

Адаптивная система распознавания по акустическим признакам на базе генеративно-сопоставительных сетей и обучения с подкреплением**Н. А. Верзун¹, М. О. Колбанев¹, А. Р. Салиева^{1✉}, Р. М. Мухтарбекова²**¹ Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В. И. Ульянова (Ленина), Санкт-Петербург, Россия² Кыргызский государственный технический университет им. И. Раззакова, Бишкек, Киргизия
✉ rustamovna.a3@gmail.com

Аннотация. Цель настоящего исследования состоит в разработке адаптивной системы акустической диагностики, способной выявлять редкие и критически важные неисправности промышленного оборудования на основе анализа аудиосигналов. В качестве методологической основы используется совмещение модифицированной архитектуры генеративно-сопоставительной сети WaveGAN, алгоритмов обучения с подкреплением Deep Q-Network и многоагентного анализа с динамической адаптацией. Такая интеграция позволяет не только генерировать физически достоверные синтетические сигналы для компенсации дисбаланса классов, но и адаптировать поведение системы в зависимости от характеристик акустической среды и типа оборудования. Эксперименты, проведенные с использованием набора данных MIMII, продемонстрировали высокие показатели точности (до 96 %) и полноты классификации на различных типах оборудования. Полученные результаты свидетельствуют о высокой устойчивости системы к внешним шумам и ее способности к своевременному обнаружению импульсных и переходных дефектов. Научная новизна работы заключается в синтезе генеративного подхода и мультиагентной архитектуры с контекстной адаптацией к производственным условиям, что обеспечивает комплексный и интерпретируемый анализ аудиосигналов. Практическая значимость обусловлена возможностью внедрения разработанной системы в состав интеллектуальных платформ мониторинга технического состояния оборудования в условиях реального производства.

Ключевые слова: акустическое распознавание, генеративно-сопоставительные сети, обучение с подкреплением, мультиагентные системы, предиктивное обслуживание, WaveGAN, промышленное оборудование, интеллектуальный мониторинг

Для цитирования: Адаптивная система распознавания по акустическим признакам на базе генеративно-сопоставительных сетей и обучения с подкреплением / Н. А. Верзун, М. О. Колбанев, А. Р. Салиева, Р. М. Мухтарбекова // Изв. СПбГЭТУ «ЛЭТИ». 2026. Т. 19, № 1. С. 15–24. doi: 10.32603/2071-8985-2026-19-1-15-24.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Scientific article

Adaptive Acoustic Recognition System Based on Generative Adversarial Networks Reinforcement Learning**N. A. Verzun¹, M. O. Kolbanev¹, A. R. Salieva^{1✉}, R. M. Mukhtarbekova²**¹ Saint Petersburg Electrotechnical University, Saint Petersburg, Russia² Razzakov Kyrgyz State Technical University, Bishkek, Kyrgyzstan
✉ rustamovna.a3@gmail.com

Abstract. The purpose of this study is to develop an adaptive acoustic diagnostic system capable of detecting rare critically important malfunctions of industrial equipment based on audio signal analysis. The combination of the modified architecture of the generative adversarial network WaveGAN, reinforcement learning algo-

rhythms Deep Q-Network multi-agent analysis with dynamic adaptation is used as a methodological basis. This integration allows not only to generate physically reliable synthetic signals to compensate for class imbalance, but also to adapt the behavior of the system depending on the characteristics of the acoustic environment the type of equipment. Experiments conducted using the MIMII dataset demonstrated high accuracy (up to 96 %) classification completeness on various types of equipment. The results obtained indicate the high stability of the system to external noise its ability to timely detect pulse transient defects. The scientific novelty of the work lies in the synthesis of a generative approach a multi-agent architecture with contextual adaptation to production conditions, which provides a comprehensive interpretable analysis of audio signals. The practical significance is due to the possibility of introducing the developed system into intelligent platforms for monitoring the technical condition of equipment in real-world production.

Keywords: acoustic recognition, generative-adversarial networks, reinforcement learning, multi-agent systems, predictive maintenance, WaveGAN, industrial equipment, intelligent monitoring

For citation: Adaptive Acoustic Recognition System Based on Generative Adversarial Networks Reinforcement Learning / N. A. Verzun, M. O. Kolbanev, A. R. Salieva, R. M. Mukhtarbekova // LETI Transactions on Electrical Engineering & Computer Science. 2026. Vol. 19, no. 1. P. 15–24. doi: 10.32603/2071-8985-2026-19-1-15-24.

Conflict of interest. The authors declare no conflicts of interest.

Введение. Современное промышленное производство характеризуется возрастающей сложностью технологических процессов, усложнением инфраструктуры оборудования и высокими требованиями к его надежности и безаварийной эксплуатации. В этом контексте особенно остро стоит задача своевременного обнаружения развивающихся дефектов, которые могут привести к непредвиденным простоям, авариям и, как следствие, значительным экономическим потерям. Традиционные методы технического обслуживания, основанные на жестких регламентах, уступают место стратегиям предиктивного (прогностического) обслуживания, опирающимся на непрерывный мониторинг состояния оборудования и анализ аномальных отклонений в поведении рабочих узлов [1]–[3].

Акустическая диагностика как метод неразрушающего контроля представляет собой одно из наиболее перспективных направлений в области мониторинга технического состояния оборудования. Она обладает рядом преимуществ: возможность дистанционного применения, высокая чувствительность к импульсным и переходным процессам. Однако применение акустических подходов в условиях реального промышленного производства связано с рядом вызовов. Во-первых, промышленная акустическая среда отличается высоким уровнем фонового шума (до 90 дБ), наличием вибрационных и электромагнитных помех, нестабильной акустической обстановкой [4], [5]. Во-вторых, реальные датасеты содержат крайне несбалансированную информацию – критические неисправности, как правило, составляют менее 5 % всех наблюдаемых случаев [6].

В-третьих, акустические сигналы различных видов оборудования значительно различаются, что требует от систем диагностики высокой адаптивности и универсальности [7].

В последние годы широкое распространение получают технологии искусственного интеллекта, в частности, методы машинного обучения, способные анализировать большие объемы данных и выявлять скрытые зависимости. Генеративно-состязательные сети (Generative Adversarial Networks, GAN), например модифицированные архитектуры WaveGAN, показали высокую эффективность при генерации синтетических аудиосигналов, способных компенсировать дисбаланс обучающей выборки без потери физических свойств сигналов [8], [9]. В то же время, методы обучения с подкреплением, включая Deep Q-Network (DQN), позволяют динамически адаптировать поведение диагностических систем под изменяющиеся условия внешней среды [10].

Научные исследования последних лет демонстрируют перспективность гибридных решений, в которых сочетаются генеративные модели, глубокие нейронные сети и адаптивные механизмы принятия решений [11]–[13]. Однако большинство существующих решений либо сосредоточены на ограниченном классе оборудования, либо требуют большого количества размеченных данных, что затрудняет их практическое применение.

Настоящие исследования направлены на разработку интеллектуальной системы акустической диагностики, способной эффективно функционировать в условиях ограниченных данных, сильного шума и разнообразия оборудования. Предлагается комплексный подход, сочетающий генеративное

усиление сигналов (на базе модифицированной архитектуры WaveGAN), мультиагентный анализ (спектральный, временной, контекстный) и механизм динамической адаптации весов агентов с использованием глубокого обучения с подкреплением.

Описание алгоритма диагностики. Алгоритм функционирования интеллектуальной системы акустической диагностики реализует последовательную обработку входных акустических сигналов с применением методов генеративного моделирования, мультиагентного анализа и механизмов обучения с подкреплением. Архитектура системы разработана для работы в условиях промышленного шума, ограниченного числа размеченных данных и необходимости высокой чувствительности к различным типам неисправностей. Общая схема алгоритма представлена на рис. 1.

1-й этап. Первичной задачей системы служит подготовка входных данных для дальнейшего анализа. Алгоритм реализует следующие операции:

– *адаптивная фильтрация шума:* применяется, например, фильтр Винера для подавления стационарных и нестационарных помех, характерных для производственной среды;

– *извлечение спектральных признаков:* аудиосигнал преобразуется в Mel-спектрограмму, обеспечивающую более устойчивое представление акустического сигнала в частотной области;

– *нормализация амплитуд:* выполняется для унификации динамического диапазона всех входных сигналов.

– *стохастические искажения:* реализуются путем применения аугментационных методов (например, временной маскировки или частотного шумового вмешательства) с целью повышения обобщающей способности модели на этапе обучения.

Генеративное усиление данных. Для компенсации дисбаланса обучающих выборок используется модуль генерации синтетических данных, основанный на модифицированной архитектуре WaveGAN, преимущества которой включают [13]:

– *генерация сигналов во временной области:* позволяет сохранить фазовую структуру и импульсные характеристики, важные для диагностики кратковременных дефектов;

– *частотно-селективная регуляризация:* применяется к синтетическим сигналам для приближения их характеристик к реальным данным в целевом частотном диапазоне;

– *физическая достоверность сигналов:* обеспечивается за счет использования улучшенной версии модели с градиентным штрафом (Wasserstein GAN with Gradient Penalty, WGAN-GP).

Архитектура включает два основных компонента – генератор и дискриминатор. На вход генератору подается случайный шумовой вектор z , который преобразуется в сигнал $x(t)$ – синтетический акустический фрагмент, соответствующий временной структуре реальных данных. Дискриминатор получает на вход как реальные, так и сгенерированные сигналы и возвращает вероятность принадлежности сигнала к реальному рас-

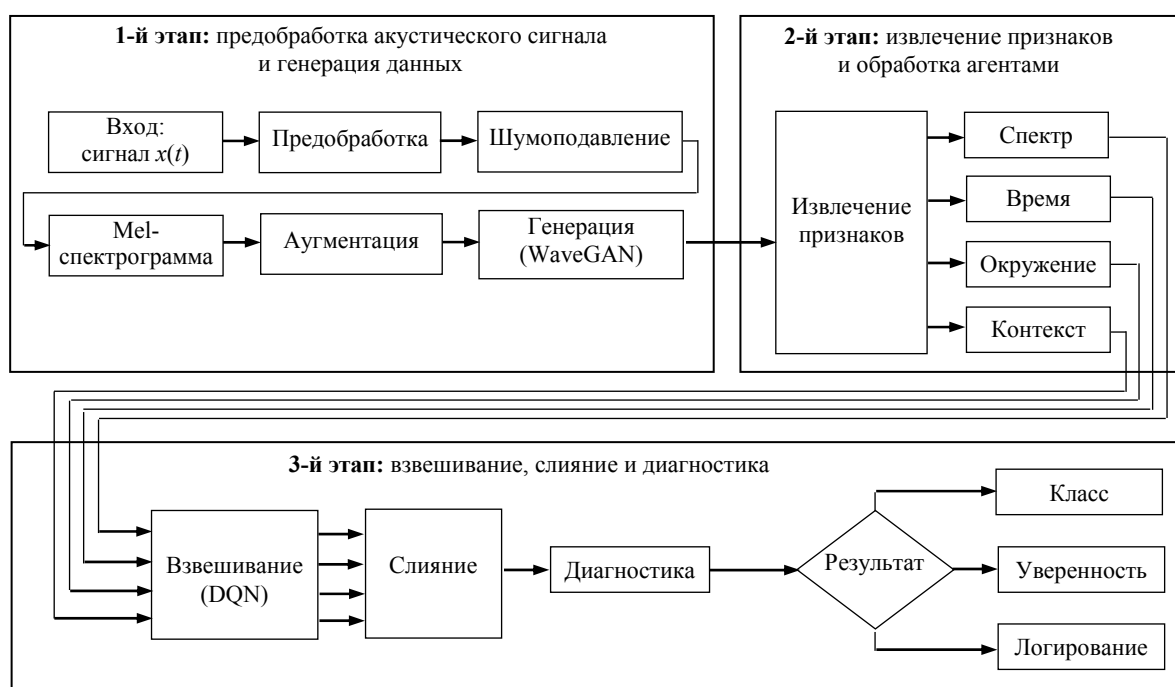


Рис. 1. Общая схема алгоритма диагностики
Fig. 1. General diagram of the diagnostic algorithm

пределению. Обучение осуществляется в состязательной форме, где цель генератора – «обмануть» дискриминатор, а задача дискриминатора – точно различать реальные и синтетические сигналы.

2-й этап. Обработка сигнала осуществляется параллельно четырьмя агентами, каждый из которых специализируется на определенном аспекте анализа:

- спектральный агент: анализирует частотные характеристики сигнала, включая пики, полосы пропускания и энергию по частотным полосам;

- временной агент: извлекает временные признаки – пики амплитуды, длительность импульсов, энтропию и автокорреляционные характеристики;

- агент окружения: использует сведения о внешнем контексте – например, идентификатор оборудования, текущую производственную фазу, интенсивность шума;

- контекстный агент: интегрирует историю предшествующих наблюдений, оценивает изменения в сигнале по времени и формирует временные зависимости.

Каждый агент формирует вектор вероятностей по классам (напр. «норма», «дефект 1», «дефект 2»).

3-й этап. На завершающем этапе результаты, полученные от агентов, передаются в модуль принятия решений с *динамическим взвешиванием*. Данный модуль:

- обучается с использованием глубокого Q -обучения;

- получает на вход векторы вероятностей от агентов, а также метаинформацию о сигнале (уровень шума, достоверность спектра);

- вычисляет оптимальное распределение весов между агентами для конкретного входного сигнала;

- формирует итоговое диагностическое решение на основе агрегированного вероятностного вектора с максимальной уверенностью.

Диагностический вывод. На выходе алгоритма формируется диагностическое заключение – метка класса технического состояния (например, «норма», «износ подшипников», «кавитация»). Результат дополнительно может включать коэффициент доверия, указывающий на уровень уверенности модели в принятом решении, а также диагностическую трассировку по каждому агенту, что способствует интерпретируемости алгоритма.

Математическая модель акустического сигнала. Математическое описание структуры акустического сигнала, регистрируемого в усло-

виях промышленной эксплуатации, подробно представлено в [14], включая аддитивную модель, методы спектрального анализа, описание шумовых компонентов и фильтрацию. В настоящей статье сохраняется базовая архитектура модели, но основное внимание уделяется новому компоненту – генеративному усилению данных с помощью модифицированной модели WaveGAN, разработанному [15] специально для условий промышленной диагностики.

Генеративное усиление данных. Архитектура модели WaveGAN. Для решения задачи недостатка размеченных акустических данных применяется модифицированная архитектура WaveGAN, адаптированная к специфике промышленной акустики. В отличие от моделей, оперирующих спектрограммами, WaveGAN работает напрямую с временными рядами, что обеспечивает:

1. Сохранение фазовой информации – важно для корректной генерации импульсных дефектов.

2. Высокое временное разрешение – позволяет воспроизводить кратковременные переходные процессы.

3. Физическую достоверность сигналов, соответствующих законам распространения звука.

Генератор $G(z)$ принимает на вход случайный вектор шума и возвращает реалистичный акустический сигнал $x = G(z)$, имитирующий реальные данные. Дискриминатор $D(x, \theta_D)$ оценивает вероятность того, что входной сигнал реален. Обучение происходит в соответствии с функцией потерь θ_D – это вектор весов и смещений нейронной сети дискриминатора, который оптимизируется в процессе обучения для точного различения реальных и сгенерированных данных [16], [17]:

$$\min_G \max_D E_{x \sim p_{\text{real}}} [\log D(x)] + \\ + E_{z \sim p(z)} [\log(1 - D(G(z)))],$$

где E_x – математическое ожидание по всем реальным примерам x ; E_z – математическое ожидание по всем векторам шума z ; p_{real} – распределение реальных данных, из которого взяты обучающие примеры; $p(z)$ – априорное распределение шума генератора, обычно многомерное гауссово или равномерное распределение.

Для повышения устойчивости и качества генерации используется улучшенная версия – Wasserstein GAN с градиентным штрафом (WGAN-GP). Соответствующая функция потерь:

$$\mathcal{L}_{\text{WGAN-GP}} = \mathbb{E}_{x \sim p_g} [D(x_g)] - \mathbb{E}_{x \sim p_r} [D(x_r)] + \\ + \lambda \mathbb{E}_{\hat{x} \sim p_{\hat{x}}} (\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2,$$

где x_g – сгенерированные данные из распределения p_g ; $D(x_g)$ – значение дискриминатора на сгенерированных данных; $\mathbb{E}[\dots]$ – математическое ожидание (усреднение по многим примерам); x_r – реальные данные из распределения p_r ; \hat{x} – точки, полученные линейной интерполяцией между реальным и сгенерированным примерами; $\nabla_{\hat{x}} D(\hat{x})$ – градиент дискриминатора по входу \hat{x} ; $\|\dots\|_2$ – норма градиента (длина вектора градиента); λ – коэффициент градиентного штрафа (gradient penalty weight). Это гиперпараметр, который контролирует вес (важность) компонента градиентного штрафа в общей функции потерь.

Экспериментальные исследования. *Источники акустических данных и объекты диагностики:* для оценки работоспособности предложенного алгоритма (на базе алгоритма была написана программа) использовались открытые акустические базы, содержащие аудиофрагменты работы промышленного оборудования в различных состояниях. Одним из основных источников данных стал открытый датасет MIMII Dataset (Malfunctioning Industrial Machine Investigation Inspection), опубликованный исследовательской группой Hitachi и доступный на платформе Zenodo. Данный набор включает более 1000 аудиофрагментов в формате WAV, записанных с различными типами оборудования, в том числе:

- центробежные насосы (pump);
- осевые и радиальные вентиляторы (fan);
- клапанные узлы (valve);
- винтовые компрессоры (slide compressor).

Каждый тип оборудования представлен в двух режимах: нормальном и аномальном, включая такие неисправности, как повышенное трение, разбалансировка, перегрев, кавитационные шумы и нестабильные вибрации. Записи осуществлялись в реальных производственных условиях с добавлением фонового шума, характерного для машиностроительных, энергетических и химических объектов. В библиотеке MIMII представлены, в частности, акустические сигналы от оборудования со следующими типами неисправностей:

– кавитация и заклинивание сопровождаются выраженными переходными процессами, что увеличивает нагрузку на временной агент и приводит к большей вычислительной стоимости;

– аэродинамические шумы и утечка воздуха имеют более устойчивый спектральный характер, что снижает общий объем обработки.

Программное обеспечение и инструменты анализа. Все эксперименты проводились в среде Python 3.8 с использованием следующих библиотек:

- Librosa – для загрузки и преобразования аудиосигналов (расчет спектрограмм, фильтрация, извлечение признаков);
- Scipy, Pytorch и NumPy – для цифровой обработки сигналов;
- TensorFlow 2.6 и Keras – для построения и обучения нейросетевых моделей;
- Scikit-learn – для расчета метрик классификации;
- Psutil – для мониторинга потребления системных ресурсов центрального процессорного устройства (ЦПУ) и оперативного запоминающего устройства (ОЗУ);
- Matplotlib, Seaborn – для визуализации результатов.

Модель тестировалась на компьютере с GPU NVIDIA RTX 3090 и 32 Гбайт оперативной памяти.

Для оценки качества акустического распознавания типа неисправности оборудования использовались метрики [18]: доля верных ответов, точность, полнота, F_1 -мера.

Результаты и обсуждение эксперимента.

Эксперимент включал следующие этапы:

Подготовка входных данных. Все аудиофайлы из набора MIMII были приведены к единому формату: длительность 5 с, частота дискретизации 16 кГц, моно. Далее проводилась нормализация амплитуды и извлечение признаков (Mel-спектрограммы и временные характеристики), а также формирование синтетических примеров с помощью генеративной модели WaveGAN. Генерация осуществлялась отдельно по каждому типу неисправности.

Организация обучающей и тестовой выборки. Данные были разделены в пропорции 80/20 на обучающую и тестовую части. Синтетические данные включались только в обучающую часть и не использовались при тестировании, чтобы избежать утечки данных.

Выход модели. Выходом модели служил вероятностный вектор по классам (норма, дефект 1, дефект 2), который затем преобразовывался в финальный диагноз на основе максимального значения. В процессе адаптации использовалась оценка доверия агентов, определяемая динамически через DQN.

Табл. 1. Показатели эффективности распознавания неисправностей
Tab. 1. Fault recognition efficiency indicators

Тип оборудования	Тип неисправности	Доля верных ответов, %	Точность, %	Полнота, %	F_1 -мера, %
Насос	Кавитация	95.2	93.7	96.5	95.1
Насос	Износ подшипников	94.6	91.9	95.8	93.8
Вентилятор	Разбалансировка	92.8	89.6	93.2	91.3
Вентилятор	Аэродинамический шум	91.1	88.2	92.4	90.2
Клапан	Заклинивание	96.4	94.5	97.8	96.1
Компрессор	Повышенное трение	93.7	90.8	94.1	92.4
Компрессор	Утечка воздуха	94.1	91.2	95.0	93.0

Табл. 2. Сравнение загрузки CPU, RAM и времени выполнения по компонентам
Tab. 2. Comparison of CPU, RAM execution time load by components

Компонент	Время обработки, мс	Использование ЦПУ, %	Использование ОЗУ, Мбайт
Предобработка сигнала	11.2	8.3	42
Агент спектрального анализа	26.4	17.5	238
Агент временного анализа	15.1	10.8	114
Агент анализа окружения	22.9	14.6	178
Агент контекстной оценки	30.3	16.7	192
Модуль динамического взвешивания	3.6	3.2	31
Итого	109.5	70.1	795

Анализ ошибок. Особое внимание уделялось случаям ложноотрицательных решений, так как они представляют наибольшую опасность при эксплуатации.

Анализ результатов эксперимента показал, что предложенная модель, основанная на ансамбле агентов с генеративным усилением и динамической адаптацией весов, обеспечивает хорошую точность распознавания неисправностей по акустическим данным в широком диапазоне типов оборудования. Табл. 1 представляет сводные значения метрик для различных классов промыш-

ленного оборудования. Как видно из таблицы, модель демонстрирует стабильные показатели F_1 -меры на уровне 90–96 %, что подтверждает ее высокую чувствительность.

Оценка вычислительной нагрузки. Для оценки применимости модели в условиях ограниченных вычислительных ресурсов была измерена нагрузка на компьютер всех компонентов (табл. 2). Рис. 2 отражает среднюю нагрузку на систему при обработке одного аудиофрагмента (длительность – 1 с, частота – 16 кГц).

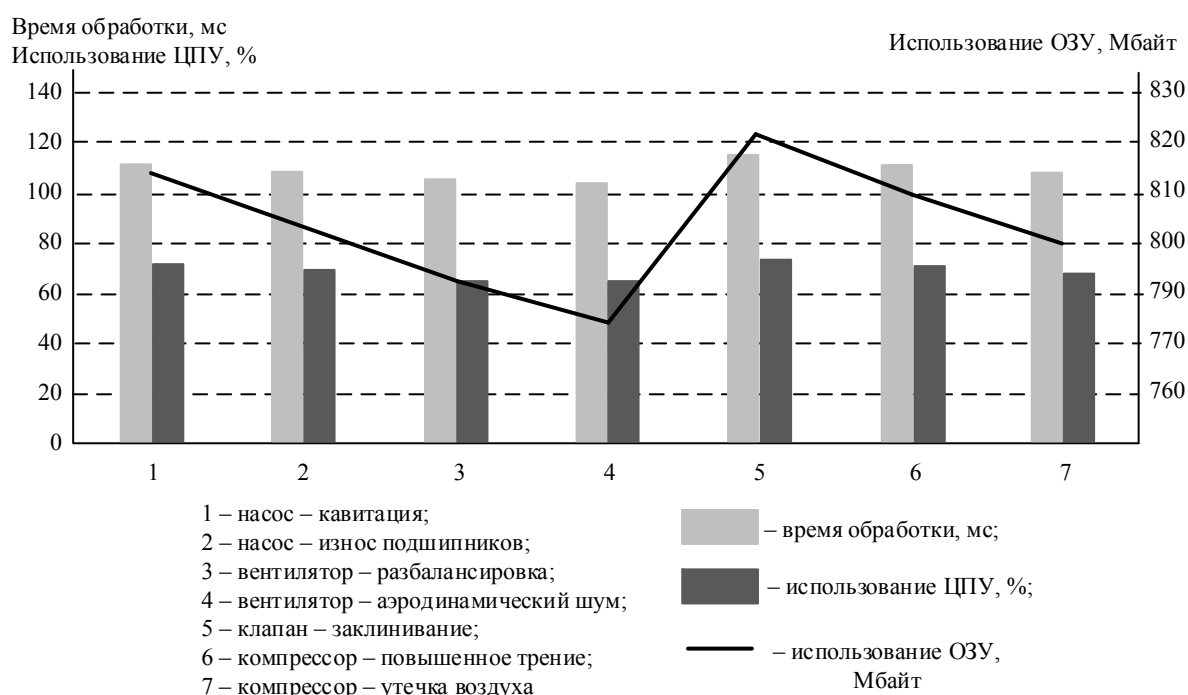


Рис. 2. Сравнение загрузки ЦПУ, ОЗУ на систему при диагностике различных типов неисправностей
Fig. 2. Comparison of CPU, RAM load on the system when diagnosing different types of faults

Модель обеспечивает анализ в режиме, близком к реальному времени (около 10 обработок в секунду) и может быть реализована на промышленных контроллерах или встраиваемых системах с поддержкой ускорителей искусственного интеллекта.

На рис. 2 показано сравнение загрузки ЦПУ, ОЗУ на систему при диагностике различных типов неисправностей.

Проведенные экспериментальные исследования подтвердили эффективность предложенного алгоритма акустической диагностики, особенно в условиях ограниченности обучающих данных и высокой шумовой загруженности. Применение генеративно-сопоставительной сети WaveGAN позволило достоверно синтезировать сигналы, сохраняющие ключевые физико-акустические характеристики, что способствовало увеличению полноты и точности распознавания различных типов неисправностей. Особенно значительное улучшение было зафиксировано для классов с ярко выраженной импульсной структурой – таких, как кавитация и заклинивание клапанов.

Многоагентная архитектура анализа продемонстрировала способность к дифференцированной обработке различных аспектов акустического сигнала. За счет разделения функциональности между спектральным, временным, контекстным агентами и агентом окружения удалось повысить интерпретируемость принимаемых системой решений, а также обеспечить более устойчивую реакцию на нестабильные внешние условия.

Ключевым элементом системы стал модуль динамического взвешивания, реализованный на основе алгоритма Deep Q-Network. Его использование обеспечило адаптацию модели к изменяющимся параметрам входных данных без необходимости повторного обучения всех агентов. Это свойство особенно актуально для применения в реальных производственных системах, где условия акустической среды могут меняться в непредсказуемом режиме.

Однако, несмотря на полученные положительные результаты, выявлены и некоторые ограничения. Так, точность классификации снижается при наличии слабовыраженных дефектов, чьи спектральные признаки слабо отличаются от фоновых шума. Кроме того, валидация алгоритма производилась на одном открытом наборе данных (MIMII), что ограничивает возможность прямой экстраполяции полученных результатов на иные производственные условия без дополнительной адаптации.

Заключение. В данной статье предложена и апробирована интеллектуальная система акустической диагностики промышленного оборудования, основанная на интеграции генеративно-сопоставительных сетей, обучения с подкреплением и мультиагентного анализа. Научная новизна исследования заключается в синтезе генеративного подхода (модифицированная архитектура WaveGAN), глубокого Q-обучения и мультиагентной архитектуры с контекстной адаптацией, что позволило реализовать гибкий и адаптивный механизм диагностики в условиях нестабильной акустической среды.

Разработанный алгоритм обеспечивает высокую точность и полноту распознавания неисправностей, устойчив к акустическим помехам и характеризуется высокой степенью интерпретируемости принимаемых решений. Практическая значимость разработанного решения заключается в возможности его интеграции в существующие системы предиктивного обслуживания, что обеспечит своевременное обнаружение критических неисправностей, минимизацию времени простоя и повышение надежности эксплуатации оборудования.

Перспективы дальнейших исследований связаны с расширением перечня типов диагностируемого оборудования, повышением устойчивости модели к слабовыраженным аномалиям, а также разработкой облегченной версии алгоритма для встраиваемых вычислительных платформ с ограниченными ресурсами.

Список литературы

1. Li X., Ding Q., Sun J.-Q. Remaining useful life estimation in prognostics using deep convolution neural networks // Reliability Engin. & System Safety. 2018. Vol. 172. P. 1–11. doi: 10.1016/j.ress.2017.11.021.
2. Nagdev A., Tarun G. A research study on unsupervised machine learning algorithms for fault detection in predictive maintenance // 2018 5th Intern. Conf. on Industrial Engin. and Appl. (ICIEA). Singapore: IEEE, 2018. P. 1–7. doi: 10.1109/IEA.2018.8387124.

3. Application of deep learning in fault diagnosis of rotating machinery / W. Jiang, Ch. Wang, J. Zou, Sh. Zhang // Proc. 2021. Vol. 9, no. 6. P. 1–22. doi: 10.3390/pr9060919.
4. A review on data-driven fault severity assessment in rolling bearings / M. Cerrada, R.-V. Sánchez, Ch. Li, F. Pacheco, D. Cabrera, J. V. de Oliveira, R. E. Vásquez // Mechan. Syst. and Signal Proc. 2018. Vol. 99. P. 169–196. doi: 10.1016/j.ymssp.2017.06.012.

5. Deep learning and Its applications to machine health monitoring: A survey / R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, R. X. Gao // J. of LaTeX Class Files. 2016. Vol. 14, no. 8. P. 1–14. URL: <https://arxiv.org/pdf/1612.07640> (дата обращения: 27.06.2025).
6. Hinton G. E., Osindero S., Teh Y.-W. A fast learning algorithm for deep belief nets // Neural Computation. 2006. Vol. 18, no. 7. P. 1527–1554. doi: 10.1162/neco.2006.18.7.1527.
7. Reinforcement learning for predictive maintenance: A systematic technical review / R. Siraskar, S. Kumar, S. Patil, K. Bhosale, S. Jaiswal // Artificial Intelligence Rev. 2023. Vol. 56. P. 12885–12947. doi: 10.1007/s10462-023-10468-6.
8. Improved parallel waveGAN vocoder with perceptually weighted spectrogram loss / E. Song, R. Yamamoto, M. Hwang, J.-S. Kim, O. Kwon, J.-M. Kim // 2021 IEEE Spoken Language Technol. Workshop (SLT). Shenzhen, China: IEEE, 2021. P. 470–476. doi: 10.1109/SLT48900.2021.9383549.
9. GAN-based data augmentation strategy for sensor anomaly detection in industrial robots / H. Lu, M. Du, K. Qian, X. He, K. Wang // IEEE Sensors J. 2021. Vol. 22, no. 18. P. 17464–17474. doi: 10.1109/JSEN.2021.3069452.
10. An introduction to deep reinforcement learning / V. François-Lavet, P. Henderson, R. Islam, M. G. Belle-mare, J. Pineau // Foundations and Trends in Machine Learning. 2018. Vol. 11, no. 3–4. P. 219–354. doi: 10.1561/22000000071.
11. Nalla N. R. Predictive maintenance in fleet management: Analyzing audio data for hazard detection // J. of Electr. Syst. 2025. Vol. 21, no. 1s. P. 132–136. doi: 10.52783/jes.8362.
12. Fault diagnosis of automobile power seat with acoustic analysis and retrained SVM based on smart-phone / X.-L. Huang, Zh. Teng, Q. Tang, Z. Yu, J. Hua, X. Wang // Measurement. 2022. Vol. 202, no. 4. Art. 111699. doi: 10.1016/j.measurement.2022.111699.
13. Self-supervised learning: A succinct review / M. Jindal, S. Arora, R. Garg, P. Madan, M. Kumar // Archives of Comp. Methods in Engin. 2023. Vol. 30, no. 4. P. 2761–2775. doi: 10.1007/s11831-023-09884-2.
14. Верзун Н. А., Колбанев М. О., Салиева А. Р. Многоагентный ансамблевый алгоритм акустического распознавания нарушений работоспособности автономного технологического оборудования // Информационно-управляющие системы. 2025. № 3. С. 14–24. doi: 10.31799/1684-8853-2025-3-14-24.
15. Yamamoto R., Song E., Kim J.-M. Parallel WaveGAN: A fast waveform generation model based on generative adversarial networks with multi-resolution spectrogram // Proc. ICASSP 2020. Barcelona, Spain: IEEE, 2020. P. 6199–6203. doi: 10.1109/ICASSP40776.2020.9053795.
16. Rall R. B. Vibration-based condition monitoring: Industrial, aerospace automotive applications. Chichester, United Kingdom: Wiley, 2011. 285 p. URL: <https://download.e-bookshelf.de/download/0000/5827/06/L-G-000582706-0015293164.pdf> (дата обращения: 07.07.2025).
17. Mizuta K., Koriyama T., Saruwatari H. Harmonic WaveGAN: GAN-Based speech waveform generation model with harmonic structure discriminator // Proc. Conf. Interspeech. Brno, Czech Republic: ISCA, 2021. P. 2192–2196. doi: 10.21437/Interspeech.2021-583.
18. Сравнение алгоритмов обучения с подкреплением и многоагентного ансамблевого алгоритма в задачах акустического распознавания / Н. А. Верзун, М. О. Колбанев, А. Р. Салиева, К. Н. Егоров // XXVIII Междунар. конф. по мягким вычислениям и измерениям (SCM'2025). СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2025. Т. 1. С. 411–416.

Информация об авторах

Верзун Наталья Аркадьевна – канд. техн. наук, доцент кафедры информационных систем СПбГЭТУ «ЛЭТИ».

E-mail: verzun.n@unecon.ru

<https://orcid.org/0000-0002-0126-2358>

Колбанев Михаил Олегович – д-р техн. наук, профессор кафедры информационных систем СПбГЭТУ «ЛЭТИ».

E-mail: mokolbanev@mail.ru

<https://orcid.org/0000-0003-4825-6972>

Салиева Аделина Рустамовна – аспирант гр. 3933, кафедра информационных систем СПбГЭТУ «ЛЭТИ».

E-mail: rustamovna.a3@gmail.com

Мухтарбекова Расита Мухтарбековна – аспирант кафедры метрологии и стандартизации. Кыргызский гос. технический университет имени И. Раззакова, пр. Чингиза Айтматова, 66/1, Первомайский р-н, г. Бишкек, 720044, Кыргызская Республика.

E-mail: rasita@kstu.kg

Вклад авторов:

Верзун Н. А. – общее руководство исследованием, постановка научной задачи, разработка концепции многоагентной системы, анализ результатов экспериментов, подготовка и редактирование текста статьи.

Колбанев М. О. – научная консультация, разработка математической модели и методологии исследования, валидация результатов, критический анализ и доработка текста статьи.

Салиева А. Р. – проведение программной реализации алгоритма, постановка и проведение вычислительных экспериментов, сбор и анализ данных, подготовка первоначального варианта текста статьи, визуализация результатов.

Мухтарбекова Р. М. – участие в проведении экспериментов, анализ литературных источников, подготовка обзора состояния вопроса, участие в обсуждении результатов.

References

1. Li X., Ding Q., Sun J.-Q. Remaining useful life estimation in prognostics using deep convolution neural networks // *Reliability Engin. & System Safety*. 2018. Vol. 172. P. 1–11. doi: 10.1016/j.res.2017.11.021.
2. Nagdev A., Tarun G. A research study on unsupervised machine learning algorithms for fault detection in predictive maintenance // 2018 5th Intern. Conf. on Industrial Engin. and Appl. (ICIEA). Singapore: IEEE, 2018. P. 1–7. doi: 10.1109/IEA.2018.8387124.
3. Application of deep learning in fault diagnosis of rotating machinery / W. Jiang, Ch. Wang, J. Zou, Sh. Zhang // *Proc.* 2021. Vol. 9, no. 6. P. 1–22. doi: 10.3390/pr9060919.
4. A review on data-driven fault severity assessment in rolling bearings / M. Cerrada, R.-V. Sánchez, Ch. Li, F. Pacheco, D. Cabrera, J. V. de Oliveira, R. E. Vásquez // *Mechan. Syst. and Signal Proc.* 2018. Vol. 99. P. 169–196. doi: 10.1016/j.ymssp.2017.06.012.
5. Deep learning and Its applications to machine health monitoring: A survey / R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, R. X. Gao // *J. of LaTeX Class Files*. 2016. Vol. 14, no. 8. P. 1–14. URL: <https://arxiv.org/pdf/1612.07640> (дата обращения: 27.06.2025).
6. Hinton G. E., Osindero S., Teh Y.-W. A fast learning algorithm for deep belief nets // *Neural Computation*. 2006. Vol. 18, no. 7. P. 1527–1554. doi: 10.1162/neco.2006.18.7.1527.
7. Reinforcement learning for predictive maintenance: A systematic technical review / R. Siraskar, S. Kumar, S. Patil, K. Bhosale, S. Jaiswal // *Artificial Intelligence Rev.* 2023. Vol. 56. P. 12885–12947. doi: 10.1007/s10462-023-10468-6.
8. Improved parallel waveGAN vocoder with perceptually weighted spectrogram loss / E. Song, R. Yamamoto, M. Hwang, J.-S. Kim, O. Kwon, J.-M. Kim // 2021 IEEE Spoken Language Technol. Workshop (SLT). Shenzhen, China: IEEE, 2021. P. 470–476. doi: 10.1109/SLT48900.2021.9383549.
9. GAN-based data augmentation strategy for sensor anomaly detection in industrial robots / H. Lu, M. Du, K. Qian, X. He, K. Wang // *IEEE Sensors J.* 2021. Vol. 22, no. 18. P. 17464–17474. doi: 10.1109/JSEN.2021.3069452.
10. An introduction to deep reinforcement learning / V. François-Lavet, P. Henderson, R. Islam, M. G. Belle-
mare, J. Pineau // *Foundations and Trends in Machine Learning*. 2018. Vol. 11, no. 3–4. P. 219–354. doi: 10.1561/22000000071.
11. Nalla N. R. Predictive maintenance in fleet management: Analyzing audio data for hazard detection // *J. of Electr. Syst.* 2025. Vol. 21, no. 1s. P. 132–136. doi: 10.52783/jes.8362.
12. Fault diagnosis of automobile power seat with acoustic analysis and retrained SVM based on smartphone / X.-L. Huang, Zh. Teng, Q. Tang, Z. Yu, J. Hua, X. Wang // *Measurement*. 2022. Vol. 202, no. 4. Art. 111699. doi: 10.1016/j.measurement.2022.111699.
13. Self-supervised learning: A succinct review / M. Jindal, S. Arora, R. Garg, P. Madan, M. Kumar // *Archives of Comp. Methods in Engin.* 2023. Vol. 30, no. 4. P. 2761–2775. doi: 10.1007/s11831-023-09884-2.
14. Verzun N. A., Kolbanov M. O., Salieva A. R. Mno-goagentnyj ansamblevyj algoritm akusticheskogo raspoznaniya narushenij rabotosposobnosti avtonomnogo tehnologicheskogo oborudovaniya // *Informacionno-upravljajushhie sistemy*. 2025. № 3. С. 14–24. doi: 10.31799/1684-8853-2025-3-14-24. (In Russ.).
15. Yamamoto R., Song E., Kim J.-M. Parallel WaveGAN: A fast waveform generation model based on generative adversarial networks with multi-resolution spectrogram // *Proc. ICASSP 2020. Barcelona, Spain: IEEE*, 2020. P. 6199–6203. doi: 10.1109/ICASSP40776.2020.9053795.
16. Rall R. B. Vibration-based condition monitoring: Industrial, aerospace automotive applications. Chichester, United Kingdom: Wiley, 2011. 285 p. URL: <https://download.e-bookshelf.de/download/0000/5827/06/L-G-000582706-0015293164.pdf> (дата обращения: 07.07.2025).
17. Mizuta K., Koriyama T., Saruwatari H. Harmonic WaveGAN: GAN-Based speech waveform generation model with harmonic structure discriminator // *Proc. Conf. Interspeech. Brno, Czech Republic: ISCA*, 2021. P. 2192–2196. doi: 10.21437/Interspeech.2021-583.
18. Srovnání algoritmov obučení s podkrep-
lením i mnogoagentnogo ansamblevogo algoritma v zadachah akusticheskogo rozpoznaniya / N. A. Verzun, M. O. Kolbanov, A. R. Salieva, K. N. Egorov // XXVIII Mezhdunar. konf. po mjagkim vychislenijam i izmerek-nijam (SCM'2025). SPb.: Izd-vo SPbGJeTU «LJeTI», 2025. T. 1. С. 411–416. (In Russ.).

Information about the authors

Natalia A. Verzun – Cand. Sci. (Eng.), Associate Professor of the Department of Information systems, Saint Petersburg Electrotechnical University.

E-mail: verzun.n@unecon.ru

<https://orcid.org/0000-0002-0126-2358>

Mikhail O. Kolbanev – Dr Sci. (Eng.), Professor of the Department of Information systems, Saint Petersburg Electrotechnical University.

E-mail: mokolbanev@mail.ru

<https://orcid.org/0000-0003-4825-6972>

Adelina R. Salieva – postgraduate student gr. 3933, Department of Information systems, Saint Petersburg Electrotechnical University.

E-mail: rustamovna.a3@gmail.com

Rasita M. Mukhtarbekova – postgraduate student of the Department of Metrology and Standardization. Kyrgyz State Technical University named after I. Razzakov, 66/1 Chyngyz Aitmatov Ave., Pervomaisky district, Bishkek, 720044, Kyrgyz Republic.

E-mail: rasita@kstu.kg

Author contribution statement:

Verzun N. A. – overall research supervision, formulation of the scientific problem, development of the multi-agent system concept, analysis of experimental results, preparation and editing of the article text.

Kolbanev M. O. – scientific consulting, development of the mathematical model and research methodology, validation of the results, critical analysis and revision of the article text.

Salieva A. R. – software implementation of the algorithm, design and conduct of computational experiments, data collection and analysis, preparation of the initial draft of the article, visualization of the results.

Mukhtarbekova R. M. – participation in conducting experiments, analysis of literary sources, preparation of the literature review, participation in the discussion of the results.

Статья поступила в редакцию 11.09.2025; принята к публикации после рецензирования 22.10.2025; опубликована онлайн 29.01.2026.

Submitted 11.09.2025; accepted 22.10.2025; published online 29.01.2026.
