System for Monitoring of Road Surface Quality by Mobile Devices // IEEE Transactions on Mobile Computing. — 2015. — Vol. 14, № 5. — Pp. 1026–1039. — DOI: 10.1109/TMC.2014.2346772.
2. Кузнецов А.С., Петров Д.И., Смирнов Е.А. Распознавание и отслеживание дефектов дорожного полотна в реальном времени на основе комплексного использования стандартных вычислительных процедур и глубоких нейронных сетей // Системы и средства информатики. — 2024. — Т. 34, № 3. — С. 421–430. — URL: <https://swsys.ru/files/2024-3/421-430.pdf>.
3. Иванов П.К., Сидоров М.В. Разработка информационной системы мониторинга состояния дорожного покрытия // Современные информационные технологии. — 2023. — № 2. — С. 45–52. — URL: <https://cyberleninka.ru/article/n/razrabotka-informatsionnoy-sistemy-monitoringa-sostoyaniya-dorozhnogo-pokrytiya>.

© Сорокин Владимир Борисович (vovochka.sorokin.2001@mail.ru)

Журнал «Современная наука: актуальные проблемы теории и практики»