

АЛГОРИТМЫ ПРЕДОБРАБОТКИ ЗВУКОВЫХ СИГНАЛОВ В СИСТЕМАХ НЕЙРОСЕТЕВОЙ ДИАГНОСТИКИ

Храпов Артемий Андреевич

студент, ФГБОУ ВО Московский государственный технологический университет «СТАНКИН»,
РФ, г. Москва

Аннотация. В статье рассмотрены этапы предобработки звуковых сигналов, применяемые перед подачей в нейросетевые модели для диагностики технического состояния станков. Описаны алгоритмы фильтрации, нормализации, оконного разбиения и формирования мел-спектрограмм, а также их влияние на устойчивость и точность классификации. Приведены аргументы в пользу использования адаптивной обработки и пояснено, как именно предобработка повышает качество нейросетевого анализа в шумовых условиях промышленного цеха.

Ключевые слова: акустическая диагностика, предобработка, мел-спектрограмма, нормализация, фильтрация, нейросети.

1. Введение.

Нейросетевые методы технической диагностики в последние годы активно внедряются в промышленность. Однако их эффективность зависит от качества входных данных, особенно в условиях шумного производственного фона. Сырые звуковые сигналы требуют тщательной предобработки, иначе модель будет воспринимать шумы как часть информативного сигнала. Именно поэтому разработка надёжного алгоритма фильтрации и преобразования звука является неотъемлемой частью всей интеллектуальной системы.

2. Адаптивная фильтрация. Для выделения полезного сигнала используется полосовой фильтр в диапазоне 50–14000 Гц, охватывающий основные частоты механических дефектов. В отличие от жёстких фильтров с фиксированной полосой, адаптивный фильтр изменяет параметры в зависимости от энергетики соседних фреймов, что особенно эффективно при нестабильных режимах работы оборудования. Применение таких фильтров позволяет уменьшить влияние высокочастотного фона, сохраняя информативные составляющие сигнала.

3. Нормализация сигнала

Фазонезависимая нормализация сводит амплитуду в интервал от -1 до 1, устраняя зависимость от расстояния до микрофона и других внешних факторов. Дополнительно вычитается скользящее среднее, что устраняет постоянную составляющую и помогает выделить именно колебательные элементы сигнала. Это обеспечивает унификацию амплитудных характеристик записей, что критически важно при обучении модели на данных, полученных из разных источников.

4. Формирование окон и спектрограмм

Сигнал разбивается на перекрывающиеся окна длительностью 25 мс с шагом 10 мс. Это позволяет захватывать как кратковременные микроудары, так и длительные изменения

спектра. Для каждого окна строится мел-спектрограмма на основе 128 фильтров, логарифм мощности которых отражает интенсивность в логарифмическом масштабе. Получившаяся матрица размером 128×992 адаптирована под архитектуру PaSST и позволяет использовать патч-разбиение 16×16.

Более того, мел-спектрограмма считается наиболее подходящей формой представления аудиосигнала для нейросетевой обработки, так как имитирует восприятие звука человеком и обеспечивает сохранение наиболее значимых для диагностики признаков. Дополнительно, при необходимости, сигнал может быть преобразован в MFCC-представление, что дополняет признаки и снижает вероятность ошибок классификации.

5. Проверка качества представления. Перед подачей в модель вычисляется энергетический профиль спектрограммы. Если профиль выходит за доверительный интервал, сигнал помечается на пересмотр. Такая проверка позволяет исключить артефакты, связанные с перегрузкой микрофона или внешними ударами, не относящимися к состоянию станка. Автоматизированные скрипты удаляют записи с нулевым или чрезмерно высоким уровнем шума, а также проводят сверку длительности сигнала с заданным стандартом.

На практике это позволяет существенно повысить качество обучающей выборки, исключив из неё случаи, потенциально способные привести к переобучению или нестабильности модели. Такая фильтрация перед обучением также повышает скорость сходимости и стабильность метрик при валидации.

6. Заключение

Комплексная предобработка звуковых сигналов — необходимое условие эффективной работы нейросетевой системы диагностики. Адаптивные фильтры, нормализация и корректное формирование спектрограмм повышают устойчивость модели к фоновым шумам и обеспечивают точное различие между штатным и аварийным режимами. Такая методика позволяет успешно применять трансформерные архитектуры даже в условиях тяжёлой производственной среды.

В дальнейшем планируется внедрение более интеллектуальных методов контроля качества — в том числе автоматическое определение аномалий в спектре и обучение вспомогательной модели для отбраковки шумовых примеров ещё до начала основного обучения. Это обеспечит ещё более высокую надёжность диагностических систем.

Список литературы:

1. Александров В. П. Методы обработки сигналов в системах технической диагностики. — М.: Машиностроение, 2020. — 248 с.
2. Васвани А., Шазир Н., Пармар Н. и др. Всё, что вам нужно — это внимание // Материалы NIPS. — 2017.
3. Гудфеллоу И., Бенжио Й., Курвиль А. Глубокое обучение. — Кембридж: MIT Press, 2016. — 775 с.
4. Документация PaSST [Электронный ресурс]. — Режим доступа: <https://passt.org/> (дата обращения: 20.05.2025).
5. Хуанг Х. и др. Обзор трансформерных моделей для аудиозадач // IEEE Transactions on Audio, Speech, and Language Processing. — 2021.