

DOI: <https://doi.org/10.30898/1684-1719.2025.11.4>

УДК: 621.396.96

ИНТЕРПРЕТИРУЕМАЯ АКУСТИЧЕСКАЯ КЛАССИФИКАЦИЯ МАЛОРАЗМЕРНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ С ИСПОЛЬЗОВАНИЕМ МЕТОДА GRAD-CAM

Е.И. Минаков¹, В.Н. Гридин², Д.Г. Андроничев¹,
В.И. Солодовников², А.А. Сычуг¹, Ю.В. Французова¹

¹Тульский государственный университет,
300012, Тульская область, город Тула, пр-кт Ленина, д.92

²Центр информационных технологий в проектировании РАН,
143003, Московская область, город Одинцово, улица Маршала Бирюзова, д. 7А

Статья поступила в редакцию 6 октября 2025 г.

Аннотация. Задача своевременного обнаружения и классификации малоразмерных летательных аппаратов осложняется ограниченной точностью традиционных спектральных методов и их низкой помехоустойчивостью. Цель работы – разработать интерпретируемую модель акустической классификации, которая сочетает в себе высокую точность и прозрачность принятия решений. В работе предложена интерпретируемая модель акустической классификации на основе сверточной нейронной сети, обученной на MFCC-признаках и оптимизированной с помощью Keras Tuner. Для визуализации значимых частотно-временных областей используется метод Grad-CAM. Экспериментальная оценка показала высокое качество распознавания: точность составила 0.97 (95 % ДИ: 0.94-0.99), macro-F1 = 0.98, ROC-AUC = 0.998. Модель продемонстрировала устойчивость к шумовым искажениям до уровня –5 дБ SNR, а карты активации подтвердили, что решения основаны на физических характеристиках сигналов – фундаментальных частотах роторов и их гармониках. Разработанный

подход может применяться в системах мониторинга воздушного пространства, защиты объектов и обеспечения безопасности, предоставляя оператору как точный прогноз, так и его интерпретацию.

Ключевые слова: акустическая классификация, сверточные нейронные сети, признаки MFCC, Grad-CAM, визуализация активации, глубокое обучение, спектральный анализ, обработка аудиосигнала, интерпретируемость модели.

Финансирование: Работы проводились при финансовой поддержке Гранта в форме субсидии на проведение научных исследований Комитета по науке и инноватике Тульской области №15 от 21 июня 2024 г.

Автор для переписки: Минаков Евгений Иванович, EMinakov@bk.ru

Введение

Современные методы классификации акустических сигналов от малоразмерных летательных аппаратов включают как традиционные спектральные подходы, так и передовые методы глубокого обучения. Выбор метода определяется качеством данных, доступными вычислительными ресурсами и требованиями к точности классификации.

Недавние исследования продемонстрировали эффективность глубоких нейронных сетей в этой области. Например, Rakshit и Bagheri Zadeh (2024) предложили модель на основе глубоких сверточных сетей и FIR-фильтрации спектрограмм, которая позволила повысить устойчивость к шуму и улучшить классификацию аудиовизуальных данных с дронов [1].

Wang и соавторы (2025) сравнили три метода извлечения признаков: спектрограмму, Fbank и MFCC, и показали, что MFCC обеспечивает наилучшую производительность в условиях низкого соотношения сигнал/шум (SNR), особенно при использовании облегченной модели ResNet10-CBAM с модулями внимания [2,3].

Yi и др. (2023) разработали многозадачную модель глубокой нейронной сети (1DCNN-ResNet), способную не только классифицировать неисправности

дронов по звуку, но и определять направление их полета. Модель показала стабильность при соотношении сигнал/шум до 10-15 дБ [4].

Gola и соавторы (2024) экспериментально подтвердили, что при акустических измерениях в городских условиях модели нейронных сетей, работающие со спектрограммами, эффективно отделяют сигналы малоразмерных летательных аппаратов от фонового шума [5].

В обзоре Seidaliyeva и др. (2023) систематизированы современные подходы к обнаружению малоразмерных летательных аппаратов, в том числе акустические методы, а также гибридные решения с системой объединения датчиков [6].

Rahman (2024) представил всесторонний обзор технологий машинного и глубокого обучения для обнаружения и классификации малоразмерных летательных аппаратов, подчеркнув растущий интерес и важность акустических подходов среди других методов (радиочастотных, визуальных, радиолокационных) [7].

Предложенный компанией open-review метод слияния результатов на уровне CNN, направленный на обнаружение малоразмерных летательных аппаратов в окружающей среде, также показал высокую эффективность [8].

Zaman и др. (2025) рассмотрели проблемы небольшого объема обучающих данных в задачах классификации звука малоразмерных летательных аппаратов; комбинаторное применение CNN и трансформаторов (Audio Spectrogram Transformer) дало хорошие результаты при ограниченной выборке [9].

В масштабном исследовании Semenyuk (2025) представлен обзор современных технологий обнаружения и классификации малоразмерных летательных аппаратов, в том числе акустических, с точки зрения их актуальности в 2020-2025 годах [10].

Сверточные нейронные сети обеспечивают высокую точность классификации и устойчивость к шумовым помехам, но глубокие модели часто воспринимаются как «черные ящики», поскольку процессы принятия решений внутри сети остаются недостаточно прозрачными. Это вызывает определенные трудности при интерпретации результатов и требует

дополнительных исследований в области интерпретируемости моделей глубокого обучения.

В последние годы проблема интерпретируемости нейронных сетей приобретает все большую актуальность. Эта задача важна для понимания того, как именно сети принимают решения, и почему они ошибаются. Непрозрачность нейронных моделей вызывает вопросы доверия к их результатам, особенно в критических областях, таких как медицина, юриспруденция и автономные системы.

Исследователи Зейлер и Фергус в 2014 году предложили метод визуализации сверточных слоев [11], который стал важным шагом в понимании внутренних механизмов нейронных сетей. Их подход, основанный на визуализации промежуточных активаций, помог улучшить понимание того, как сеть обрабатывает информацию на разных уровнях. Однако этот метод не всегда давал достаточно точную картину того, какие именно области изображения важны для принятия решения.

В 2016 году Чжоу и ее соавторы представили метод Class Activation Mapping (CAM) [12], который стал прорывом в области интерпретируемости нейронных сетей. CAM позволяет визуализировать наиболее значимые области входного изображения для конкретного класса, что делает его полезным инструментом для анализа и отладки моделей. Этот метод был разработан специально для сверточных нейронных сетей (CNN) и показал высокую эффективность в задачах классификации изображений.

В 2017 году группа исследователей под руководством Селвараджу предложила метод Grad-CAM [13], который использует градиенты модели для выявления ключевых областей изображения, влияющих на принятие решений. Этот подход стал еще более мощным инструментом для интерпретации работы нейронных сетей. Grad-CAM не только позволяет визуализировать важные области, но и предоставляет информацию о том, как модель оценивает различные части изображения.

Grad-CAM был успешно использован не только в задачах классификации изображений, но и в более сложных областях, таких как мультимодальный анализ и обработка аудио. Это расширило возможности его применения и сделало его важным инструментом для исследователей и разработчиков.

В 2018 году Чаттопадхай и ее коллеги предложили усовершенствование Grad-CAM, которое повысило его точность и информативность [14]. Их подход, известный как Grad-CAM++, использует дополнительные механизмы для более точного выявления ключевых областей, что сделало Grad-CAM еще более эффективным инструментом для обеспечения прозрачности и достоверности решений нейронных сетей.

За последние годы был достигнут значительный прогресс в области интерпретируемости нейронных сетей. Методы визуализации, такие как CAM и Grad-CAM, стали важными инструментами для понимания работы нейронных моделей и повышения доверия к их результатам. Эти достижения открывают новые возможности для применения нейронных сетей в различных областях и способствуют развитию более прозрачных и надежных систем искусственного интеллекта.

Цель этой работы – разработать интерпретируемую модель акустической классификации для малоразмерных летательных аппаратов, которая сочетает в себе высокую точность распознавания и прозрачность принятия решений, обеспечиваемую методом Grad-CAM.

Задача состоит в том, чтобы создать и обучить сверточную нейронную сеть на основе признаков MFCC, оптимизированных с помощью автоматического выбора гиперпараметров, и интегрировать Grad-CAM для визуализации наиболее значимых частотно-временных областей акустического сигнала.

Цели исследования:

- 1) Разработать архитектуру сверточной нейронной сети, устойчивую к шумовым искажениям в акустических сигналах.

- 2) Внедрить автоматизированный подбор гиперпараметров (Keras Tuner) для оптимизации архитектуры.

3) Интегрировать Grad-CAM для интерпретации производительности модели и проверки точности обнаруженных признаков.

4) Провести экспериментальную оценку точности классификации и устойчивости к шуму на тестовом наборе данных.

1. Данные

Записи аудио фрагментов были сделаны в полигонных условиях с использованием двух микрофонов МК-207 и МД-307. Запись велась на открытой площадке, где присутствовали фоновые шумы (ветер, транспорт, отдаленные разговоры), что обеспечивает приближение к реальным условиям эксплуатации. Частота дискретизации составила 22050 Гц, длительность каждого фрагмента – 2 секунды. Дополнительно в набор был включен класс «Прочие звуки», содержащий громкие хлопки, звуки транспортных средств, работу двигателей и другие шумы из открытых источников.

Для обеспечения корректной оценки данные были разделены на три части: 75 % – обучающая, 20 % – валидационная, 5 % – тестовая. При этом разбиение выполнялось с учетом условий записи и оборудования, что исключало попадание записей одного и того же сеанса или устройства одновременно в разные выборки. Такой протокол предотвращает утечки и обеспечивает независимость теста.

Сводная статистика классов представлена в таблице 1. В ней приведены названия классов, количество и общая длительность записей по каждому классу. Названия классов приведены в условном виде («Тип 1-9») и соответствуют различным моделям малоразмерных летательных аппаратов и прочим звукам.

Таблица 1. Распределение записей по классам, длительности и количеству клипов.

Класс	Количество записей	Длительность (ч)
Тип 1	397	0.194
Тип 2	52	0.024
Тип 3	1089	0.577
Тип 4	64	0.031
Тип 5	490	0.250
Тип 6	1223	0.653
Тип 7 (Прочие звуки)	292	0.161
Тип 8	64	0.031
Тип 9	30	0.015
Итого	3701	1.94

2. Предобработка

Аудиозаписи были приведены к фиксированной длительности в 2 секунды при частоте дискретизации 22050 Гц. Для построения спектрального представления применялось преобразование Фурье с окном 1024 отсчета, шагом 256 и функцией окна Ханна. Из логарифмированных мел-спектрограмм, рассчитанных по 64 полосам, извлекались 40 коэффициентов MFCC, формирующих компактное описание временно-частотных характеристик. Полученные признаки подвергались стандартизации по каждому коэффициенту на основе статистик обучающей выборки, что позволило устранить различия в масштабах и смещения. В результате входные данные модели имели форму $40 \times 173 \times 1$, где 40 соответствует числу коэффициентов, 173 – количеству временных кадров, а единица – числу каналов.

Для оценки устойчивости модели к шумам использовался дополнительный набор зашумленных спектрограмм, в которые добавлялся гауссов шум при различных соотношениях сигнал/шум (–5, 0, 5, 10 и 15 дБ). Анализ показал, что модель сохраняет высокое качество вплоть до уровня 0 дБ и начинает заметно деградировать лишь при –5 дБ (табл. 2).

Таблица 2. Точность и macro-F1 модели при различных уровнях SNR.

SNR (дБ)	Accuracy	Macro-F1
-5	0.85	0.84
0	0.97	0.96
5	0.99	0.99
10	1.00	1.00
15	1.00	1.00

3. Архитектура и обучение

Для задачи классификации использовалась сверточная нейронная сеть, включающая четыре сверточных блока (Conv2D → ReLU → BatchNorm → MaxPooling) с числом фильтров 16, 32, 192 и 128 и размером ядра 5. После блока GlobalAveragePooling2D применялся полносвязный слой из 64 нейронов с L2-регуляризацией (0.001) и Dropout (0.45). Финальный слой SoftMax обеспечивал классификацию по девяти классам. Общее число обучаемых параметров модели составило около 792 тысяч.

Обучение проводилось в течение 40 эпох с размером батча 16, оптимизатором Adam и скоростью обучения 0.005. Для предотвращения переобучения использовались EarlyStopping (по функции потерь на валидации, patience = 5) и снижение learning rate на плато (factor = 0.5, patience = 2).

Подбор гиперпараметров выполнялся с помощью Keras Tuner (RandomSearch, 200 испытаний). Пространство поиска включало вариации числа фильтров в каждом блоке, размер ядра свертки, размерность полносвязного слоя, коэффициенты регуляризации, dropout и скорость обучения.

Итоговые оптимальные значения: kernel size = 5; filters = (16, 32, 192, 128); dense = 64; L2 = 0.001; dropout = 0.45; learning rate = 0.005.

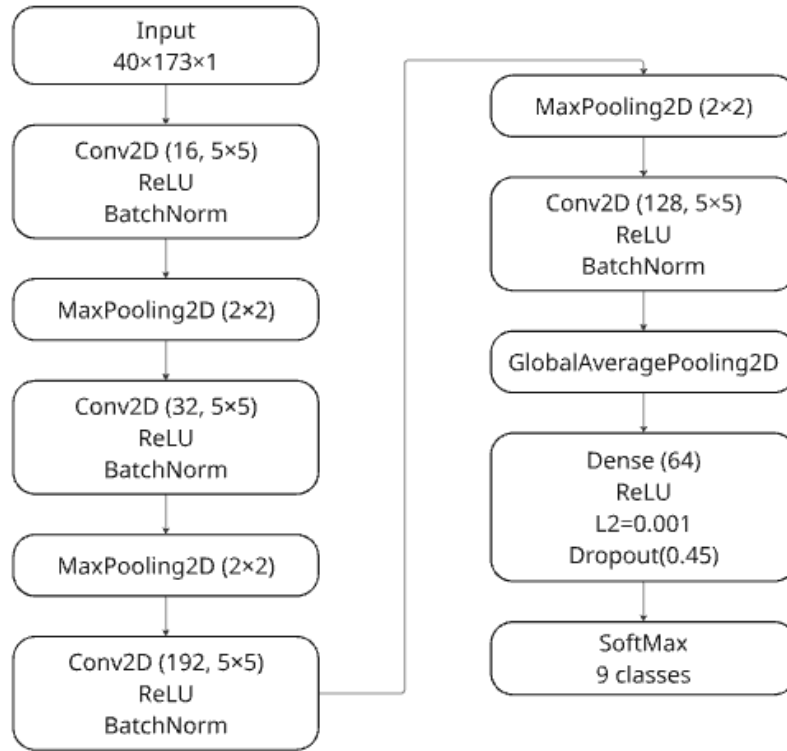


Рис. 1. Архитектура созданной сети.

Для наглядности архитектура сети представлена на схеме (рис. 1), где показана последовательность слоев и точки применения нормализации, активаций и операций усреднения.

4. Интерпретируемость

В качестве формального инструмента анализа использован метод Grad-CAM, который позволяет визуализировать наиболее значимые частотно-временные области MFCC-представлений.

Пусть последний сверточный слой формирует карты признаков $A^k \in R^{(H \times W)}$ ($k=1, \dots, K$). Если после него стоит Global Average Pooling (GAP) и сразу классификатор, логит класса y^c выражается через усредненные по пространству карты:

$$y^c = \sum_{k=1}^K \omega_k^c \left(\frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W A_{ij}^k \right), \quad (1)$$

где ω_k^c – обучаемый вес, соединяющий k -ю карту с классом c .

Тогда CAM-карта важности для класса c :

$$CAM^c(i, j) = \sum_{k=1}^K \omega_k^c A_{ij}^k, (i, j) \in \{1...H\} \times \{1...W\} \quad (2)$$

Метод ограничен архитектурами вида «Conv $\rightarrow$ GAP $\rightarrow$ Softmax» и плохо совместим с дополнительными «головами».

Grad-CAM снимает архитектурное ограничение, используя градиенты логита y^c по активациям произвольного сверточного слоя:

$$\alpha_k^c = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W \frac{\partial y^c}{\partial A_{ij}^k}, \quad (3)$$

$$L_{Grad-CAM}^c(i, j) = ReLU \left(\sum_{k=1}^K \alpha_k^c A_{ij}^k \right)$$

Здесь α_k^c – веса каналов, вычисленные усреднением градиента по пространству; $ReLU$ оставляет только поддерживающий вклад. Если в сети есть GAP, то Grad-CAM переходит в CAM.

В процессе работы с Grad-CAM мы начинаем с выбора слоя, чаще всего это последний сверточный блок. Его карты содержат уже обобщенные и семантически насыщенные признаки, при этом пространственное разрешение остается приемлемым.

Для оценки корректности получаемых карт применены две независимые проверки.

Во-первых, проведен deletion test (табл. 4): маскирование 20-30 % наиболее «горячих» областей, выделенных Grad-CAM, приводило к резкому снижению медианной уверенности модели в истинном классе ($c \approx 0.998$ до 0.285 при 20 % и до 0.022 при 30 %). Для случайного маскирования тех же долей падение было существенно слабее (0.753 и 0.335 соответственно). Таким образом, модель действительно опирается на признаки, выделяемые Grad-CAM, а не на случайные фрагменты спектра.

Во-вторых, выполнена оценка стабильности карт по чекпоинтам обучения. Сравнение Grad-CAM, полученных для разных эпох одной тренировки, показало достаточно высокое совпадение (средний IoU > 0.6 для top-20 % и ~ 0.65 для top-30 %). Это свидетельствует о воспроизводимости «горячих» областей при изменении параметров модели в процессе обучения.

В совокупности данные проверки подтверждают, что Grad-CAM-визуализации отражают физически значимые закономерности (фундаментальные частоты роторов и гармоники) и обеспечивают интерпретируемость модели.

5. Результаты

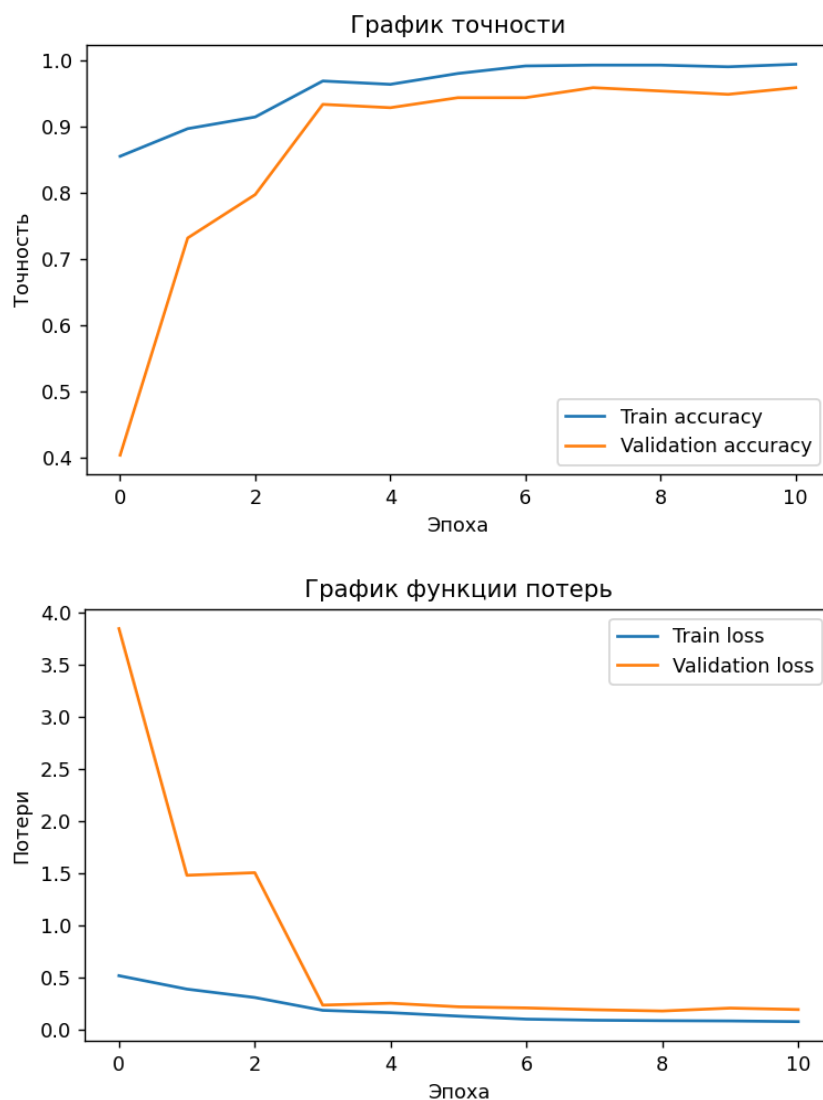


Рис. 2. Динамика функции потерь и точности на обучающем и валидационном множествах.

Модель показала стабильную сходимость обучения (рис. 2): валидационная точность превысила 0.95 уже к 6-й эпохе и стабилизировалась на уровне 0.97. Потери на обучении и валидации демонстрируют быстрое уменьшение и отсутствие признаков переобучения.

Таблица 3. Классовые метрики (Precision, Recall, F1).

Класс	Precision	Recall	F1-score
Тип 1	0.99	0.95	0.97
Тип 2	1.00	1.00	1.00
Тип 3	1.00	1.00	1.00
Тип 4	1.00	1.00	1.00
Тип 5	0.95	0.81	0.87
Тип 6	0.95	0.95	0.95
Тип 7 (Прочие звуки)	1.00	1.00	1.00
Тип 8	1.00	1.00	1.00
Тип 9	1.00	1.00	1.00
Среднее значение	0.99	0.97	0.98

Таблица 3 содержит сводные метрики по всем девяти классам, названия классов приведены в условном виде («Тип 1-9») и соответствуют различным моделям малоразмерных летательных аппаратов.

Для большинства категорий значения Precision, Recall и F1-score близки к единице. Наибольшие трудности наблюдаются при классификации «Малоразмерный летательный аппарат тип 5» (снижение Recall до 0.81), что подтверждается пониженными значениями на матрице ошибок (рис. 3).

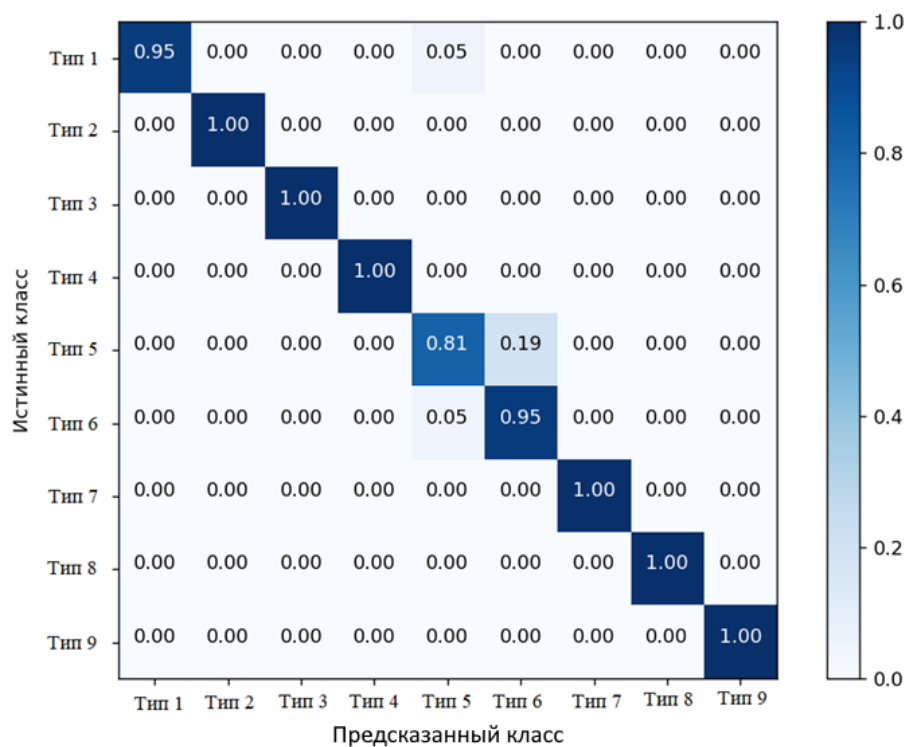


Рис. 3. Нормализованная матрица ошибок.

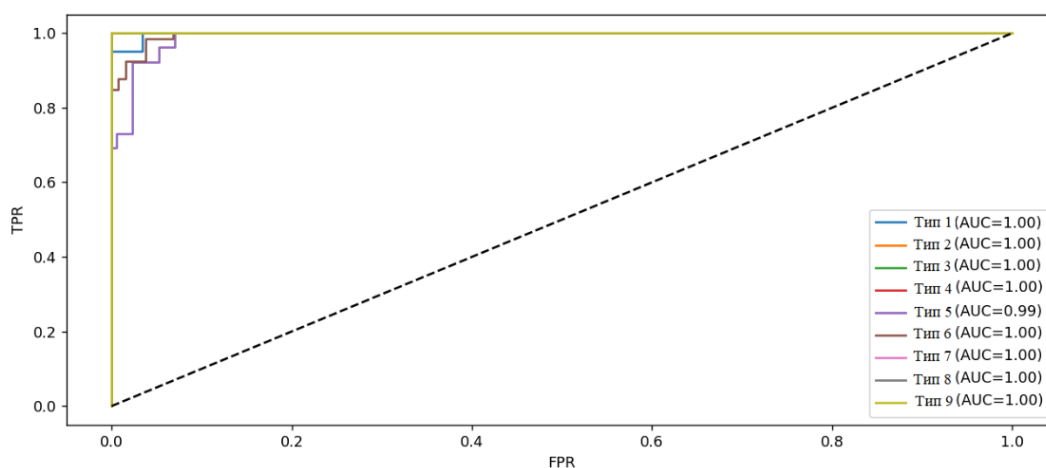


Рис. 4. ROC-кривые для девяти классов.

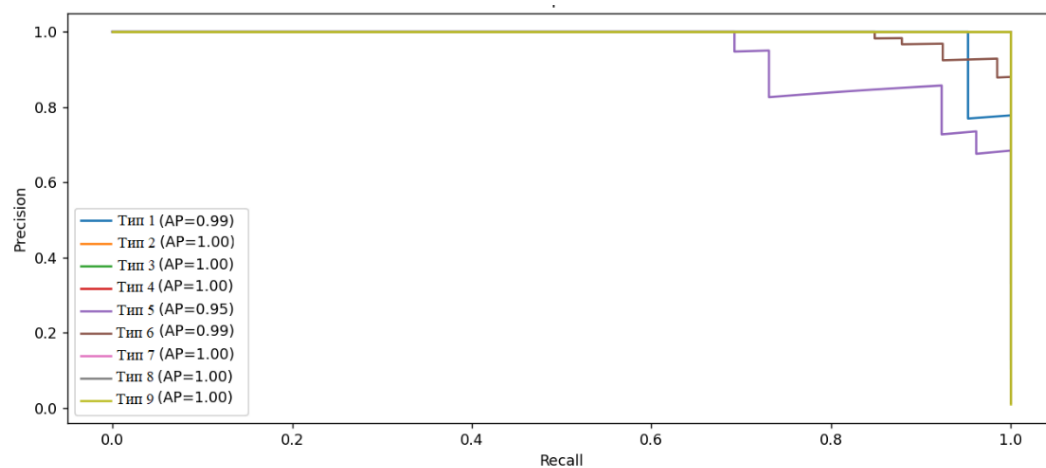


Рис. 5. Precision-Recall кривые для девяти классов.

ROC-кривые (рис. 4) демонстрируют $AUC \approx 1.0$ для всех классов (минимальное значение 0.95 у «Малоразмерный летательный аппарат тип 5»), что указывает на высокую разделимость признаков. Аналогично, PR-кривые (рис. 5) подтверждают высокий уровень Precision и Recall даже при сильном дисбалансе классов.

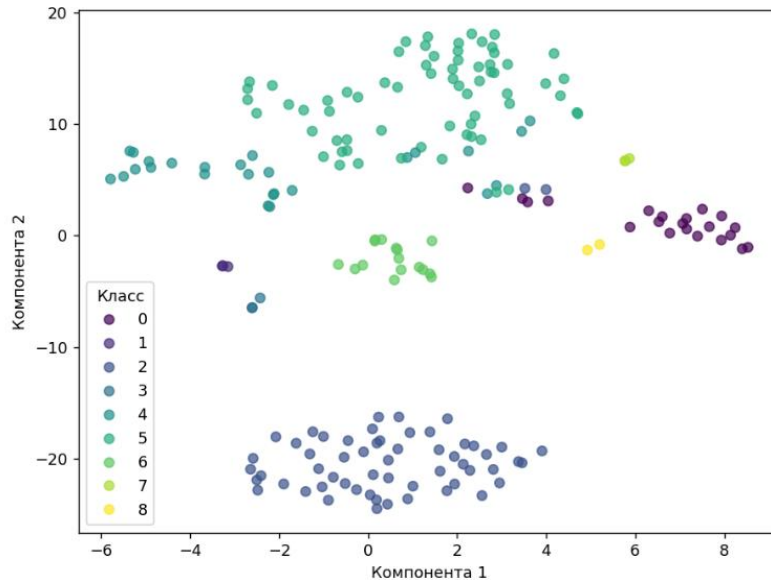


Рис. 6. t-SNE-визуализация эмбедингов последнего скрытого слоя.

t-SNE визуализация эмбедингов (рис. 6) показывает четкое разделение классов в пространстве признаков, без значительного перекрытия, за исключением частичного сближения классов «Малоразмерный летательный аппарат тип 4» и «Малоразмерный летательный аппарат тип 5».

Для проверки устойчивости к шумам была проведена оценка точности при различных уровнях SNR (от -5 до $+15$ дБ). Результаты (табл. 2) показывают стабильную работу модели: даже при $SNR = -5$ дБ точность остается выше 0.85.

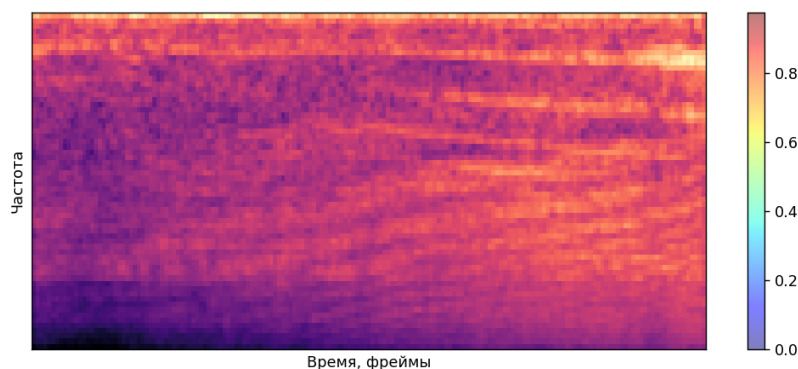


Рис. 7. Пример Grad-CAM-карты на лог-мел спектрограмме.

Визуализации Grad-CAM (рис. 7) подтверждают, что модель основывает решения на физических признаках сигналов – концентрируясь на гармониках ротора и их обертонах, игнорируя низкоэнергетичные области спектра.

Дополнительно была проведена количественная оценка интерпретируемости через Deletion test (табл. 4). При маскировании наиболее «горячих» областей, выделенных Grad-CAM, уверенность модели падает существенно сильнее (-0.82 при 30 % удаленных областей), чем при случайном удалении (≈ 0.00). Это подтверждает информативность карт активации.

Таблица 4. Результаты Deletion test.

Метод	Δ Confidence (top-20 %)	Δ Confidence (top-30 %)
Grad-CAM	-0.14	-0.82
Random	≈ 0.00	≈ 0.00

Таким образом, модель достигает точности 0.97 (95 % ДИ: 0.94-0.99), обладает устойчивостью к шуму до -5 дБ, а также подтвержденной интерпретируемостью карт активации, что делает ее применимой в практических системах мониторинга воздушного пространства.

6. Обсуждение

Полученные результаты демонстрируют, что интеграция методов интерпретируемости в задачу акустической классификации малоразмерных летательных аппаратов повышает доверие к модели и облегчает ее практическое применение. Deletion-тест количественно подтвердил, что ключевые зоны, выделяемые Grad-CAM, действительно содержат критическую информацию для распознавания, а анализ IoU показал устойчивость этих карт между разными этапами обучения.

Таким образом, интерпретации, предоставляемые моделью, не являются артефактами конкретной инициализации или случайных шумов в данных, а опираются на реальные акустические признаки. Это делает предложенный подход более надежным для использования в системах мониторинга воздушного

пространства и обеспечения безопасности, где важно не только высокое качество классификации, но и прозрачность принимаемых решений.

7. Ограничения

Несмотря на высокую точность классификации, предложенный подход имеет ряд ограничений.

Во-первых, набор данных ограничен по объему (около 2 часов записей) и охватывает лишь девять классов сигналов. Для более надежной оценки и повышения обобщающей способности модели требуется расширение корпуса с включением большего числа типов малоразмерных аппаратов и разнообразных условий записи.

Во-вторых, акустические данные собирались в контролируемых условиях полигона с использованием двух микрофонов (МК-207 и МД-307). В реальной обстановке (городская среда, сильный ветер, транспортный шум) характеристики сигналов могут существенно отличаться.

В-третьих, добавленные шумы и аугментации лишь частично имитируют реальные условия. Модель не тестировалась на воздействие других источников помех, таких как пролет самолетов, громкие импульсные звуки или комбинированные шумовые сценарии. Кроме того, интерпретируемость проверялась преимущественно на лог-мел спектрограммах; прямое применение Grad-CAM к MFCC требует дополнительного обоснования.

Наконец, модель оптимизирована под фиксированное окно анализа (40×173 кадров), что может ограничивать ее применение для длинных непрерывных записей и потоковой обработки.

Заключение

Для обеспечения воспроизводимости экспериментов в работе подробно описаны архитектура модели, гиперпараметры обучения, а также методика подготовки и нормализации данных.

Использовались открытые и широко применяемые инструменты машинного обучения, что обеспечивает переносимость результатов на различные вычислительные платформы.

Все ключевые этапы – от формирования признаков до обучения и интерпретации модели – задокументированы и могут быть повторены при наличии аналогичных аудиозаписей.

Таким образом, полученные результаты воспроизводимы и могут служить основой для дальнейших исследований в области акустической классификации малоразмерных летательных аппаратов.

Финансирование: Работы проводились при финансовой поддержке Гранта в форме субсидии на проведение научных исследований Комитета по науке и инноватике Тульской области №15 от 21 июня 2024 г.

Литература

1. Rakshit H., Bagheri Zadeh P. A New Approach to Classify Drones Using a Deep Convolutional Neural Network //Drones. – 2024. – Т. 8. – №. 7. – С. 319. <https://doi.org/10.3390/drones8070319>.
2. Berg A. P., Zhang Q., Wang M. Y. 4,500 Seconds: Small Data Training Approaches for Deep UAV Audio Classification //arXiv preprint arXiv:2505.23782. – 2025. <https://doi.org/10.5220/0013462400003967>.
3. Berg A. P., Zhang Q., Wang M. Y. 15,500 Seconds: Lean UAV Classification Leveraging PEFT and Pre-Trained Networks //arXiv preprint arXiv:2506.11049. – 2025.
4. Yi W., Choi J. W., Lee J. W. Sound-based drone fault classification using multitask learning //arXiv preprint arXiv:2304.11708. – 2023.

5. Paszkowski W., Gola A., Świć A. Acoustic-Based Drone Detection Using Neural Networks—A Comprehensive Analysis //Advances in Science and Technology Research Journal. – 2024. – T. 18. – №. 1. – C. 36-47. <https://doi.org/10.12913/22998624/175863>.
6. Seidaliyeva U. et al. Advances and challenges in drone detection and classification techniques: A state-of-the-art review //Sensors. – 2023. – T. 24. – №. 1. – C. 125. <https://doi.org/10.3390/s24010125>.
7. Rahman M. H. et al. A comprehensive survey of unmanned aerial vehicles detection and classification using machine learning approach: Challenges, solutions, and future directions //Remote Sensing. – 2024. – T. 16. – №. 5. – C. 879. <https://doi.org/10.3390/rs16050879>.
8. Ni J., Zhou Z. Blind source separation and unmanned aerial vehicle classification using CNN with hybrid cross-channel and spatial attention module //Scientific Reports. – 2025. – T. 15. – №. 1. – C. 21905.
9. Zaman B., Al-Dulaimi A., Hussain M., Nguyen Q., Han S., Zhang Y. Audio Spectrogram Transformer with Convolutional Frontend for UAV Sound Classification in Low-Data Regime // Applied Sciences. – 2025. – Vol. 15, № 7. – Art. 1234.
10. Semenyuk S. Modern Technologies for Drone Detection and Classification: 2020–2025 Overview // Sensors. – 2025. – Vol. 25, № 3. – Art. 678.
11. Zeiler M. D., Fergus R. Visualizing and understanding convolutional networks //European conference on computer vision. – Cham : Springer International Publishing, 2014. – C. 818-833.https://doi.org/10.1007/978-3-319-10590-1_53
12. Zhou B. et al. Learning deep features for discriminative localization //Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – C. 2921-2929. <https://doi.org/10.1109/CVPR.2016.319>
13. Selvaraju R. R. et al. Grad-cam: Visual explanations from deep networks via gradient-based localization //Proceedings of the IEEE international conference on computer vision. – 2017. – C. 618-626. <https://doi.org/10.1109/ICCV.2017.74>

14. Chattopadhyay A. et al. Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks //2018 IEEE winter conference on applications of computer vision (WACV). – IEEE, 2018. – С. 839-847..
<https://doi.org/10.1109/WACV.2018.00097>.

Для цитирования:

Минаков Е.И., Гридин В.Н., Андронычев Д.Г., Солодовников В.И., Сычугов А.А., Французова Ю.В. Интерпретируемая акустическая классификация малоразмерных летательных аппаратов с использованием метода Grad-CAM. // Журнал радиоэлектроники. – 2025. – №. 11.
<https://doi.org/10.30898/1684-1719.2025.11.4>