УДК 681.518.5, 004.89 DOI: 10.14529/cmse260306

НОВЫЙ МЕТОД АДАПТАЦИИ МОДЕЛЕЙ ДИАГНОСТИКИ СИНХРОННЫХ ЭЛЕКТРОДВИГАТЕЛЕЙ ПРИ ОТСУТСТВИИ ДАННЫХ О ДЕФЕКТЕ В ЦЕЛЕВОМ ДОМЕНЕ

© 2026 А.Л. Шестаков, Д.В. Галышев, В.А. Мартьянова, О.Л. Ибряева

Южно-Уральский государственный университет

(454080 Челябинск, пр. им. В.И. Ленина, д. 76)

E-mail: a.l.shestakov@susu.ru, galyshevdv@susu.ru, eremeevava@susu.ru, ibriaevaol@susu.ru

Поступила в редакцию: 30.09.2026

Среди подходов к диагностике неисправностей вращающегося оборудования в условиях доменного сдвига остается малоизученным сценарий ввода оборудования в эксплуатацию, когда можно собрать данные только о нормальном состоянии, а примеры целевых неисправностей полностью отсутствуют. Стандартные методы доменной адаптации в условиях неразмеченных данных здесь неприменимы, так как отсутствуют целевые дефекты для кластеризации, в то время как методы обнаружения аномалий на основе одного класса игнорируют ценные знания о неисправностях в исходном домене. Методы, использующие цифровые двойники или генеративно-состязательные сети, требуют дорогостоящего физического моделирования и сталкиваются с проблемой расхождения между синтетическими и реальными данными. Мы предлагаем новый подход к доменной адаптации FD-GAP (Fréchet Distance and Gap-based Penalty) в условиях, когда целевой домен содержит исключительно данные нормальной работы оборудования. В отличие от предыдущих работ, где адаптация ограничивается выравниванием распределений, мы переносим геометрическую границу класса соответствующего дефекту из исходного домена в целевой, используя расстояние Фреше и функцию зазора для отталкивания состояний с дефектом. Для компенсации стохастической нестабильности модели при различных случайных инициализациях вводится новый механизм ансамблевой фильтрации на основе расхождения, работающий исключительно с данными о нормальном состоянии целевого объекта, что обеспечивает надежность внедрения с заданным уровнем ложноположительных срабатываний.

Ключевые слова: FD-GAP, зазор, доменная адаптация, обучение на одном классе, межвитковые замыкания, синхронные электродвигатели, диагностика по токам статора, расстояние Фреше, ансамбль нейронных сетей, фильтрация нестабильных моделей.

ОБРАЗЕЦ ЦИТИРОВАНИЯ

Шестаков А.Л., Галышев Д.В., Мартьянова В.А., Ибряева О.Л. Новый метод адаптации моделей диагностики синхронных электродвигателей при отсутствии данных о дефекте в целевом домене // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2026. Т. 15, № 3. С. 117–142. DOI: 10.14529/cmse260306.

Введение

Вращающееся оборудование составляет основу промышленного производства. Его безотказная работа важна для непрерывности технологических процессов, а непредвиденные отказы приводят к существенным финансовым потерям из-за простоев и затрат на ремонт [1, 2]. Выявление ранних признаков неисправностей по временным рядам измеряемых сигналов остается сложной задачей. Поэтому в последние годы широкое распространение получили методы глубокого обучения, позволяющие автоматически извлекать диагностические признаки из многоканальных сигналов [1, 3].

Качество диагностических моделей, обученных на данных одного объекта, может снижаться при изменении распределения входных данных. В реальных условиях различия

между оборудованием и режимами его работы вызывают доменный сдвиг, затрудняющий перенос модели на новый объект. При этом целенаправленное воспроизведение дефектов на промышленных объектах затруднено или невозможно по соображениям безопасности, а собираемые данные преимущественно отражают штатный режим работы. Для компенсации доменного сдвига развиваются методы адаптации доменов (Domain Adaptation, DA) [1, 4, 5].

Наиболее изучен сценарий адаптации без учителя (Unsupervised Domain Adaptation, UDA). В классической постановке с общим набором классов целевой домен содержит неразмеченные данные штатной работы и при дефекте. Для выравнивания распределений применяются состязательное обучение и объединение информации из нескольких источников [6], псевдометки с учетом физических закономерностей [7], а также контрастивное обучение [8, 9]. Более широкие постановки, такие как частичная (Partial DA) и универсальная (Universal DA) адаптация, рассматривают несовпадение наборов классов источника и цели, включая отсутствие части исходных классов или появление неизвестных категорий [10–12]. Если при адаптации доступны только штатные данные цели, предположения методов о представленности нескольких целевых классов требуют пересмотра. В частности, механизмы кластеризации и формирования псевдометок, рассчитанные на наличие состояний с дефектом в целевой выборке, могут ошибочно разделять штатные данные на несколько классов. Отсутствие целевых примеров с дефектом следует отличать от отсутствия их разметки.

В сценарии адаптации без доступа к исходным данным (Source-Free Domain Adaptation, SFDA), например из-за требований конфиденциальности, используются только предварительно обученная модель и неразмеченные целевые данные. Разработаны методы на основе прототипов и псевдометок [13–15], контрастивного обучения [16], оценки неопределенности [17], адаптации на этапе тестирования [18] и универсальной адаптации на базе Transformer [19]. Для методов, использующих предположение о разнообразии целевых классов, наличие только штатных данных также меняет условия применения. В частности, стимулирование разнообразия предсказаний при максимизации информации может приводить к присвоению ложных меток дефекта. Однако такой риск связан с конкретной функцией потерь и предположениями метода, а не с постановкой SFDA в целом.

Сценарий обобщения доменов (Domain Generalization, DG) предполагает отсутствие целевых данных на этапе обучения. Для выделения признаков, сохраняющих диагностическую информативность в новых условиях, используются состязательное обучение [20, 21], аугментация [22–24], разделение доменных признаков и признаков дефекта [25, 26], выравнивание глобальных распределений [27], перенос отношений [28] и обучение без примеров целевых категорий (zero-shot learning) [29]. Валидация моделей и подбор гиперпараметров в DG выполняются на исходных доменах. Такая оценка не гарантирует сохранения качества на целевом объекте и затрудняет анализ влияния случайной инициализации на результат переноса, поскольку целевые данные для проверки недоступны.

Отдельное направление связано с генерацией синтетических данных с помощью цифровых двойников, метода конечных элементов и генеративно-состязательных сетей для компенсации нехватки образцов с дефектом [30–33]. В работе [34] цифровой двойник настраивается по штатным данным целевого объекта для генерации синтетических дефектов. Подходы, основанные на физическом моделировании, требуют знаний о механизмах отказов и могут быть вычислительно затратными. При использовании синтетических данных

также необходимо учитывать возможное расхождение между моделируемыми и реальными сигналами.

В данной работе предлагается метод FD-GAP для адаптации диагностической модели, обученной на размеченных данных источника соответствующих работе в штатном режиме и с дефектом, к целевому оборудованию, для которого на этапах адаптации, валидации и калибровки доступны только данные штатной работы. Метод сочетает выравнивание латентных распределений штатных состояний с помощью расстояния Фреше и фнукции зазора, отталкивающий примеры с дефектом источника от области штатных состояний. Перенос разделяющей границы опирается на предположение о сходстве физических механизмов дефектов и их влияния на признаки в обоих доменах. Дополнительно выполняются пересчет статистик слоев нормализации методом AdaBN и калибровка решающего порога по штатным данным цели.

Отсутствие целевых примеров с дефектом не позволяет напрямую оценивать качество их обнаружения при выборе модели. Для снижения зависимости результата от случайной инициализации используется ансамбль независимо обученных моделей с медианным агрегированием логитов (необработанные предсказания модели). Фильтрация по расхождению предсказаний на штатных данных цели позволяет исключать модели, отклоняющиеся от большинства ансамбля. Такая проверка оценивает согласованность моделей, но не заменяет валидацию на целевых дефектах. Калибровка порога задает целевой уровень частоты ложных срабатываний (False Positive Rate, FPR) по отдельной штатной выборке и сама по себе не гарантирует сохранения этого уровня на будущих данных.

Статья организована следующим образом. В разделе 1 приведена формальная постановка задачи. В разделе 2 описан предлагаемый метод, включая составную функцию потерь, Фреше-выравнивание, функцию зазора и процедуру ансамблирования. В разделе 3 представлены экспериментальные результаты для синхронного двигателя. В разделе 4 обсуждаются результаты и ограничения метода.

1. Постановка задачи

В данной статье мы рассматриваем адаптацию модели диагностики в условиях, когда на целевом оборудовании отсутствуют данные с дефектом — доступны только данные штатного (нормального) состояния. Такая постановка отражает практический сценарий развертывания диагностической системы на новом оборудовании, когда сбор данных с дефектами затруднен или невозможен на этапе ввода в эксплуатацию.

Пусть исходный домен $\mathcal{D}_S$ содержит размеченный набор данных $\mathcal{S} = \{(x_i, y_i)\}_{i=1}^{N_S}$ объемом N_S, где $y_i \in \{0, 1\}$, $x_i \in \mathcal{X} \subseteq \mathbb{R}^d$ — d-мерный вектор признаков i-го объекта. Метка $y_i = 0$ соответствует штатному режиму работы, $y_i = 1$ — наличию дефекта. Целевой домен $\mathcal{D}_T$ содержит выборку $\mathcal{T} = \{x_j\}_{j=1}^{N_T}$, состоящую только из данных штатного режима.

Требуется обучить классификатор $f_\theta : \mathcal{X} \to [0, 1]$, способный обнаруживать дефекты в целевом домене, не имея доступа к целевым примерам с дефектом на этапах обучения и адаптации.

Классификатор f_θ имеет составную структуру $f_\theta(x) = \sigma \leqslant ft(h_w \leqslant ft(g_\phi(x))$, где:

- $g_\phi : \mathcal{X} \to \mathbb{R}^k$ — нелинейный экстрактор признаков (энкодер), отображающий исходные данные в k-мерное латентное (скрытое) пространство;
- $h_w : \mathbb{R}^k \to \mathbb{R}$ — многослойный перцептрон, преобразующий латентное представление $z = g_\phi(x)$ в скалярную оценку $\ell(x) = h_w(z)$ (логиты);

- $\sigma(\cdot)$ — сигмоидальная функция активации $\sigma(\ell) = \frac{1}{1+e^{-\ell}}$, преобразующая выход последнего линейного слоя перцептрона в вероятность.

Такое разделение модели на энкодер и классифицирующий модуль необходимо для реализации предлагаемого метода, поскольку механизмы адаптации применяются именно к латентным представлениям $z = g_\phi(x)$.

Отметим, что в классической постановке адаптации доменов без учителя (unsupervised domain adaptation) обычно предполагается доступность неразмеченных целевых данных, покрывающих все классы. В этом случае адаптация решает задачу отсутствия меток, а не отсутствия целого класса. Полное отсутствие примеров с дефектом в $\mathcal{T}$ не только усложняет задачу адаптации, но и делает невозможной оценку точности модели по целевому классу в процессе обучения.

Как будет показано в разделе 3, качество отдельного запуска существенно варьируется в зависимости от случайной инициализации. Это наблюдение мотивирует переход к ансамблевой схеме, описанной в разделе 2.5 и являющейся неотъемлемой частью предлагаемого метода.

2. Метод

2.1. Общая схема метода FD-GAP

Метод **FD-GAP** состоит из четырех последовательных стадий, схематично изображенных на рис. 1:

1. **Предобучение.** На этом этапе модель $f_\theta(x) = \sigma \leqslant ft(h_w(g_\phi(x))$ обучается как стандартный классификатор на размеченных данных источника $\mathcal{S}$, минимизируя бинарную кросс-энтропию (Binary Cross Entropy Loss BCE-Loss) между предсказаниями и истинными метками классов (штатный режим / дефект). Цель — научиться разделять классы в исходном домене.
2. **Адаптация.** Параметры модели оптимизируются путем минимизации составной функции потерь, включающей три компонента: (1) ошибку классификации на размеченных данных источника, предотвращающую деградацию разделяющей способности модели; (2) ошибку выравнивания распределений штатных состояний источника и цели в латентном пространстве; (3) функцию зазора, отталкивающую примеры с дефектом источника от границы кластера с данными штатной работы.
3. **Постобработка и калибровка порога.** Статистики слоев нормализации пересчитываются на данных штатного режима целевого домена без обновления весов модели (метод AdaBN [35]), а решающий порог классификации калибруется на данных штатной работы целевого домена под заданный уровень FPR. Детали калибровки порога приведены в разделе 2.5.
4. **Ансамблирование.** Процедура этапов 1–3 повторяется K раз (в экспериментах $K = 30$) с различными случайными инициализациями весов, после чего предсказания моделей агрегируются через медиану логитов. Фильтр, оценивающий согласие каждой модели с большинством ансамбля на целевых данных штатной работы (раздел 2.5), может отклонить часть моделей. Такой подход компенсирует невозможность прямой валидации отдельной модели по целевому дефектному классу.

Разделение модели на энкодер и классифицирующий модуль принципиально для предлагаемого метода, т.к. компоненты выравнивания распределений (раздел 2.3) и переноса границы дефектов (раздел 2.4) применяются к латентным представлениям $z = g_\phi(x)$.

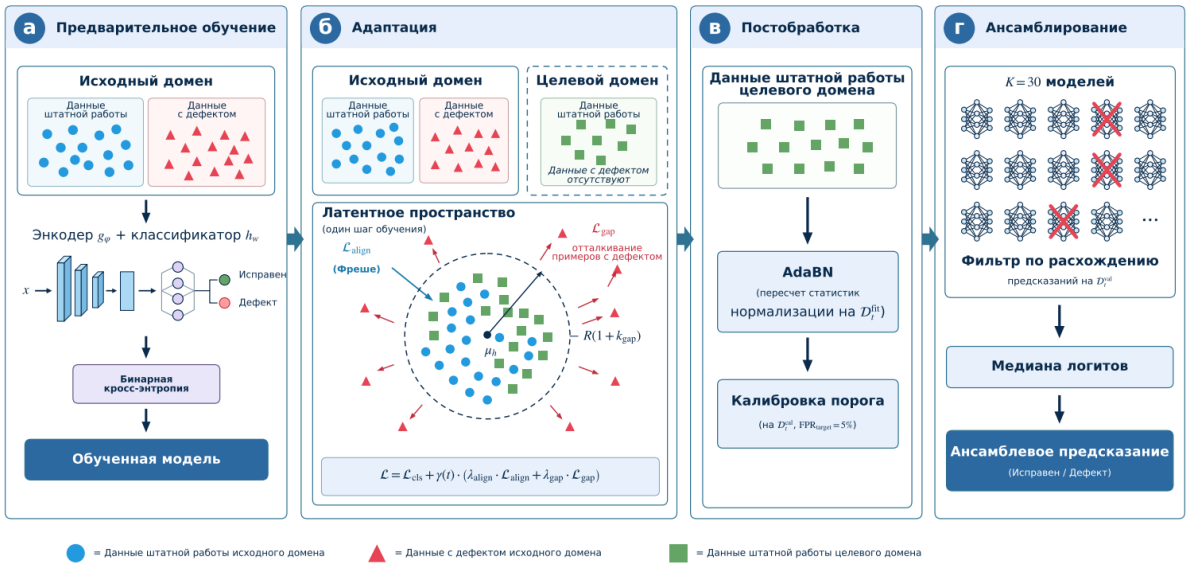


Рис. 1. Общая схема метода FD-GAP

На рис. 2 показана конкретная реализация энкодера g_ϕ и классификатора h_w для задачи диагностики электродвигателей (подробнее в разделе 3). Входом модели служит матрица признаков размером $F \times 3$, где F — число признаков, выделенных из сигнала тока, а 3 — количество фаз.

Модель построена на базе сверточной архитектуры с ядрами 3×3 и последовательным уменьшением пространственной размерности (stride $(2, 1)$). Использование сверток позволяет улавливать локальные связи между признаками и каналами (например, между фазами тока). Применение пакетной нормализации после каждого сверточного слоя необходимо для последующего использования метода AdaBN (раздел 2.1), который пересчитывает статистику нормализации на целевых данных без обновления весов. Кроме того, прореживание ($p = 0.3$ после слоев с уменьшением размерности и $p = 0.5$ в классификаторе) используется как регуляризатор, предотвращающий переобучение и повышающий обобщающую способность модели.

Отметим, что приведенная на рис. 2 архитектура является частным случаем, выбранным специально для формата наших экспериментальных данных. Однако сам предложенный метод универсален: для анализа данных другого типа достаточно лишь заменить энкодер на соответствующую архитектуру.

2.2. Составная функция потерь на этапе адаптации

На стадии адаптации, в отличие от предобучения, где используется только BCE-Loss, параметры модели оптимизируются путем минимизации составной функции потерь:

$$\mathcal{L} = \mathcal{L}_{\text{cls}} + \gamma(e) \cdot \big(\lambda_{\text{align}} \cdot \mathcal{L}_{\text{align}} + \lambda_{\text{gap}} \cdot \mathcal{L}_{\text{gap}}\big), \tag{1}$$

где $\mathcal{L}_{\text{cls}}$ — ошибка классификации на размеченных данных источника (предотвращает деградацию разделяющей способности модели); $\mathcal{L}_{\text{align}}$ — ошибка выравнивания распределений штатных состояний источника и цели в латентном пространстве (раздел 2.3); $\mathcal{L}_{\text{gap}}$ — штраф, отталкивающий состояния соответствующие дефекту источника от границы кластера данных штатной работы (раздел 2.4).

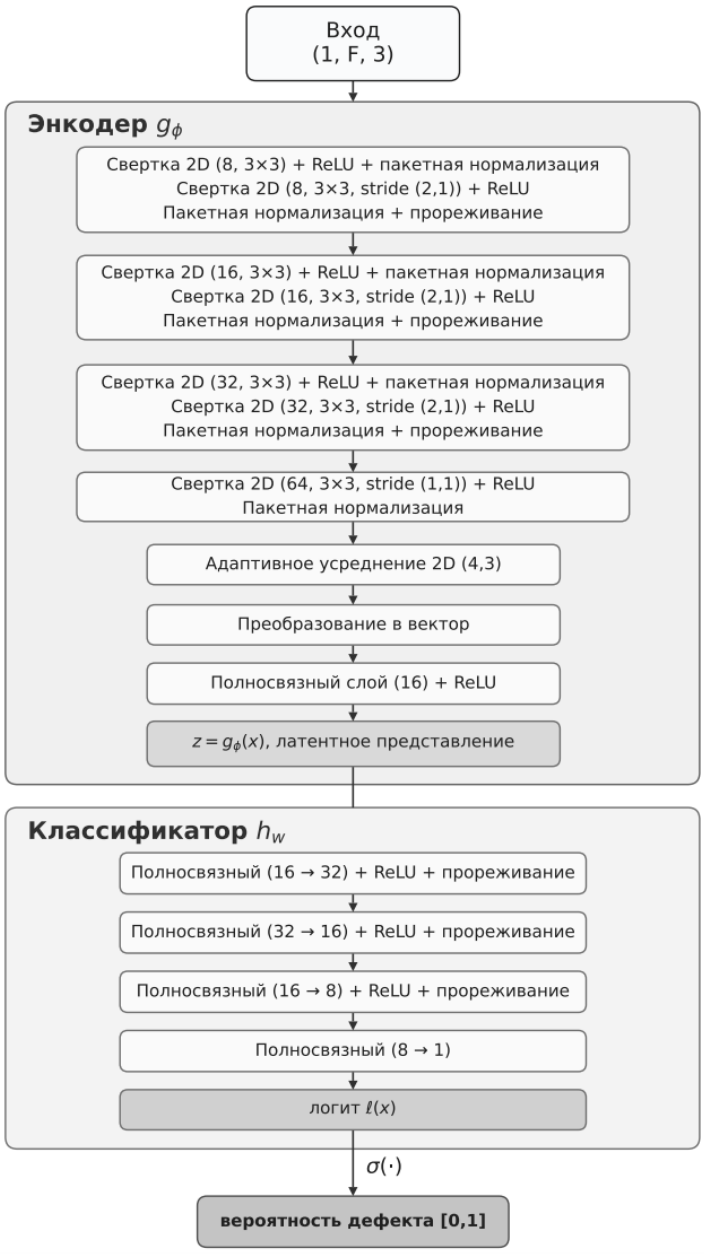


Рис. 2. Архитектура модели. Энкодер g_ϕ извлекает k-мерные латентные признаки ($k = 16$ в экспериментах), а Классификатор h_w преобразует их в логит $\ell(x)$ для вычисления вероятности дефекта

Коэффициент $\gamma(e)$ обеспечивает линейное наращивание вклада адаптационных компонентов от нуля до единицы в течение первых T_{warmup} эпох (в экспериментах $T_{\text{warmup}} = 5$):

$$\gamma(e) = \min\Big(1,\ \frac{e}{T_{\text{warmup}}}\Big), \qquad e = 1, 2, \dots \tag{2}$$

Такая схема позволяет плавно включить компоненты выравнивания и зазора, чтобы не разрушить представления признаков, уже сформированные на этапе предобучения.

Совместная оптимизация разнородных по масштабу и динамике слагаемых (1) без дополнительных мер склонна к неустойчивости: один из компонентов может доминировать

над остальными или, наоборот, перестать вносить вклад в градиент. Для предотвращения этого применяется адаптивная нормировка функции ошибки.

Каждый компонент функции ошибки делится на оценку его типичного значения s, которая обновляется в процессе обучения. Нормировка ошибки для задачи классификации имеет вид:

$$\bar{\mathcal{L}}_{\text{cls}}^{(n)} = \frac{\mathcal{L}_{\text{cls}}^{(n)}}{s_{\text{cls}}^{(n)}}, \qquad s_{\text{cls}}^{(n)} = \max\Big(\rho\, s_{\text{cls}}^{(n-1)} + (1-\rho)\,\mathcal{L}_{\text{cls}}^{(n)},\ s_{\text{cls}}^{\min}\Big), \tag{3}$$

где n — номер шага оптимизации, $\rho = 0.95$ — коэффициент экспоненциального сглаживания (Exponential Moving Average, EMA), а $s_{\text{cls}}^{\min} = 10^{-3}$ — нижняя граница масштаба, предотвращающая деление на чрезмерно малые значения. Аналогичная нормировка применяется ко всем остальным компонентам ($\mathcal{L}_{\text{mean}}$, $\mathcal{L}_{\text{cov}}$, $\mathcal{L}_{\text{gap}}$), но с нижней границей $s^{\min} = 10^{-2}$.

На первых $N_{\text{init}} = 10$ шагах обучения оценка s вычисляется как среднее арифметическое наблюдаемых значений функции ошибки для получения надежной начальной оценки, после чего переключается на экспоненциальное сглаживание с коэффициентом $\rho = 0.95$.

Для предотвращения переобучения в процессе адаптации применяется механизм ранней остановки (early stopping). Критерием остановки служит сглаженная сумма ошибок классификации на валидационных выборках источника и цели: если эта оценка не улучшается в течение заданного числа эпох (в наших экспериментах patience = 30), обучение прерывается, а модель возвращается к состоянию с наилучшей валидационной оценкой.

2.3. Фреше-выравнивание распределений

Для выравнивания распределений штатных состояний источника и цели в латентном пространстве мы используем расстояние Фреше между аппроксимирующими гауссовскими распределениями. Это расстояние, также известное как 2-е расстояние Вассерштейна между гауссианами [36], естественным образом учитывает как положение облаков точек (через средние значения), так и их форму (через ковариационные структуры).

Пусть μ_s, Σ_s и μ_t, Σ_t — выборочные средние и ковариации признаков штатного режима источника и цели соответственно, вычисленные в k-мерном латентном пространстве. Тогда функция потерь выравнивания имеет вид:

$$\mathcal{L}_{\text{align}} = \mathcal{L}_{\text{mean}} + \mathcal{L}_{\text{cov}}, \tag{4}$$

$$\mathcal{L}_{\text{mean}} = \frac{1}{k}\|\mu_s - \mu_t\|_2^2, \tag{5}$$

$$\mathcal{L}_{\text{cov}} = \frac{1}{k}\Big(\operatorname{Tr}(\Sigma_s) + \operatorname{Tr}(\Sigma_t) - 2\operatorname{Tr}\big(\sqrt{\sqrt{\Sigma_s}\,\Sigma_t\,\sqrt{\Sigma_s}}\big)\Big). \tag{6}$$

Слагаемые (5) и (6), где $\operatorname{Tr}(\cdot)$ обозначает след матрицы, представляют собой квадрат 2-го расстояния Вассерштейна между гауссовыми распределениями $\mathcal{N}(\mu_s, \Sigma_s)$ и $\mathcal{N}(\mu_t, \Sigma_t)$. Деление на размерность латентного пространства k обеспечивает инвариантность масштаба функции ошибки к выбору этой размерности.

Ковариации оцениваются с регуляризацией Ледуа—Вольфа [37] — стандартным методом улучшения обусловленности ковариационных матриц при малом объеме выборки. Вычисление матричного квадратного корня, входящего в формулу (6), выполняется с помощью итераций Ньютона—Шульца [38, 39].

2.4. Функция зазора: перенос границы класса дефекта без целевых примеров

Главная идея компонента зазора — перенести геометрическую границу между штатным режимом и дефектом из исходного домена в целевой, не имея ни одного примера целевого дефекта. Интуиция метода проиллюстрирована на рис. 3: пока кластеры штатного режима источника и цели разнесены, граница, построенная вокруг источника, не применима к целевому домену. После Фреше-выравнивания (раздел 2.3) оба кластера штатного режима занимают близкую область общего пространства. В результате построенная вокруг них граница становится валидной и для (невиданных) целевых дефектов — в предположении, что дефекты в обоих доменах вызваны схожими физическими механизмами и смещают признаки в похожем направлении относительно своего штатного состояния.

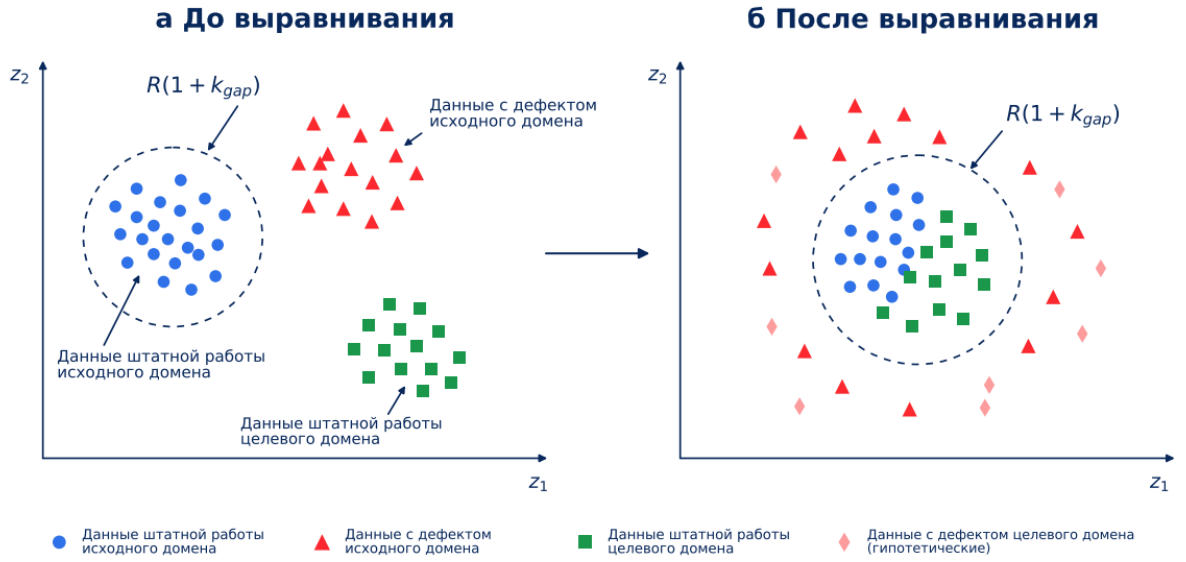


Рис. 3. Схематическая иллюстрация работы функции зазора в 2D-проекции латентного пространства

На каждом шаге адаптации для каждого объекта x вычисляем его признак в латентном пространстве $z = g_\phi(x)$ и нормируем по L2-норме: $\tilde{z} = z/\|z\|_2$.

Пусть $\tilde{Z}_h$ — множество нормированных признаков штатного режима источника в текущей мини-выборке. На каждом шаге сначала вычисляем локальные оценки центра и радиуса по батчу:

$$\mu_h^{\text{batch}} = \text{mean}\big(\tilde{Z}_h\big), \tag{7}$$

$$R^{\text{batch}} = \text{Quantile}_{0.95}\big(\|\tilde{Z}_h - \mu_h^{\text{batch}}\|_2\big), \tag{8}$$

а затем обновляем EMA оценки:

$$\mu_h^{(n)} = \beta\,\mu_h^{(n-1)} + (1-\beta)\,\mu_h^{\text{batch}}, \tag{9}$$

$$R^{(n)} = \beta\,R^{(n-1)} + (1-\beta)\,R^{\text{batch}}, \tag{10}$$

где n — номер шага оптимизации, $\beta = 0.9$ — коэффициент EMA. Радиус R соответствует 95-му процентилю расстояний от локального центра до точек штатного режима в батче.

Нормирование устраняет зависимость R и коэффициента запаса k_{gap} от масштаба выходных значений энкодера.

Функция зазора отталкивает состояния соответствующие дефекту источника от сферы радиуса $R\,(1 + k_{\text{gap}})$ с центром в μ_h:

$$\mathcal{L}_{\text{gap}} = \mathbb{E}_{x\sim\mathcal{D}_s^{\text{defect}}}\Big[\max\big(0,\ R\,(1 + k_{\text{gap}}) - \|\tilde{z}(x) - \mu_h\|_2\big)\Big], \tag{11}$$

где $k_{\text{gap}} = 0.5$ — коэффициент запаса, расширяющий защитную зону сверх естественного разброса штатных данных. Радиус R определяется данными как 95-й процентиль расстояний от центра до точек штатного режима и отражает их естественный разброс. Использование 95-го процентиля вместо максимума расстояний делает оценку радиуса устойчивой к редким выбросам среди данных штатной работы.

В формуле (11) используется классическая функция потерь hinge loss (аналогичная применяемой в One-Class SVM и Deep SVDD): для каждой точки соответствующей дефекту вычисляется расстояние до центра кластера штатных данных, и если оно меньше порогового значения $R(1 + k_{\text{gap}})$, Функция ошибки принимает положительное значение, пропорциональное «глубине» проникновения точки внутрь сферы, если точка находится вне сферы, вклад в итоговую функцию ошибки равен нулю.

Оптимальное значение параметра $k_{\text{gap}} = 0.5$ было установлено эмпирически в ходе предварительных экспериментов на валидационных выборках.

Полная процедура обучения одной модели приведена в алг. 1.

2.5. Ансамблирование и фильтрация моделей (Фильтр согласованности)

В постановке нашей задачи целевые данные с дефектом отсутствуют, поэтому невозможно оценить качество отдельной адаптированной модели: нет критерия, по которому один запуск можно признать успешным, а другой — нет. При этом стохастичность обучения (случайная инициализация весов, порядок обхода данных, выборка мини-батчей) приводит к тому, что разные запуски одной и той же конфигурации дают существенно разные результаты (раздел 3). Для получения надежного и воспроизводимого предсказания мы обучаем ансамбль из $K = 30$ моделей с различными случайными инициализациями, каждая из которых обучается по процедуре, описанной в алг. 1 (раздел 2.4). Модели, чьи предсказания на неразмеченных данных штатной работы из целевого домена наиболее сильно расходятся с медианым предсказанием ансамбля, исключаются из финального голосования.

Агрегирование предсказаний. Итоговая вероятность принадлежности объекта x к классу дефекта вычисляется как значение сигмоиды от медианы логитов (значений на выходе классификатора до применения функции активации) всех K моделей ансамбля:

$$\hat{p}(x) = \sigma\Big(\operatorname{median}_{k=1}^{K}\ \ell_k(x)\Big), \tag{12}$$

где $\ell_k(x)$ — логит k-й модели. Агрегация выполняется именно в пространстве логитов, а не вероятностей, из-за ограничений машинной точности формата float32. При больших по модулю значениях логитов функция сигмоиды округляется до предельных значений 0 или 1. Если вычислять медиану уже преобразованных вероятностей, она также примет это предельное значение, что приведет к потере информации о различиях в степени уверенности отдельных моделей. Применение сигмоиды однократно, после нахождения медианы логитов, позволяет сохранить эту информацию и обеспечить численную устойчивость процедуры агрегирования.

Алгоритм 1. Алгоритм обучения одной модели FD-GAP

Require: Исходный домен $\mathcal{D}_s = \{(x_i^s, y_i^s)\}$, целевые данные штатной работы $\mathcal{D}_t^{\text{fit}}$ и $\mathcal{D}_t^{\text{val}}$, калибровочные данные $\mathcal{D}_t^{\text{cal}}$

Предобработка: Предобучить базовую модель f_θ на $\mathcal{D}_s$ (минимизация BCE-Loss).
Инициализировать параметры оптимизатора и адаптивный нормализатор функции ошибки (раздел 2.2).
for эпоха $e = 1, \ldots, T$ **do**
  $\gamma(e) \leftarrow \min(1, e/T_{\text{warmup}})$
  for мини-батч источника $(x^s, y^s) \sim \mathcal{D}_s$ и независимый мини-батч цели $x^t \sim \mathcal{D}_t^{\text{fit}}$ **do**
    $\mathcal{L}_{\text{cls}} \leftarrow \text{BCE}(f_\theta(x^s), y^s)$
    $\mathcal{L}_{\text{align}} \leftarrow$ Фреше-расстояние между $g_\phi(x^s_{\text{healthy}})$ и $g_\phi(x^t)$ (раздел 2.3)
    Вычислить локальные $\mu_h^{\text{batch}}, R^{\text{batch}}$ по $g_\phi(x^s_{\text{healthy}})$ и обновить глобальные EMA-оценки μ_h, R (раздел 2.4)
    $\mathcal{L}_{\text{gap}} \leftarrow$ по формуле (11) на $g_\phi(x^s_{\text{defect}})$ с использованием обновленных μ_h, R
    Применить адаптивную нормировку ко всем компонентам функции ошибки
    $\mathcal{L} \leftarrow \mathcal{L}_{\text{cls}} + \gamma(e) \cdot (\lambda_{\text{align}}\mathcal{L}_{\text{align}} + \lambda_{\text{gap}}\mathcal{L}_{\text{gap}})$
    Шаг оптимизатора (обновление параметров ϕ, w)
  end for
  Оценить потери на $\mathcal{D}_s^{\text{val}}$ и $\mathcal{D}_t^{\text{val}}$; обновить критерий ранней остановки
  if критерий ранней остановки сработал **then**
    break
  end if
end for
Постобработка: Пересчитать статистики пакетной нормализации на $\mathcal{D}_t^{\text{fit}}$ (AdaBN)
Калибровать порог τ на $\mathcal{D}_t^{\text{cal}}$ под заданный уровень FPR (раздел 2.5)
return Адаптированная модель f_θ и порог τ

Калибровка порога классификации. В условиях отсутствия размеченных данных с дефектом в целевом домене FPR является единственной метрикой, которую можно достоверно контролировать. Мы не можем настраивать порог по F_1-мере или полноте (Recall), так как для их вычисления требуются метки дефектов. Поэтому порог классификации τ устанавливается на отдельном подмножестве данных штатной работы из целевого домена $\mathcal{D}_t^{\text{cal}}$ с целью гарантировать, что доля объектов штатной работы, ошибочно классифицированных как с дефектом, не превысит заданный предел $\text{FPR}_{\text{target}}$.

В данной работе мы полагаем $\text{FPR}_{\text{target}} = 0.05$. Этот выбор обусловлен не только классическим уровнем значимости $\alpha = 0.05$ в статистической проверке гипотез, но и строгими требованиями промышленной эксплуатации систем прогнозного обслуживания. Данный порог широко используется в современных исследованиях как оптимальный компромисс между надежностью и экономикой обслуживания. Например, в работе [40] при диагностике зарождающейся кавитации насосов целевой уровень ложных срабатываний также жестко фиксировался на уровне 5% для обеспечения непрерывности работы оборудования. Аналогично, в недавнем исследовании [41] по предиктивному обслуживанию интеллектуальных энергосетей ограничение FPR на уровне 5% было выбрано как ключевой критерий для минимизации затрат на ложные срабатывания при сохранении высокой полноты обнаружения дефектов.

Отметим, что предложенная схема калибровки не привязана к конкретному числовому значению: параметр $\mathrm{FPR}_{\mathrm{target}}$ может быть изменен в зависимости от приоритетов конкретной задачи — например, ужесточен до 0.01 в системах, где критична стоимость ложных срабатываний, или, напротив, смягчен до 0.10–0.15 в сценариях, где пропуск развивающегося дефекта существенно опаснее внеплановой диагностики.

Порог выбирается как квантиль уровня $1 - \mathrm{FPR}_{\mathrm{target}}$ распределения предсказанных вероятностей ансамбля:

$$\tau = \mathrm{Quantile}_{1-\mathrm{FPR}_{\mathrm{target}}}\big(\{\hat{p}(x) : x \in \mathcal{D}_t^{\mathrm{cal}}\}\big). \tag{13}$$

Объект x признается дефектным, если $\hat{p}(x) \geqslant \tau$.

Фильтрация нестабильных моделей. При обучении ансамбля из K моделей с разными случайными инициализациями часть запусков может сходиться к неудовлетворительным решениям: некоторые модели коллапсируют (выдают почти константные предсказания), другие расходятся с консенсусом большинства. Поскольку целевые данные с дефектом недоступны, мы не можем проверить модели на тесте — вместо этого мы оцениваем их согласованность на валидационном наборе данных штатной работы $\mathcal{D}_t^{\mathrm{val}}$.

Для каждой модели k вычисляются две характеристики ее согласованности с медианным предсказанием ансамбля:

$$\mathrm{dev}_k = \mathbb{E}_{x \sim \mathcal{D}_t^{\mathrm{val}}} \big|\hat{p}_k(x) - \mathrm{median}_j\ \hat{p}_j(x)\big|, \tag{14}$$

$$\mathrm{corr}_k = \mathrm{Pearson}\big(\hat{p}_k(\mathcal{D}_t^{\mathrm{val}}),\ \mathrm{median}_j\ \hat{p}_j(\mathcal{D}_t^{\mathrm{val}})\big), \tag{15}$$

где $\hat{p}_k(x) = \sigma(\ell_k(x))$ — вероятность дефекта, предсказанная моделью k. Величина dev_k измеряет среднее абсолютное отклонение предсказаний модели k от медианы ансамбля, а corr_k — коэффициент корреляции Пирсона между вектором предсказаний модели k и медианным вектором ансамбля на наборе $\mathcal{D}_t^{\mathrm{val}}$.

Модель исключается из ансамбля, если выполняется хотя бы одно из двух условий:

$$z\text{-score}(\mathrm{dev}_k) > 2 \quad \text{или} \quad \mathrm{corr}_k \leqslant 0, \tag{16}$$

где z-score — стандартизированная оценка, которая вычисляется относительно распределения $\mathrm{dev}_1, \ldots, \mathrm{dev}_K$ по всему ансамблю.

Первое условие (z-score отклонения) отфильтровывает модели с аномально высокой дисперсией предсказаний на данных штатной работы из целевого домена. Такие модели демонстрируют хаотичное, нестабильное поведение: их выходы значительно отклоняются от стабильного консенсуса ансамбля, что указывает на нарушение сходимости процесса адаптации и искажение латентного пространства.

Второе условие (отрицательная корреляция Пирсона) выявляет модели с инвертированными предсказаниями: они выдают высокие вероятности дефекта для объектов, которые большинство ансамбля считает нормальными. Такое поведение возникает при неудачной сходимости адаптации, когда целевые нормальные данные смещаются в область латентного пространства, соответствующую дефектам источника.

Описанные критерии по построению чувствительны только к поведению, отклоняющемуся от консенсуса большинства членов ансамбля, и поэтому надежно выявляют лишь коллапс, обусловленный случайностью инициализации отдельных запусков. Если деградация воспроизводится синхронно у большинства моделей — например, из-за структурных

свойств целевого домена — медиана смещается вместе с отказавшими моделями, и критерии теряют чувствительность. В таких условиях, при полном отсутствии данных цели с дефектом, задача выбора между корректной и вырожденной моделью принципиально не разрешима на основе доступных данных.

В заключение отметим, что использование ансамблей независимо обученных нейронных сетей для оценки неопределенности предсказаний является хорошо изученным подходом [42, 43]. Однако классические методы опираются на доступность разметки хотя бы на этапе валидации. В рассматриваемой постановке UDA с полным отсутствием целевых примеров с дефектом прямая оценка уверенности невозможна. Мы предлагаем использовать разногласие моделей не для квантификации неопределенности отдельного объекта, а как критерий отбраковки нестабильных моделей на этапе формирования ансамбля. Насколько нам известно, предложенная комбинация — агрегация в пространстве логитов, фильтрация по корреляции с медианой для выявления «инвертированных» моделей и применение данной схемы к задаче диагностики с нулевым числом примеров целевого класса — ранее в литературе не описывалась и составляет практический вклад данной работы.

3. Экспериментальные результаты

3.1. Описание датасета

Используется открытый набор данных [44], содержащий записи токов статора синхронных электродвигателей с постоянными магнитами двух номинальных мощностей: 3 кВт и 1.5 кВт. Электродвигатели работали при скорости вращения 3000 об/мин под нагрузкой 15% от номинального момента.

Для каждого двигателя записаны сигналы в восьми состояниях: одном штатном и семи степенях тяжести межвиткового замыкания (МВЗ) в обмотке статора. Каждая запись представляет собой четырехканальный сигнал длиной 480 000 отсчетов при частоте дискретизации 4000 Гц: три фазных тока статора и канал вибрации. В данной работе используются только три канала фазных токов.

Сигналы сегментируются на окна длиной 400 отсчетов с шагом 200 отсчетов, что дает 2399 окон на каждое состояние. Для устранения дисбаланса классов из каждого состояния соответсвующего дефекту используется лишь часть окон, что дает суммарно 2400 окон с дефектом, сопоставимых по объему с 2399 штатными.

3.2. Предобработка и извлечение признаков

Для обеспечения инвариантности к нагрузке вместо исходных фазных токов используются их попарные разности:

$$s_1 = I_a - I_b, \quad s_2 = I_b - I_c, \quad s_3 = I_c - I_a, \tag{17}$$

где I_a, I_b, I_c — токи фаз статора. Каждый разностный сигнал нормируется независимо в пределах окна.

Для каждого из трех разностных сигналов извлекается 156 признаков во временной, статистической и частотной областях с использованием библиотеки TSFEL [45]. Применяется многоступенчатая процедура отбора: оценка значимости с помощью случайного леса из 200 деревьев, голосование по трем перестановкам разностных сигналов для устранения зависимости от порядка фаз, и исключение частотно-зависимых спектральных признаков.

В результате для каждого разностного сигнала отбирается 26 информативных признаков (табл. 1).

Таблица 1. Группы признаков, используемых для диагностики

Группа признаков	Количество
Статистические признаки временной области (эксцесс, асимметрия, медиана, максимум и др.)	8
Кепстральные коэффициенты линейного предсказания (LPCC)	6
Мел-кепстральные коэффициенты (MFCC)	4
Спектральные характеристики (центроид, дисперсия, roll-off и др.)	8
Итого	**26**

Финальная нормализация выполняется методом z-нормировки, где статистики вычисляются по обучающей выборке исходного домена.

Поскольку процедура извлечения признаков применяется независимо к каждому из трех разностных сигналов, итоговый вектор признаков для каждого окна содержит $26 \times 3 = 78$ компонент.

3.3. Разбиение данных

Рассматривается задача адаптации при переносе с двигателя 3 кВт на двигатель 1.5 кВт. Разбиение данных представлено в табл. 2.

Таблица 2. Разбиение данных по доменам

Домен	Набор	Штатная работа	С дефектом	Всего
Исходный (3 кВт)	Обучение	1680	1680	3360
	Валидация	–	–	480
	Тест (source)	–	–	959
Целевой (1.5 кВт)	Адаптация ($\mathcal{D}_t^{\text{fit}}$)	1500	–	1500
	Валидация ($\mathcal{D}_t^{\text{val}}$)	200	–	200
	Калибровка ($\mathcal{D}_t^{\text{cal}}$)	200	–	200
	Тест (target)	499	499	998

Разбиение выполняется с учетом временного порядка записи, что исключает утечку информации между соседними окнами. Окна целевого домена с дефектом используются исключительно для финальной оценки качества и не участвуют в адаптации. Доменный сдвиг обусловлен различием номинальных мощностей, что приводит к изменению амплитудных и частотных характеристик сигналов.

Отметим различие в использовании тестовых наборов. Тестовый набор исходного домена используется для верификации качества базовой модели до начала адаптации. Что касается целевого домена, то его данные с дефектом используются нами исключительно для финальной оценки качества в разделе 3. Ни на одном этапе предобучения, адаптации, ранней остановки или калибровки порога алгоритм не имеет доступа к разметке целевых дефектов, что полностью соответствует постановке нашей задачи.

3.4. Результаты

Все эксперименты проводились на 30 независимых запусках с различными случайными инициализациями весов. Для оценки качества диагностики использовались метрики Precision, Recall, F_1, FPR и ROC-AUC. Положительному классу соответствует дефектное состояние ($y = 1$). Обозначим через TP и TN числа правильно распознанных примеров дефектной и штатной работы, через FP и FN числа ложных срабатываний и пропущенных дефектов соответственно. Тогда точность, полнота и F_1-мера и частота ложных срабатываний определяются как:

$$\text{Precision} = \frac{TP}{TP + FP}, \qquad \text{Recall} = \frac{TP}{TP + FN}, \tag{18}$$

$$\text{F}_1 = \frac{2TP}{2TP + FP + FN}, \qquad \text{FPR} = \frac{FP}{FP + TN}. \tag{19}$$

Площадь под ROC-кривой, полученной при изменении порога классификации, вычисляется как

$$\text{ROC-AUC} = \int_0^1 \text{Recall}(FPR)\, dFPR. \tag{20}$$

Все модели и алгоритмы адаптации реализованы на фреймворке **PyTorch**. Основные гиперпараметры метода приведены в табл. 3. Все гиперпараметры фиксировались на основе предварительных экспериментов на валидационной выборке исходного домена и не подбирались на целевых данных, что исключает утечку информации. Порог классификации τ калибровался строго по неразмеченным данным штатной работы из целевого домена для обеспечения заданного уровня частоты ложных срабатываний $\text{FPR}_{\text{target}} = 0.05$.

Таблица 3. Основные гиперпараметры метода FD-GAP

Параметр	**Значение**
Размер латентного пространства	16
Оптимизатор / Learning rate	Adam / 1×10^{-4}
Весовое затухание	1×10^{-4}
Размер мини-батча (классификация / выравнивание)	256 / 512
Макс. число эпох адаптации / Patience	200 / 30
λ_{align} (вес Фреше-выравнивания)	1.0
λ_{gap} (вес функции зазора)	1.0
k_{gap} (коэффициент запаса радиуса)	0.5
Число моделей в ансамбле (K)	30
Целевой уровень FPR для калибровки порога	0.05

В табл. 4 представлены усредненные метрики качества для отдельных моделей. Анализ показывает, что конфигурации, включающие компонент выравнивания распределений (`frechet_only`, `frechet_gap`), обеспечивают статистически значимый прирост полноты обнаружения дефектов (Recall ≈ 0.44) по сравнению с базовой моделью (`baseline`), сохраняя при этом FPR на заданном уровне ≈ 0.05. При этом различия между конфигурациями находятся в пределах статистической погрешности.

Таблица 4. Результаты адаптации отдельных моделей (среднее ± стандартное отклонение по 30 запускам)

Конфигурация	F_1	**Recall**	**Precision**	**FPR**	**ROC-AUC**
`baseline` (без AdaBN)	0.277 ±0.128	0.177 ±0.091	0.702 ±0.187	0.056 ±0.023	0.572 ±0.046
`baseline` (с AdaBN)	0.318 ±0.125	0.207 ±0.090	0.739 ±0.176	0.056 ±0.022	0.596 ±0.046
`gap_only` (с AdaBN)	0.287 ±0.122	0.183 ±0.086	0.709 ±0.192	0.057 ±0.024	0.582 ±0.050
frechet_only (с AdaBN)	**0.589 ±0.125**	**0.448 ±0.107**	**0.873 ±0.167**	**0.049 ±0.021**	**0.773 ±0.046**
frechet_gap (с AdaBN)	**0.582 ±0.124**	**0.440 ±0.104**	**0.874 ±0.168**	**0.047 ±0.019**	**0.778 ±0.045**

Эффективность ансамблирования и фильтрации. Для оценки вклада ансамблирования и Фильтра согласованности (раздел 2.5) рассматриваются два варианта агрегации предсказаний для каждой конфигурации:

- **raw** — базовый ансамбль: медианное агрегирование логитов всех $K = 30$ обученных моделей без исключений;
- **filtered** — ансамбль с фильтрацией: агрегирование после применения Фильтра согласованности, который исключает из голосования модели, демонстрирующие аномальное отклонение от консенсуса (z-score > 2) или инверсию рангов предсказаний (corr $\leqslant 0$) на валидационной выборке целевого домена $\mathcal{D}_t^{\text{val}}$.

В обоих вариантах порог классификации калибруется заново на вероятностях ансамбля (формула (13)), что обеспечивает корректное сравнение с заданным уровнем $\text{FPR}_{\text{target}} = 0.05$.

Таблица 5. Результаты ансамблирования

Метод	**Вариант**	**Моделей**	F_1	**Recall**	**Precision**	**FPR**	**ROC-AUC**
`baseline`	raw	30	0.464	0.334	0.761	0.106	0.593
	filtered	29	0.473	0.347	0.745	0.120	0.590
`frechet_only`	raw	30	0.679	0.545	0.904	0.059	0.821
	filtered	**27**	**0.679**	**0.545**	**0.905**	**0.058**	**0.821**
`frechet_gap`	raw	30	0.672	0.535	0.906	0.057	0.829
	filtered	**27**	**0.668**	**0.531**	**0.904**	**0.058**	**0.829**

Анализ результатов ансамблирования показывает, что все конфигурации с компонентом выравнивания распределений статистически значимо превосходят базовый ансамбль `baseline` по ключевым метрикам качества (F_1 и ROC-AUC). При этом различия между методами выравнивания находятся в пределах статистической погрешности: конфигурации `frechet_only` и `frechet_gap` дают близкие значения F_1.

Применение фильтра согласованности в данном эксперименте оказывает умеренное влияние на итоговые метрики: из ансамбля исключается в среднем 2–3 модели (от 27 до 29 из 30), что приводит к незначительной стабилизации результатов без существенного изменения их среднего уровня. Это указывает на то, что при умеренном доменном сдвиге (перенос 3 кВт $\rightarrow$ 1.5 кВт) базовый механизм выравнивания уже обеспечивает достаточную устойчивость латентных представлений, и дополнительная фильтрация не дает существенного выигрыша.

Следует отметить, что конфигурация `frechet_gap` не является безусловным лидером по F_1, хотя и демонстрирует наилучшую ROC-AUC. При умеренном доменном сдвиге компонент выравнивания сам по себе уже достаточен для совмещения распределений штатных режимов, и функция зазора дает нейтральный или слабо неоднозначный вклад. Это согласуется с результатами на уровне отдельных моделей (табл. 4), где различия между `frechet_only` и `frechet_gap` также статистически незначимы.

В заключение отметим, что в нашей работе [46] на том же наборе данных был применен метод адаптации на основе Margin Disparity Discrepancy (MDD), достигающий $F_1 = 0.767$ в постановке с 0% размеченных данных цели. Сопоставление этого результата с $F_1 = 0.668$ для `frechet_gap` требует учета различий в доступности данных. В работе [46] при 0% размеченных данных метод MDD использует полный неразмеченный набор целевого двигателя, физически содержащий примеры с дефектом, для выравнивания условных распределений. В нашей постановке данные с дефектом целевого домена полностью отсутствуют. Более того, анализ матрицы ошибок из [46] позволяет рассчитать частоту ложных срабатываний (FPR) для их модели: из 1596 фактически состояний штатной работы модель ошибочно классифицирует как с дефектом 897, что соответствует FPR ≈ 0.56. Метод `frechet_gap`, напротив, калибрует порог классификации под заданное ограничение $\mathrm{FPR}_{\mathrm{target}} = 0.05$, что обеспечивает контролируемый уровень ложных срабатываний, приемлемый для промышленной эксплуатации, ценой некоторого снижения метрики F_1.

Метод MDD также подвержен нестабильности обучения: авторы [46] демонстрируют немонотонную зависимость качества от количества размеченных данных и отмечают тенденцию модели к классификации всех состояний штатной работы как с дефектом при недостаточном объеме размеченной выборки. Для компенсации этой нестабильности в работе [46] используется эмпирический подбор оптимальной доли размеченных данных (12.5%) и валидация на размеченном тестовом наборе целевого домена. В нашей постановке, где целевые данные с дефектом полностью отсутствуют, такая валидация невозможна. Поэтому предложенный механизм ансамблирования с фильтрацией является не просто способом повышения метрик, а единственным доступным инструментом для выявления и исключения нестабильных моделей.

Исходный код программной реализации предложенного метода и подготовленные выборки, использованные в вычислительных экспериментах, предоставляются авторами по запросу. Исходный набор экспериментальных данных доступен в открытом доступе [44].

4. Обсуждение

В условиях полного отсутствия целевых меток с дефектом качество отдельной адаптированной модели сильно зависит от случайной инициализации, поэтому ансамблирование с фильтрацией является необходимым механизмом для выявления и исключения нестабильных моделей. Этот подход принципиально отличает нашу работу от существующих методов,

таких как Margin Disparity Discrepancy (MDD) [46]. Хотя MDD достигает $F_1 = 0.767$ на том же наборе данных синхронных двигателей, он использует полный неразмеченный набор целевых данных (физически содержащий дефекты) и склонен к массовым ложным срабатываниям (расчет по матрице ошибок из [46] дает FPR ≈ 0.56). Предложенный в данной статье метод FD-GAP решает более строгую задачу, полностью исключая целевые дефекты, и калибрует порог под жесткое ограничение $\mathrm{FPR}_{\mathrm{target}} = 0.05$.

Ключевые ограничения метода носят концептуальный и практический характер. Во-первых, функция зазора неявно предполагает схожесть физических механизмов дефектов в источнике и цели — если природа дефекта в целевом домене качественно иная, перенос границы может оказаться некорректным. Во-вторых, обучение ансамбля из $K = 30$ моделей увеличивает вычислительные затраты, хотя эта процедура может быть распараллелена. В-третьих, метод требует наличия хотя бы неразмеченных данных штатной работы целевого домена, что в большинстве практических сценариев прогнозного обслуживания выполняется. Наконец, фильтр отсеивает только модели, отклоняющиеся от консенсуса меньшинства: если деградация носит систематический характер и затрагивает большинство запусков, задача отличить корректный консенсус от коллективно смещенного принципиально неразрешима без минимальной целевой разметки.

Заключение

В работе предложен метод FD-GAP для адаптации диагностических моделей в условиях полного отсутствия размеченных данных с дефектом целевого домена — постановка, отражающая реальный сценарий развертывания системы на новом оборудовании на этапе ввода в эксплуатацию. Метод сочетает три ключевых компонента. Во-первых, распределения штатных режимов источника и цели совмещаются посредством расстояния Фреше в латентном пространстве. Во-вторых, граница класса соответствующего дефекту переносится через функцию зазора, отталкивающей примеры с дефектом источника от кластера штатных данных. В-третьих, применяется ансамблирование K независимо обученных моделей с фильтрацией, компенсирующее невозможность прямой валидации отдельной модели по целевому дефектному классу.

Эксперименты в рамках задачи диагностики синхронных электродвигателей подтвердили эффективность предложенного подхода. В задаче переноса модели диагностики межвитковых замыканий синхронных двигателей с 3 кВт на 1.5 кВт ансамбль `frechet_gap` достигает $F_1 = 0.668$ при контролируемом уровне ложных срабатываний FPR ≈ 0.05 и ROC-AUC $= 0.829$, что превосходит базовый ансамбль на 0.204 по F_1 и на 0.236 по ROC-AUC.

Естественным направлением развития метода является переход к многоклассовой постановке, в которой целевой домен содержит несколько типов дефектов. В текущей версии выравнивание условных распределений $P(z \mid y)$ выполняется только для класса штатной работы ($y = 0$), тогда как для класса соответствующего ($y = 1$) перенос границы опирается исключительно на геометрическое предположение зазора. Если в процессе эксплуатации становится доступным даже небольшое число размеченных образцов конкретного типа дефекта в целевом домене, это предположение можно заменить прямым выравниванием: аналогично тому, как совмещаются кластеры штатного режима, становится возможным совместно адаптировать и распределения $P(z \mid y = k)$ для известного класса дефекта k. Это позволит не только повысить точность распознавания данного типа неисправности, но

и эмпирически оценить, насколько физика соответствующего дефекта согласована между источником и целью — то есть превратить одно из ключевых допущений метода в проверяемую величину. Дальнейшие исследования также будут направлены на проверку переносимости подхода на другие типы оборудования и виды сигналов.

Исследование выполнено за счет гранта Российского научного фонда № 25-29-01598, https://rscf.ru/project/25-29-01598/.

Литература

1. Li W., Huang R., Li J., Liao Y., *et al.* A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: Theories, applications and challenges // Mechanical Systems and Signal Processing. 2022. Vol. 167. DOI: 10.1016/j.ymssp.2021.108487.
2. Sullivan G., Pugh R., Melendez A.P., Hunt W.D. Operations & Maintenance Best Practices - A Guide to Achieving Operational Efficiency (Release 3): tech. rep. / Pacific Northwest National Laboratory (PNNL), Richland, WA (United States). Richland, WA, United States, Aug. 2010. DOI: 10.2172/1034595.
3. Khan S., Yairi T. A review on the application of deep learning in system health management // Mechanical Systems and Signal Processing. 2018. Vol. 107. P. 241–265. DOI: 10.1016/j.ymssp.2017.11.024.
4. Li C., Zhang S., Qin Y., Estupinan E. A systematic review of deep transfer learning for machinery fault diagnosis // Neurocomputing. 2020. Vol. 407. P. 121–135. DOI: 10.1016/j.neucom.2020.04.045.
5. Fink O., Wang Q., Svensén M., Dersin P., *et al.* Potential, challenges and future directions for deep learning in prognostics and health management applications // Engineering Applications of Artificial Intelligence. 2020. Vol. 92. DOI: 10.1016/j.engappai.2020.103678.
6. Gao T., Yang J., Tang Q. A multi-source domain information fusion network for rotating machinery fault diagnosis under variable operating conditions // Information Fusion. 2024. Vol. 106. DOI: 10.1016/j.inffus.2024.102278.
7. Jia N., Huang W., Ding C., Wang J., *et al.* Physics-informed unsupervised domain adaptation framework for cross-machine bearing fault diagnosis // Advanced Engineering Informatics. 2024. Vol. 62. DOI: 10.1016/j.aei.2024.102774.
8. Pan X., Chen H., Wang W., Su X. Adversarial domain adaptation based on contrastive learning for bearings fault diagnosis // Simulation Modelling Practice and Theory. 2025. Vol. 139. DOI: 10.1016/j.simpat.2024.103058.
9. Pang B., Liu Q., Sun Z., Xu Z., *et al.* Time-frequency supervised contrastive learning via pseudo-labeling: An unsupervised domain adaptation network for rolling bearing fault diagnosis under time-varying speeds // Advanced Engineering Informatics. 2024. Vol. 59. DOI: 10.1016/j.aei.2023.102304.
10. Lin F., Guo C., Zhao Z., Zhang X., *et al.* A Universal Domain Adaptation Method With Cluster Matching for Machinery Fault Diagnosis // IEEE Internet of Things Journal. 2025. Vol. 12, no. 6. P. 7487–7502. DOI: 10.1109/JIOT.2024.3496928.

11. Liu Y., Deng A., Chen G., Shi Y., *et al.* Universal domain adaptation in rotating machinery fault diagnosis: A self-supervised orthogonal clustering approach // Reliability Engineering & System Safety. 2025. Vol. 257. DOI: 10.1016/j.ress.2025.110828.

12. Tian J., Han D., Karimi H.R., Zhang Y., *et al.* A universal multi-source domain adaptation method with unsupervised clustering for mechanical fault diagnosis under incomplete data // Neural Networks. 2024. Vol. 173. DOI: 10.1016/j.neunet.2024.106167.

13. Jiao J., Li H., Zhang T., Lin J. Source-Free Adaptation Diagnosis for Rotating Machinery // IEEE Transactions on Industrial Informatics. 2023. Vol. 19, no. 9. P. 9586–9595. DOI: 10.1109/TII.2022.3231414.

14. Li Q., Tang B., Deng L., Yang Q., *et al.* Adaptive centroid prototype-based domain adaptation for fault diagnosis of rotating machinery without source data // Reliability Engineering & System Safety. 2024. Vol. 251. DOI: 10.1016/j.ress.2024.110393.

15. Liu Z., Zhao K., Shao H., Chen J., *et al.* Reinforced fuzzy domain adaptation: Revolutionizing data-unaccessible rotating machinery fault diagnosis across multiple domains // Expert Systems with Applications. 2024. Vol. 252. DOI: 10.1016/j.eswa.2024.124094.

16. Yu A., Cai B., Wu Q., Martínez García M., *et al.* Source-free domain adaptation method for fault diagnosis of rotation machinery under partial information // Reliability Engineering & System Safety. 2024. Vol. 248. DOI: 10.1016/j.ress.2024.110181.

17. Lin Y., Wang Y., Zhang M., Zhao M. A robust source-free unsupervised domain adaptation method based on uncertainty measure and adaptive calibration for rotating machinery fault diagnosis // Reliability Engineering & System Safety. 2025. Vol. 253. DOI: 10.1016/j.ress.2024.110516.

18. Jeong H., Kim S., Seo D., Kwon J. Source-Free Domain Adaptation Framework for Rotary Machine Fault Diagnosis // Sensors. 2025. Vol. 25, no. 14. DOI: 10.3390/s25144383.

19. Zhang Y., Ren Z., Feng K., Yu K., *et al.* Universal source-free domain adaptation method for cross-domain fault diagnosis of machines // Mechanical Systems and Signal Processing. 2023. Vol. 191. DOI: 10.1016/j.ymssp.2023.110159.

20. Chen L., Li Q., Shen C., Zhu J., *et al.* Adversarial Domain-Invariant Generalization: A Generic Domain-Regressive Framework for Bearing Fault Diagnosis Under Unseen Conditions // IEEE Transactions on Industrial Informatics. 2022. Vol. 18, no. 3. P. 1790–1800. DOI: 10.1109/TII.2021.3078712.

21. Han T., Li Y.-F., Qian M. A Hybrid Generalization Network for Intelligent Fault Diagnosis of Rotating Machinery Under Unseen Working Conditions // IEEE Transactions on Instrumentation and Measurement. 2021. Vol. 70. P. 1–11. DOI: 10.1109/TIM.2021.3088489.

22. Li X., Zhang W., Ma H., Luo Z., *et al.* Domain generalization in rotating machinery fault diagnostics using deep neural networks // Neurocomputing. 2020. Vol. 403. P. 409–420. DOI: 10.1016/j.neucom.2020.05.014.

23. Fan Z., Xu Q., Jiang C., Ding S.X. Deep Mixed Domain Generalization Network for Intelligent Fault Diagnosis Under Unseen Conditions // IEEE Transactions on Industrial Electronics. 2024. Vol. 71, no. 1. P. 965–974. DOI: 10.1109/TIE.2023.3243293.

24. Shi Y., Deng A., Deng M., Xu M., *et al.* Domain augmentation generalization network for real-time fault diagnosis under unseen working conditions // Reliability Engineering & System Safety. 2023. Vol. 235. DOI: 10.1016/j.ress.2023.109188.

25. Gao T., Yang J., Wang W., Fan X. A domain feature decoupling network for rotating machinery fault diagnosis under unseen operating conditions // Reliability Engineering & System Safety. 2024. Vol. 252. DOI: 10.1016/j.ress.2024.110449.

26. Wang R., Huang W., Lu Y., Zhang X., *et al.* A novel domain generalization network with multidomain specific auxiliary classifiers for machinery fault diagnosis under unseen working conditions // Reliability Engineering & System Safety. 2023. Vol. 238. DOI: 10.1016/j.ress.2023.109463.

27. Zhao C., Shen W. A domain generalization network combing invariance and specificity towards real-time intelligent fault diagnosis // Mechanical Systems and Signal Processing. 2022. Vol. 173. DOI: 10.1016/j.ymssp.2022.108990.

28. Qian Q., Zhou J., Qin Y. Relationship Transfer Domain Generalization Network for Rotating Machinery Fault Diagnosis Under Different Working Conditions // IEEE Transactions on Industrial Informatics. 2023. Vol. 19, no. 9. P. 9898–9908. DOI: 10.1109/TII.2022.3232842.

29. Liu X., Sun W., Li H., Li Q., *et al.* Unknown working condition fault diagnosis of rotate machine without training sample based on local fault semantic attribute // Advanced Engineering Informatics. 2024. Vol. 61. DOI: 10.1016/j.aei.2024.102515.

30. Liu C., Gryllias K. Simulation-Driven Domain Adaptation for Rolling Element Bearing Fault Diagnosis // IEEE Transactions on Industrial Informatics. 2022. Vol. 18, no. 9. P. 5760–5770. DOI: 10.1109/TII.2021.3103412.

31. Lou Y., Kumar A., Xiang J. Machinery Fault Diagnosis Based on Domain Adaptation to Bridge the Gap Between Simulation and Measured Signals // IEEE Transactions on Instrumentation and Measurement. 2022. Vol. 71. P. 1–9. DOI: 10.1109/TIM.2022.3180416.

32. Xiao Y., Shao H., Yan S., Wang J., *et al.* Domain generalization for rotating machinery fault diagnosis: A survey // Advanced Engineering Informatics. 2025. Vol. 64. DOI: 10.1016/j.aei.2024.103063.

33. Zhang Y., Ji J., Ren Z., Ni Q., *et al.* Digital twin-driven partial domain adaptation network for intelligent fault diagnosis of rolling bearing // Reliability Engineering & System Safety. 2023. Vol. 234. DOI: 10.1016/j.ress.2023.109186.

34. Azari M.S., Santini S., Edrisi F., Flammini F. Self-adaptive fault diagnosis for unseen working conditions based on digital twins and domain generalization // Reliability Engineering & System Safety. 2025. Vol. 254. DOI: 10.1016/j.ress.2024.110560.

35. Li Y., Wang N., Shi J., Hou X., *et al.* Adaptive Batch Normalization for practical domain adaptation // Pattern Recognition. 2018. Vol. 80. P. 109–117. DOI: 10.1016/j.patcog.2018.03.005.

36. Dowson D.C., Landau B.V. The Fréchet distance between multivariate normal distributions // Journal of Multivariate Analysis. 1982. Vol. 12, no. 3. P. 450–455. DOI: 10.1016/0047-259X(82)90077-X.

37. Ledoit O., Wolf M. A well-conditioned estimator for large-dimensional covariance matrices // Journal of Multivariate Analysis. 2004. Vol. 88, no. 2. P. 365–411. DOI: 10.1016/S0047-259X(03)00096-4.

38. Higham N.J. Functions of Matrices: Theory and Computation. Philadelphia: Society for Industrial, Applied Mathematics, 2008. Other Titles in Applied Mathematics. DOI: 10.1137/1.9780898717778.

39. Ionescu C., Vantzos O., Sminchisescu C. Matrix Backpropagation for Deep Networks with Structured Layers // 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 13–16 December 2015. IEEE, 2015. P. 2965–2973. DOI: 10.1109/ICCV.2015.339.

40. Chu N., Wang L., Yu L., He C., *et al.* An Adaptive Autogram Approach Based on a CFAR Detector for Incipient Cavitation Detection // Sensors. 2020. Vol. 20, no. 8. DOI: 10.3390/s20082303.

41. Cheng J., Yang D., Wang S. Unsupervised Learning Applications for Predictive Maintenance in Consumer Electronics and Intelligent Energy Networks // IEEE Transactions on Consumer Electronics. 2025. Vol. 71, no. 3. P. 8347–8354. DOI: 10.1109/TCE.2025.3573649.

42. Lakshminarayanan B., Pritzel A., Blundell C. Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles // Advances in Neural Information Processing Systems, Long Beach, California, USA, 4–9 December 2017. Vol. 30. 2017. P. 6402–6413. URL: https://papers.nips.cc/paper_files/paper/2017/hash/9ef2ed4b7fd2c810847ffa5fa85bce38-Abstract.html.

43. Tifrea A., Stavarache E., Yang F. Semi-supervised novelty detection using ensembles with regularized disagreement // Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence, Eindhoven, The Netherlands, 1–5 August 2022. Vol. 180 / ed. by J. Cussens, K. Zhang. PMLR, 2022. P. 1939–1948. Proceedings of Machine Learning Research. URL: https://proceedings.mlr.press/v180/tifrea22a.html.

44. Jung W., Kim S.-H., Yun S.-H., Bae J., *et al.* Vibration, acoustic, temperature, and motor current dataset of rotating machine under varying operating conditions for fault diagnosis // Data in Brief. 2023. Vol. 48. DOI: 10.1016/j.dib.2023.109049.

45. Barandas M., Folgado D., Fernandes L., Santos S., *et al.* TSFEL: Time Series Feature Extraction Library // SoftwareX. 2020. Vol. 11. DOI: 10.1016/j.softx.2020.100456.

46. Shestakov A.L., Galyshev D.V., Eremeeva V.A., Ibryaeva O.L. Domain-Adaptive and Multifactor Invariant Diagnosis of Inter-Turn Short Circuits in Permanent Magnet Synchronous Motors // 2026 28th International Conference on Digital Signal Processing and its Applications (DSPA), Moscow, Russian Federation, 25–27 March 2026. IEEE, 2026. P. 1–6. DOI: 10.1109/DSPA69176.2026.11476735.

Шестаков Александр Леонидович, докт. техн. наук, ведущий научный сотрудник управления научной и инновационной деятельности, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Россия)

Галышев Дмитрий Вячеславович, аспирант, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Россия)

Мартьянова Виктория Александровна, канд. техн. наук., инженер-исследователь управления научной и инновационной деятельности, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Россия)

Ибряева Ольга Леонидовна, докт. техн. наук, проф., старший научный сотрудник инновационного отдела, Южно-Уральский государственный университет (национальный исследовательский университет) (Челябинск, Россия)

DOI: 10.14529/cmse260306

A NOVEL METHOD FOR ADAPTING SYNCHRONOUS MOTOR DIAGNOSTIC MODELS IN THE ABSENCE OF TARGET-DOMAIN FAULT DATA

© 2026 A.L. Shestakov, D.V. Galyshev, V.A. Martyanova, O.L. Ibryaeva

South Ural State University

(pr. Lenina 76, Chelyabinsk, 454080 Russian Federation)

E-mail: a.l.shestakov@susu.ru, galyshevdv@susu.ru, eremeevava@susu.ru, ibriaevaol@susu.ru

Received: 30.09.2026

Among the approaches to fault diagnosis of rotating machinery under domain shift, the equipment commissioning scenario remains insufficiently studied, where only normal-state data can be collected and examples of target faults are completely unavailable. Conventional unsupervised domain adaptation methods are not directly applicable in this setting because target fault samples required for clustering are absent, while one-class anomaly detection methods ignore valuable knowledge about faults available in the source domain. Approaches based on digital twins or generative adversarial networks require computationally expensive physical modeling and may suffer from discrepancies between synthetic and real data. We propose a novel FD-GAP domain adaptation approach for the setting in which the target domain contains only normal operating data. Unlike previous approaches limited to distribution alignment, the proposed method transfers the geometric boundary of the fault class from the source domain to the target domain using the Fréchet distance and gap loss to push fault states away from the normal-state region. To compensate for stochastic model instability caused by different random initializations, a novel disagreement-based ensemble filtering mechanism is introduced. The mechanism operates exclusively on normal-state data from the target equipment and improves deployment reliability while maintaining a prescribed false-positive rate.

Keywords: FD-GAP, gap loss, domain adaptation, one-class learning, inter-turn short circuits, synchronous motors, stator current-based diagnostics, Fréchet distance, neural network ensemble, unstable model filtering.

FOR CITATION

Shestakov A.L., Galyshev D.V., Martyanova V.A., Ibryaeva O.L. A Novel Method for Adapting Synchronous Motor Diagnostic Models in the Absence of Target-domain Fault Data. Bulletin of the South Ural State University. Series: Computational Mathematics and Software Engineering. 2026. Vol. 15, no. 3. P. 117–142. (in Russian) DOI: 10.14529/cmse260306.

This paper is distributed under the terms of the Creative Commons Attribution-Non Commercial 4.0 License which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is properly cited.

References

1. Li W., Huang R., Li J., Liao Y., *et al.* A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: Theories, applications and challenges. Mechanical Systems and Signal Processing. 2022. Vol. 167. DOI: 10.1016/j.ymssp.2021.108487.
2. Sullivan G., Pugh R., Melendez A.P., Hunt W.D. Operations & Maintenance Best Practices - A Guide to Achieving Operational Efficiency (Release 3): tech. rep. / Pacific Northwest National Laboratory (PNNL), Richland, WA (United States). Richland, WA, United States, Aug. 2010. DOI: 10.2172/1034595.
3. Khan S., Yairi T. A review on the application of deep learning in system health management. Mechanical Systems and Signal Processing. 2018. Vol. 107. P. 241–265. DOI: 10.1016/j.ymssp.2017.11.024.
4. Li C., Zhang S., Qin Y., Estupinan E. A systematic review of deep transfer learning for machinery fault diagnosis. Neurocomputing. 2020. Vol. 407. P. 121–135. DOI: 10.1016/j.neucom.2020.04.045.
5. Fink O., Wang Q., Svensén M., Dersin P., *et al.* Potential, challenges and future directions for deep learning in prognostics and health management applications. Engineering Applications of Artificial Intelligence. 2020. Vol. 92. DOI: 10.1016/j.engappai.2020.103678.
6. Gao T., Yang J., Tang Q. A multi-source domain information fusion network for rotating machinery fault diagnosis under variable operating conditions. Information Fusion. 2024. Vol. 106. DOI: 10.1016/j.inffus.2024.102278.
7. Jia N., Huang W., Ding C., Wang J., *et al.* Physics-informed unsupervised domain adaptation framework for cross-machine bearing fault diagnosis. Advanced Engineering Informatics. 2024. Vol. 62. DOI: 10.1016/j.aei.2024.102774.
8. Pan X., Chen H., Wang W., Su X. Adversarial domain adaptation based on contrastive learning for bearings fault diagnosis. Simulation Modelling Practice and Theory. 2025. Vol. 139. DOI: 10.1016/j.simpat.2024.103058.
9. Pang B., Liu Q., Sun Z., Xu Z., *et al.* Time-frequency supervised contrastive learning via pseudo-labeling: An unsupervised domain adaptation network for rolling bearing fault diagnosis under time-varying speeds. Advanced Engineering Informatics. 2024. Vol. 59. DOI: 10.1016/j.aei.2023.102304.
10. Lin F., Guo C., Zhao Z., Zhang X., *et al.* A Universal Domain Adaptation Method With Cluster Matching for Machinery Fault Diagnosis. IEEE Internet of Things Journal. 2025. Vol. 12, no. 6. P. 7487–7502. DOI: 10.1109/JIOT.2024.3496928.
11. Liu Y., Deng A., Chen G., Shi Y., *et al.* Universal domain adaptation in rotating machinery fault diagnosis: A self-supervised orthogonal clustering approach. Reliability Engineering & System Safety. 2025. Vol. 257. DOI: 10.1016/j.ress.2025.110828.
12. Tian J., Han D., Karimi H.R., Zhang Y., *et al.* A universal multi-source domain adaptation method with unsupervised clustering for mechanical fault diagnosis under incomplete data. Neural Networks. 2024. Vol. 173. DOI: 10.1016/j.neunet.2024.106167.
13. Jiao J., Li H., Zhang T., Lin J. Source-Free Adaptation Diagnosis for Rotating Machinery. IEEE Transactions on Industrial Informatics. 2023. Vol. 19, no. 9. P. 9586–9595. DOI: 10.1109/TII.2022.3231414.

14. Li Q., Tang B., Deng L., Yang Q., *et al.* Adaptive centroid prototype-based domain adaptation for fault diagnosis of rotating machinery without source data. Reliability Engineering & System Safety. 2024. Vol. 251. DOI: 10.1016/j.ress.2024.110393.

15. Liu Z., Zhao K., Shao H., Chen J., *et al.* Reinforced fuzzy domain adaptation: Revolutionizing data-unaccessible rotating machinery fault diagnosis across multiple domains. Expert Systems with Applications. 2024. Vol. 252. DOI: 10.1016/j.eswa.2024.124094.

16. Yu A., Cai B., Wu Q., Martínez García M., *et al.* Source-free domain adaptation method for fault diagnosis of rotation machinery under partial information. Reliability Engineering & System Safety. 2024. Vol. 248. DOI: 10.1016/j.ress.2024.110181.

17. Lin Y., Wang Y., Zhang M., Zhao M. A robust source-free unsupervised domain adaptation method based on uncertainty measure and adaptive calibration for rotating machinery fault diagnosis. Reliability Engineering & System Safety. 2025. Vol. 253. DOI: 10.1016/j.ress.2024.110516.

18. Jeong H., Kim S., Seo D., Kwon J. Source-Free Domain Adaptation Framework for Rotary Machine Fault Diagnosis. Sensors. 2025. Vol. 25, no. 14. DOI: 10.3390/s25144383.

19. Zhang Y., Ren Z., Feng K., Yu K., *et al.* Universal source-free domain adaptation method for cross-domain fault diagnosis of machines. Mechanical Systems and Signal Processing. 2023. Vol. 191. DOI: 10.1016/j.ymssp.2023.110159.

20. Chen L., Li Q., Shen C., Zhu J., *et al.* Adversarial Domain-Invariant Generalization: A Generic Domain-Regressive Framework for Bearing Fault Diagnosis Under Unseen Conditions. IEEE Transactions on Industrial Informatics. 2022. Vol. 18, no. 3. P. 1790–1800. DOI: 10.1109/TII.2021.3078712.

21. Han T., Li Y.-F., Qian M. A Hybrid Generalization Network for Intelligent Fault Diagnosis of Rotating Machinery Under Unseen Working Conditions. IEEE Transactions on Instrumentation and Measurement. 2021. Vol. 70. P. 1–11. DOI: 10.1109/TIM.2021.3088489.

22. Li X., Zhang W., Ma H., Luo Z., *et al.* Domain generalization in rotating machinery fault diagnostics using deep neural networks. Neurocomputing. 2020. Vol. 403. P. 409–420. DOI: 10.1016/j.neucom.2020.05.014.

23. Fan Z., Xu Q., Jiang C., Ding S.X. Deep Mixed Domain Generalization Network for Intelligent Fault Diagnosis Under Unseen Conditions. IEEE Transactions on Industrial Electronics. 2024. Vol. 71, no. 1. P. 965–974. DOI: 10.1109/TIE.2023.3243293.

24. Shi Y., Deng A., Deng M., Xu M., *et al.* Domain augmentation generalization network for real-time fault diagnosis under unseen working conditions. Reliability Engineering & System Safety. 2023. Vol. 235. DOI: 10.1016/j.ress.2023.109188.

25. Gao T., Yang J., Wang W., Fan X. A domain feature decoupling network for rotating machinery fault diagnosis under unseen operating conditions. Reliability Engineering & System Safety. 2024. Vol. 252. DOI: 10.1016/j.ress.2024.110449.

26. Wang R., Huang W., Lu Y., Zhang X., *et al.* A novel domain generalization network with multidomain specific auxiliary classifiers for machinery fault diagnosis under unseen working conditions. Reliability Engineering & System Safety. 2023. Vol. 238. DOI: 10.1016/j.ress.2023.109463.

27. Zhao C., Shen W. A domain generalization network combing invariance and specificity towards real-time intelligent fault diagnosis. Mechanical Systems and Signal Processing. 2022. Vol. 173. DOI: 10.1016/j.ymssp.2022.108990.

28. Qian Q., Zhou J., Qin Y. Relationship Transfer Domain Generalization Network for Rotating Machinery Fault Diagnosis Under Different Working Conditions. IEEE Transactions on Industrial Informatics. 2023. Vol. 19, no. 9. P. 9898–9908. DOI: 10.1109/TII.2022.3232842.

29. Liu X., Sun W., Li H., Li Q., *et al.* Unknown working condition fault diagnosis of rotate machine without training sample based on local fault semantic attribute. Advanced Engineering Informatics. 2024. Vol. 61. DOI: 10.1016/j.aei.2024.102515.

30. Liu C., Gryllias K. Simulation-Driven Domain Adaptation for Rolling Element Bearing Fault Diagnosis. IEEE Transactions on Industrial Informatics. 2022. Vol. 18, no. 9. P. 5760–5770. DOI: 10.1109/TII.2021.3103412.

31. Lou Y., Kumar A., Xiang J. Machinery Fault Diagnosis Based on Domain Adaptation to Bridge the Gap Between Simulation and Measured Signals. IEEE Transactions on Instrumentation and Measurement. 2022. Vol. 71. P. 1–9. DOI: 10.1109/TIM.2022.3180416.

32. Xiao Y., Shao H., Yan S., Wang J., *et al.* Domain generalization for rotating machinery fault diagnosis: A survey. Advanced Engineering Informatics. 2025. Vol. 64. DOI: 10.1016/j.aei.2024.103063.

33. Zhang Y., Ji J., Ren Z., Ni Q., *et al.* Digital twin-driven partial domain adaptation network for intelligent fault diagnosis of rolling bearing. Reliability Engineering & System Safety. 2023. Vol. 234. DOI: 10.1016/j.ress.2023.109186.

34. Azari M.S., Santini S., Edrisi F., Flammini F. Self-adaptive fault diagnosis for unseen working conditions based on digital twins and domain generalization. Reliability Engineering & System Safety. 2025. Vol. 254. DOI: 10.1016/j.ress.2024.110560.

35. Li Y., Wang N., Shi J., Hou X., *et al.* Adaptive Batch Normalization for practical domain adaptation. Pattern Recognition. 2018. Vol. 80. P. 109–117. DOI: 10.1016/j.patcog.2018.03.005.

36. Dowson D.C., Landau B.V. The Fréchet distance between multivariate normal distributions. Journal of Multivariate Analysis. 1982. Vol. 12, no. 3. P. 450–455. DOI: 10.1016/0047-259X(82)90077-X.

37. Ledoit O., Wolf M. A well-conditioned estimator for large-dimensional covariance matrices. Journal of Multivariate Analysis. 2004. Vol. 88, no. 2. P. 365–411. DOI: 10.1016/S0047-259X(03)00096-4.

38. Higham N.J. Functions of Matrices: Theory and Computation. Philadelphia: Society for Industrial, Applied Mathematics, 2008. Other Titles in Applied Mathematics. DOI: 10.1137/1.9780898717778.

39. Ionescu C., Vantzos O., Sminchisescu C. Matrix Backpropagation for Deep Networks with Structured Layers. 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 13–16 December 2015. IEEE, 2015. P. 2965–2973. DOI: 10.1109/ICCV.2015.339.

40. Chu N., Wang L., Yu L., He C., *et al.* An Adaptive Autogram Approach Based on a CFAR Detector for Incipient Cavitation Detection. Sensors. 2020. Vol. 20, no. 8. DOI: 10.3390/s20082303.

41. Cheng J., Yang D., Wang S. Unsupervised Learning Applications for Predictive Maintenance in Consumer Electronics and Intelligent Energy Networks. IEEE Transactions on Consumer Electronics. 2025. Vol. 71, no. 3. P. 8347–8354. DOI: 10.1109/TCE.2025.3573649.

42. Lakshminarayanan B., Pritzel A., Blundell C. Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles. Advances in Neural Information Processing Systems, Long Beach, California, USA, 4–9 December 2017. Vol. 30. 2017. P. 6402–6413. URL: https://papers.nips.cc/paper_files/paper/2017/hash/9ef2ed4b7fd2c810847ffa5fa85bce38-Abstract.html.

43. Tifrea A., Stavarache E., Yang F. Semi-supervised novelty detection using ensembles with regularized disagreement. Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence, Eindhoven, The Netherlands, 1–5 August 2022. Vol. 180 / ed. by J. Cussens, K. Zhang. PMLR, 2022. P. 1939–1948. Proceedings of Machine Learning Research. URL: https://proceedings.mlr.press/v180/tifrea22a.html.

44. Jung W., Kim S.-H., Yun S.-H., Bae J., *et al.* Vibration, acoustic, temperature, and motor current dataset of rotating machine under varying operating conditions for fault diagnosis. Data in Brief. 2023. Vol. 48. DOI: 10.1016/j.dib.2023.109049.

45. Barandas M., Folgado D., Fernandes L., Santos S., *et al.* TSFEL: Time Series Feature Extraction Library. SoftwareX. 2020. Vol. 11. DOI: 10.1016/j.softx.2020.100456.

46. Shestakov A.L., Galyshev D.V., Eremeeva V.A., Ibryaeva O.L. Domain-Adaptive and Multifactor Invariant Diagnosis of Inter-Turn Short Circuits in Permanent Magnet Synchronous Motors. 2026 28th International Conference on Digital Signal Processing and its Applications (DSPA), Moscow, Russian Federation, 25–27 March 2026. IEEE, 2026. P. 1–6. DOI: 10.1109/DSPA69176.2026.11476735.