



ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР
**ИНФОРМАТИКА
И УПРАВЛЕНИЕ**
РОССИЙСКОЙ АКАДЕМИИ НАУК

Информированные нейронные сети: от детерминированных подходов к стохастическому моделированию сложных систем

Андрей Константинович Горшенин

*д.ф.-м.н., доцент
главный научный сотрудник, руководитель отдела*

4 марта 2026 г.

ОБОСНОВАННЫЕ НЕЙРОСЕТЕВЫЕ МОДЕЛИ



Актуальность

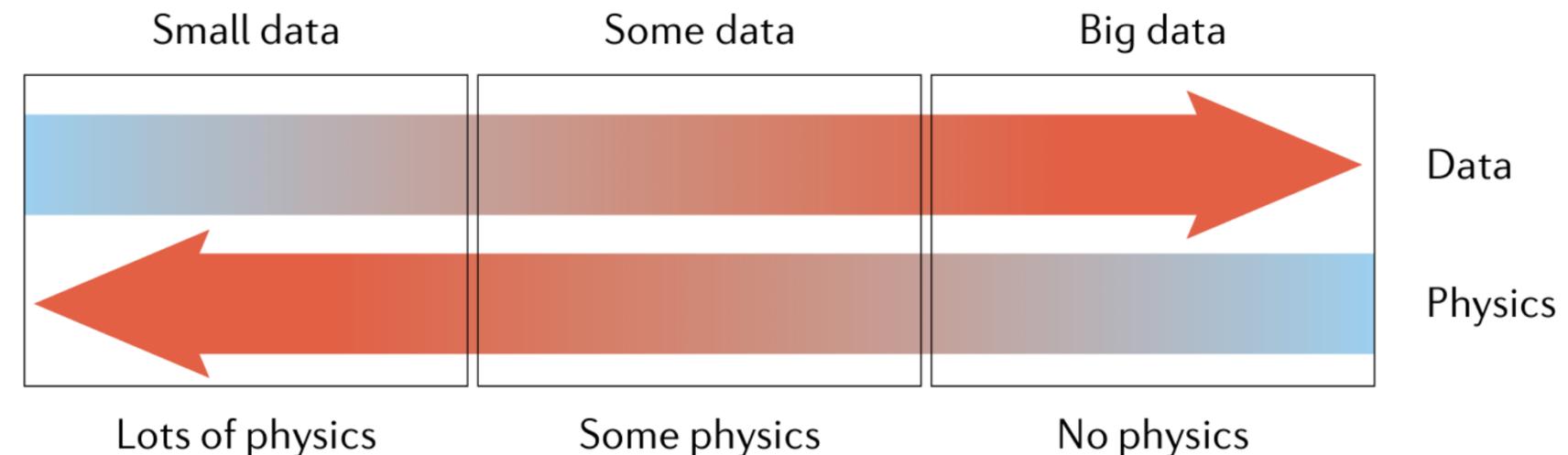
- Необходимость создания **обоснованных моделей искусственного интеллекта** для **реалистичного описания** процессов и явлений.
- Возрастающий рост **объема и сложности** собираемых сигналов, изображений и иных датасетов в **прикладных областях**.
- **Специфика данных**: **ограниченность** доступных **обучающих наборов**, наличие **неоднородностей** и **дисбаланс классов** объектов.

Текущее положение

- **Малые данные.**
 - Известна математическая формализация исследуемого физического процесса/явления (чаще всего, **детерминированная модель**).
 - Данные доступны только для начальных и граничных условий.
 - Нейросеть – **альтернатива численным методам**.
- **Промежуточный вариант.**
 - Модели известны **частично**.
 - Данных для обучения достаточно **ограничены по объему**.
 - **Симбиоз** классического обучения нейросетей и математических моделей.
- **Большие данные.**
 - Физические модели отсутствуют, но накоплен **огромный объем данных**.
 - Нейропотенциал для открытия **новых физических взаимосвязей**.

Вызовы

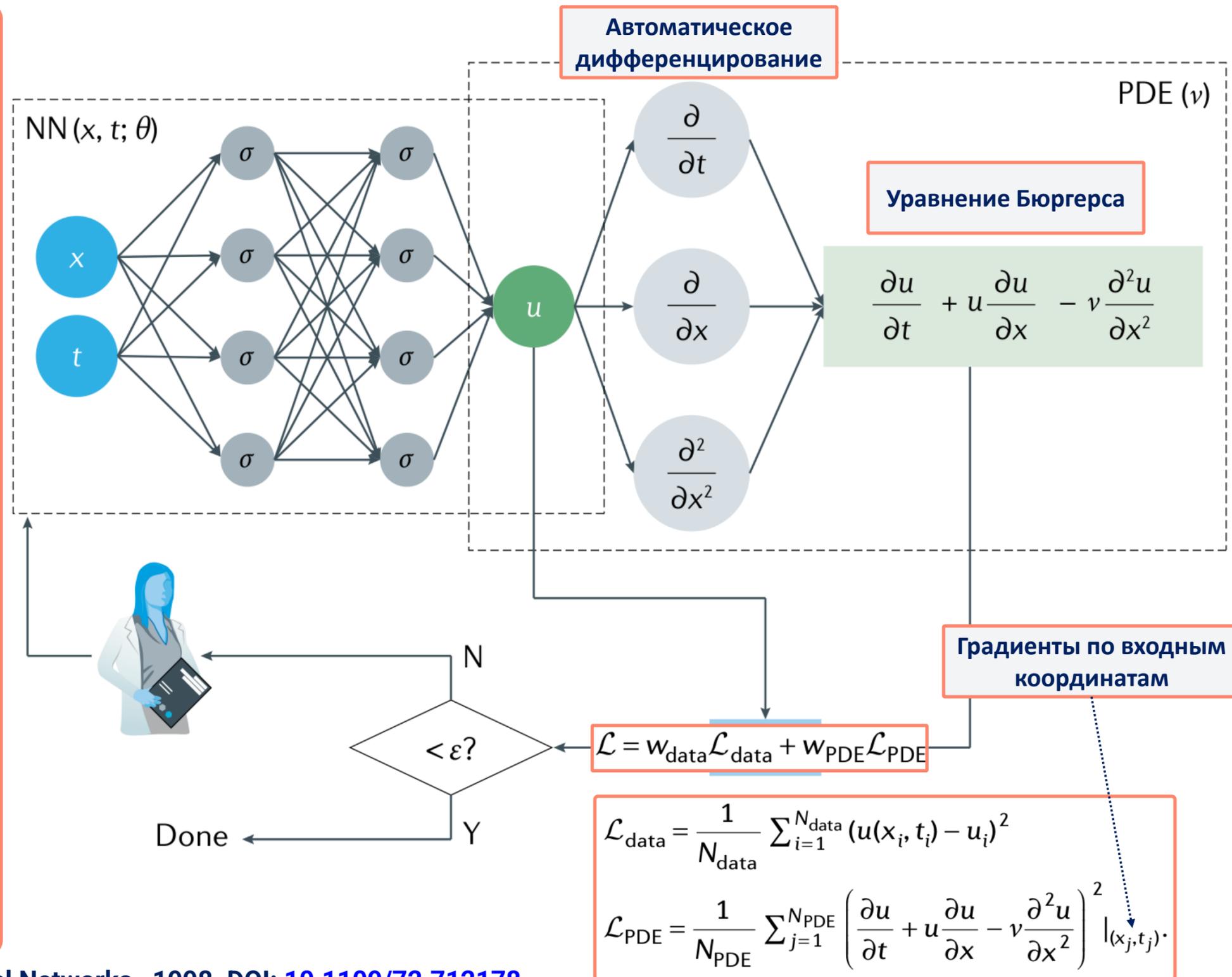
- Интеграция в архитектуру нейросетевых методов **стохастических моделей гетерогенных** (ограниченных по объему, зашумленных, существенно неоднородных) **сигналов**.
- Создание моделей меньшего размера, чем сопоставимые по качеству аналоги: **малая модель для большой задачи**.
- Теоретическое **доказательство** свойств получаемых **адаптивных** нейросетевых моделей.
- Потенциальное превосходство над **лучшими** специализированными современными архитектурами, включая трансформерные.
- **Масштабирование и интеграция** в комплексные современные архитектуры.



ФИЗИЧЕСКИ-ИНФОРМИРОВАННЫЕ НЕЙРОННЫЕ СЕТИ



- **Физически-информированные модели** используют для обработки **многомерных сигналов**, причем соответствующие **процессы** допускают **описание** с использованием **детерминированных физических моделей**.
- Объединение **данных** и **физических законов** в обучении нейронной сети.
- Физически-информированные сети используют физические законы как **регуляризатор функции потерь**:
 - Обучение модели на **малом объеме данных**.
 - Получение **физически корректных результатов**.
 - Возможность эффективного **решения** как **прямых**, так и **обратных задач моделирования**.
- Преимущества **автоматического дифференцирования**:
 - **Отсутствие ошибок аппроксимации**, характерных для численного дифференцирования на сетке.
 - **Высокая скорость вычисления производных**, сопоставимая со скоростью прямого прохода.
 - **Гибкость и адаптивность**: возможность быстрой замены класса математических уравнений в модели.
- Отказ от традиционных сеточных методов позволяет избегать «**проклятия размерности**».



□ Lagaris I.E., Likas A., & Fotiadis D.I.: «IEEE Transactions on Neural Networks» 1998. DOI: [10.1109/72.712178](https://doi.org/10.1109/72.712178)

□ Raissi M., Perdikaris P., Karniadakis, G.E.: «Journal of Computational Physics», 2019. DOI: [10.1016/j.jcp.2018.10.045](https://doi.org/10.1016/j.jcp.2018.10.045)

□ Karniadakis G.E., Kevrekidis I.G., Lu L. et al.: «Nature Reviews Physics», 2021. DOI: [10.1038/s42254-021-00314-5](https://doi.org/10.1038/s42254-021-00314-5)

ТЕОРЕТИЧЕСКИЕ РЕЗУЛЬТАТЫ ДЛЯ ИНФОРМИРОВАНИЯ НА УРОВНЕ ФУНКЦИИ ПОТЕРЬ



- **Полная ошибка аппроксимации** E_{total} (разница между истинным решением u^* и предсказанием нейросети u_θ) ограничивается суммой трех факторов:

$$E_{total} \leq E_{approx} + E_{gen} + E_{opt},$$

- E_{approx} (**ошибка аппроксимации**): минимально возможная ошибка для выбранной архитектуры сети;
 - E_{gen} (**ошибка обобщения**): разница между ошибкой на N_{PDE} обучающих точках (**коллокации**), в которых проверяется соблюдение физического закона, и ошибкой во всей непрерывной области;
 - E_{opt} (**ошибка оптимизации**): отклонение текущих весов сети от оптимальных из-за работы градиентных методов.
- При соблюдении условий регулярности, **ошибка стремится к нулю** при увеличении **ширины** (количество нейронов в слое) сети (m):

$$\|u^* - u_\theta\|_{L^2} \leq C \left(E_{approx}(m) + \frac{1}{\sqrt{N_{PDE}}} + E_{opt} \right),$$

где $C > 0$ – некоторая константа, выбирающаяся для конкретного «информирующего» уравнения. Можно отметить использование квадратур Монте-Карло для оценок в функции потерь.

- Для широкого класса PDE (включая нелинейные параболические уравнения) **сложность** сети S (количество параметров) растет **полиномиально от размерности**:

$$S(\varepsilon, d) \leq C \cdot d^q \cdot \varepsilon^{-\gamma},$$

где d – размерность задачи (количество пространственных координат), при этом величины q и γ **от нее не зависят**. Это подтверждает преимущество над методами с экспоненциальной сложностью.

ВИДЫ ИНФОРМИРОВАНИЯ НЕЙРОННЫХ СЕТЕЙ



Современное состояние исследований

- Высокая теоретическая обоснованность информирования на уровне функции потерь, прежде всего, в случае использования **детерминированных моделей**.
- Известны результаты в области **моделирования решений уравнений**, в том числе и **стохастических**, с использованием нейронных сетей. В основном – специальные блоки в архитектурах-решателях, однако этот подход **не является** в полном смысле **информированием**.
- Недостаточно широко исследованы вопросы **обоснования эффективности подходов архитектурного информирования**, прежде всего, в части использования **вероятностных моделей**. Такие блоки должны реализовывать **структурные ограничения** на решения сети.
- Недостаточно изучены подходы в области информирования в случае **принципиально ограниченных** (по ряду характеристик) анализируемых **датасетов**, в том числе для методов **информирования на уровне признаков**.

КОМБИНИРОВАННОЕ ИНФОРМИРОВАНИЕ: ПРИЗНАКИ + МОДИФИЦИРОВАННЫЕ САМООБУЧАЕМЫЕ ФУНКЦИИ ПОТЕРЬ



- Рассматривается задача прогнозирования, то есть построения модели f , которая наилучшим образом осуществляет предсказание целевой переменной Y на основании набора признаков X такая, что $f: X \rightarrow Y$.
- Рассматривается вариант информирования на основе нетривиального расширения признакового пространства X за счет выделения новых признаков X^+ и построение модифицированной модели $f^+: (X \cup X^+) \rightarrow Y$.

- Математическая модель основана на стохастическом дифференциальном уравнении Ито:

$$dX(t) = A(t) dt + B(t) dW(t),$$

где $X(t)$ – стохастический процесс, $W(t)$ – стандартный винеровский процесс, $A(t)$, $B(t)$ – случайные функции.

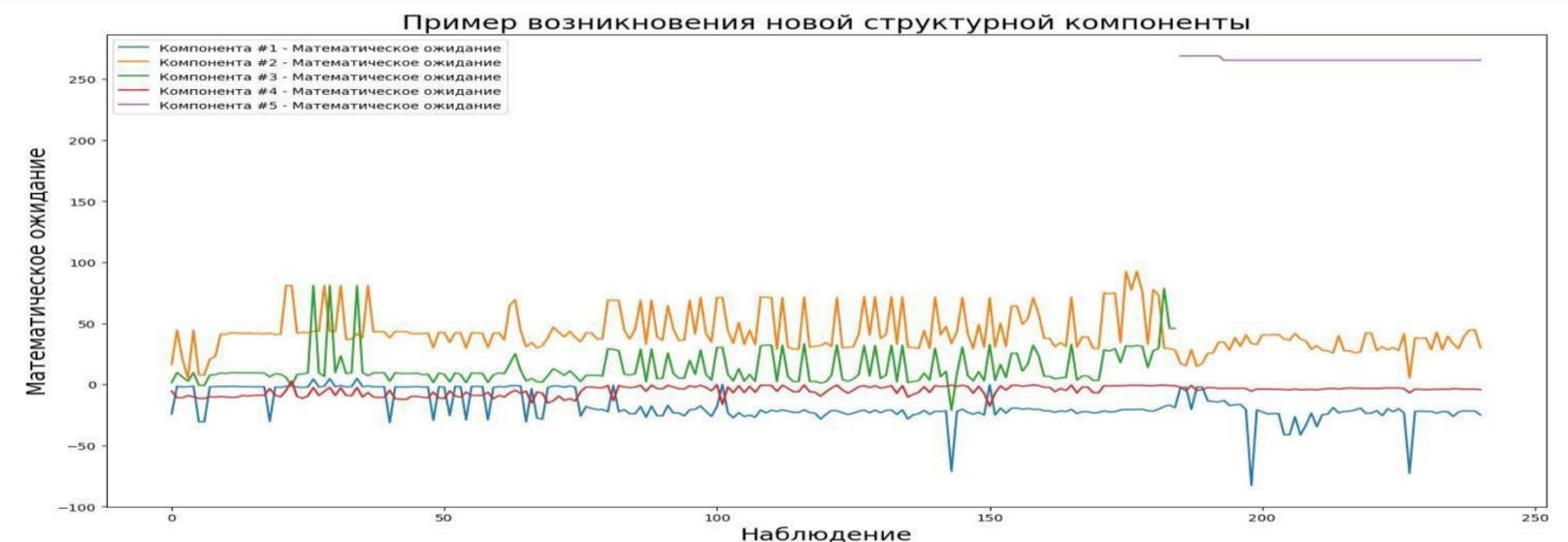
Теорема.

Пусть $f(x)$ – полносвязная однослойная нейронная сеть, использующая для восстановления дважды дифференцируемой функции $F(x)$, а h – среднее арифметическое расстояний между элементами нормализованного набора. Тогда информирование K -компонентной смесью нормальных распределений с плотностью $p(\cdot)$ на уровне признаков X в виде преобразования вектора, уменьшает погрешность восстановления $F(x)$:

$$\mathbb{E} \left(\int (F(x) - f(p(x)))^2 dx \right) = O(h^4).$$

- Информирование:

- Формирование матрицы взаимосвязей параметров смешанных моделей, получаемых в режиме скользящего окна. Выделение связанных структурных компонент с помощью кластеризации.



- Модифицированная функция потерь:

$$LOSS = (1 - \omega) \frac{1}{n} \sum (y_i - \hat{y}_i)^2 + \omega \frac{1}{n} \sum (f(y_i) - \hat{f}(\hat{y}_i))^2,$$

где, $\omega \in [0, 1]$ – вес, $f(y_i)$, $\hat{f}(\hat{y}_i)$ – гладкие функции, аппроксимирующие плотности предсказанных и настоящих значений целевой переменной.

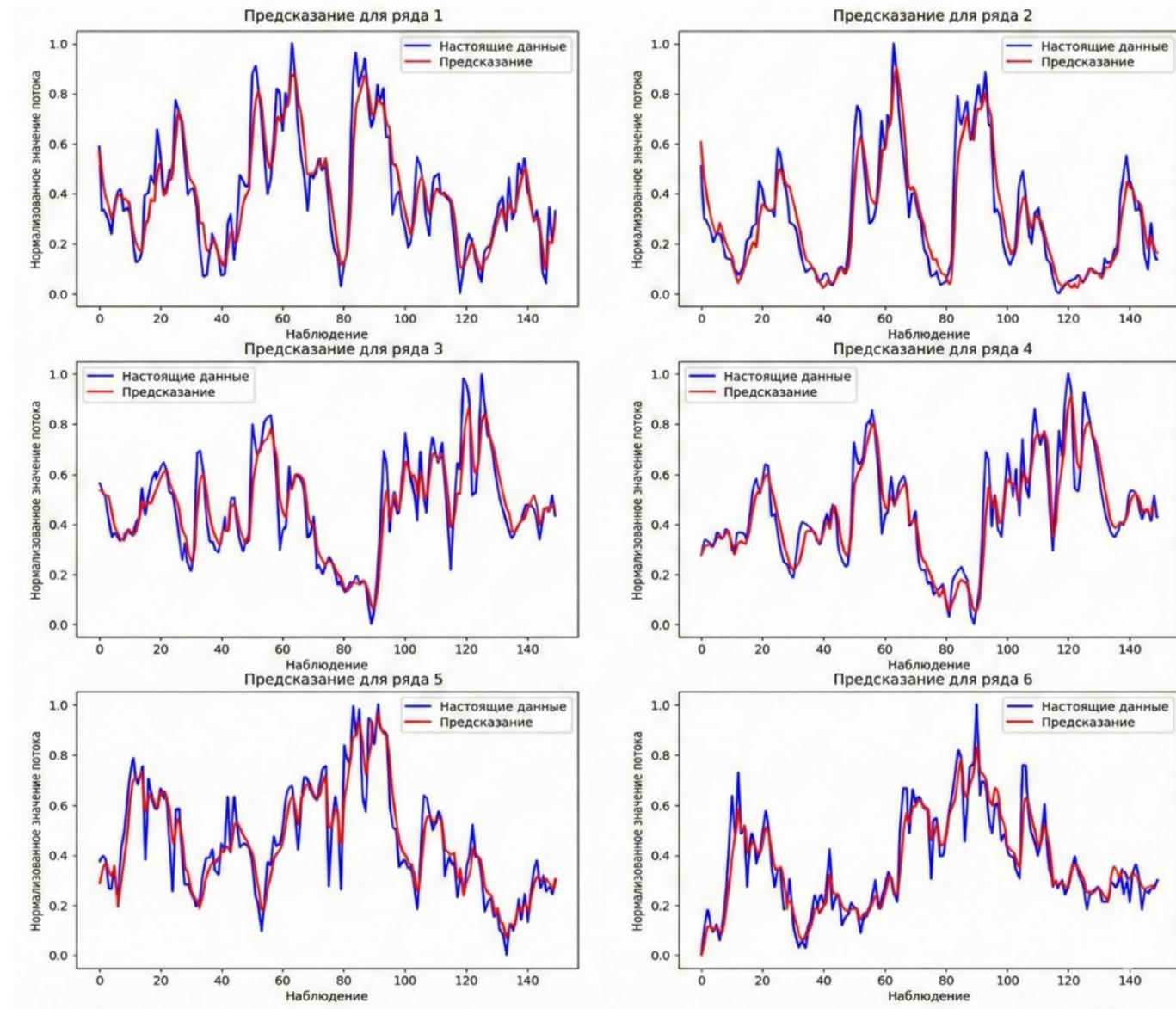
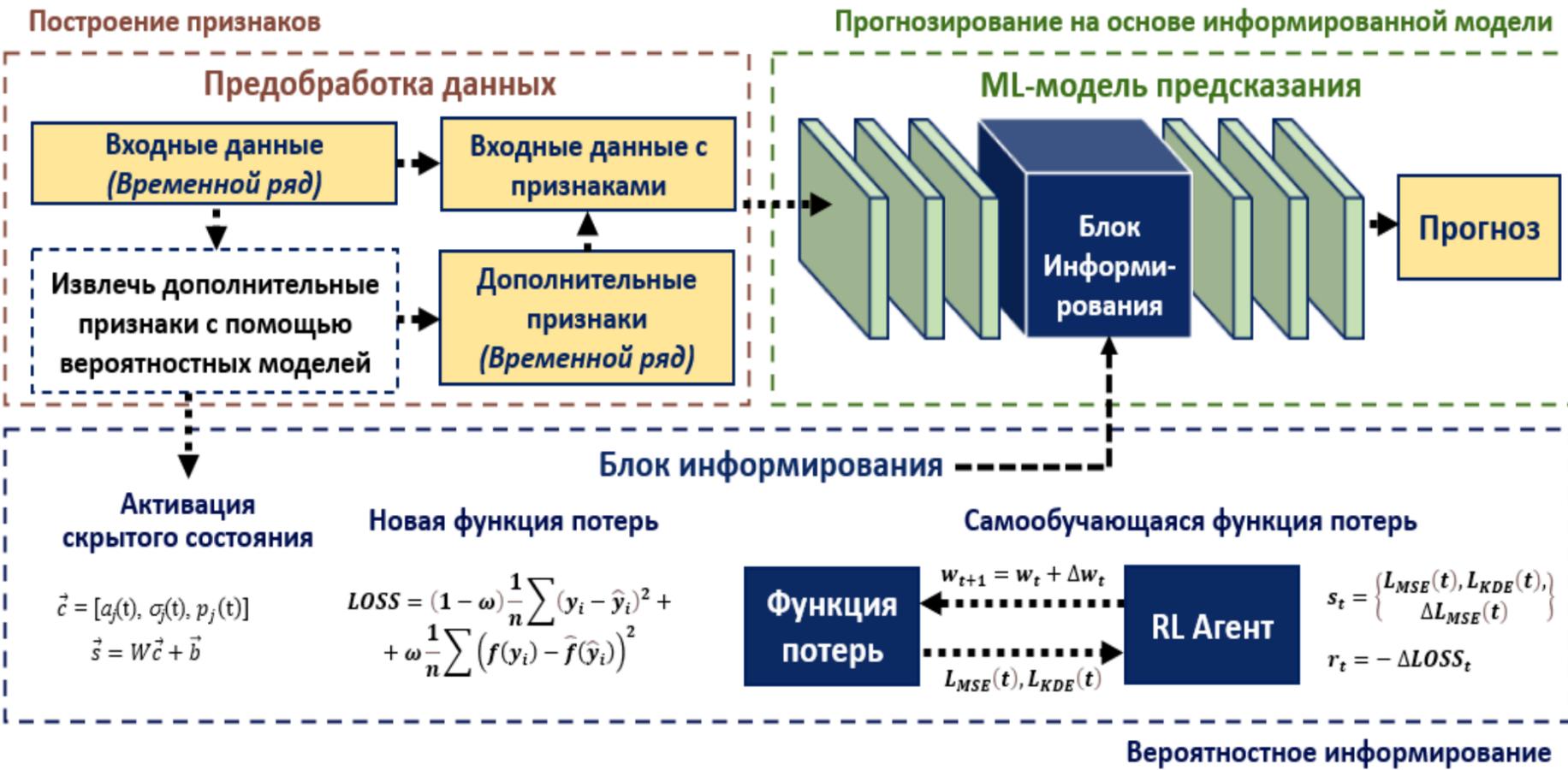
❑ Gorshenin A.K., Kuzmin V.Yu.: «Mathematics», 2022. DOI: [10.3390/math10040589](https://doi.org/10.3390/math10040589)

❑ Gorshenin A.K., Vilyaev A.L.: «AI», 2024. DOI: [10.3390/ai5040097](https://doi.org/10.3390/ai5040097)

❑ Dostovalova A., Gorshenin A.: «Neural Computing and Application», 2025. DOI: [10.1007/s00521-025-10997-x](https://doi.org/10.1007/s00521-025-10997-x)

❑ Gorshenin A.K., Vilyaev A.L.: «Pattern Recognition and Image Analysis», 2025. DOI: [10.1134/S1054661825601418](https://doi.org/10.1134/S1054661825601418)

КОМБИНИРОВАННОЕ ИНФОРМИРОВАНИЕ: ПРИЗНАКИ + МОДИФИЦИРОВАННЫЕ САМООБУЧАЕМЫЕ ФУНКЦИИ ПОТЕРЬ



Модель	Прирост RMSE
Random Forest	+22%
LSTM	+28%
minLSTM	+24%
LSTM + Трансформер	+35%
Mamba	+16%

СТОХАСТИЧЕСКИ-ОБОСНОВАННЫЕ МОДЕЛИ ПРОСТРАНСТВЕННО-ВРЕМЕННЫХ ДАННЫХ



➤ Математическая модель для параметров энергообмена между океаном и атмосферой основана на стохастическом дифференциальном уравнении Ито.

➤ В дискретной форме:

$$X(t + \Delta t) = a(t, X)\Delta t + b(t, X)(W(t + \Delta t) - W(t)).$$

➤ Конечномерное разложение Карунэна-Лозва:

$$X(t + \Delta t) = X(t) + a(t, X)\Delta t + \sum_{i=1}^N \lambda_i e_i \eta_i,$$

где $\eta_i = (e_i, W(t + \Delta t) - W(t))$.

Вызовы

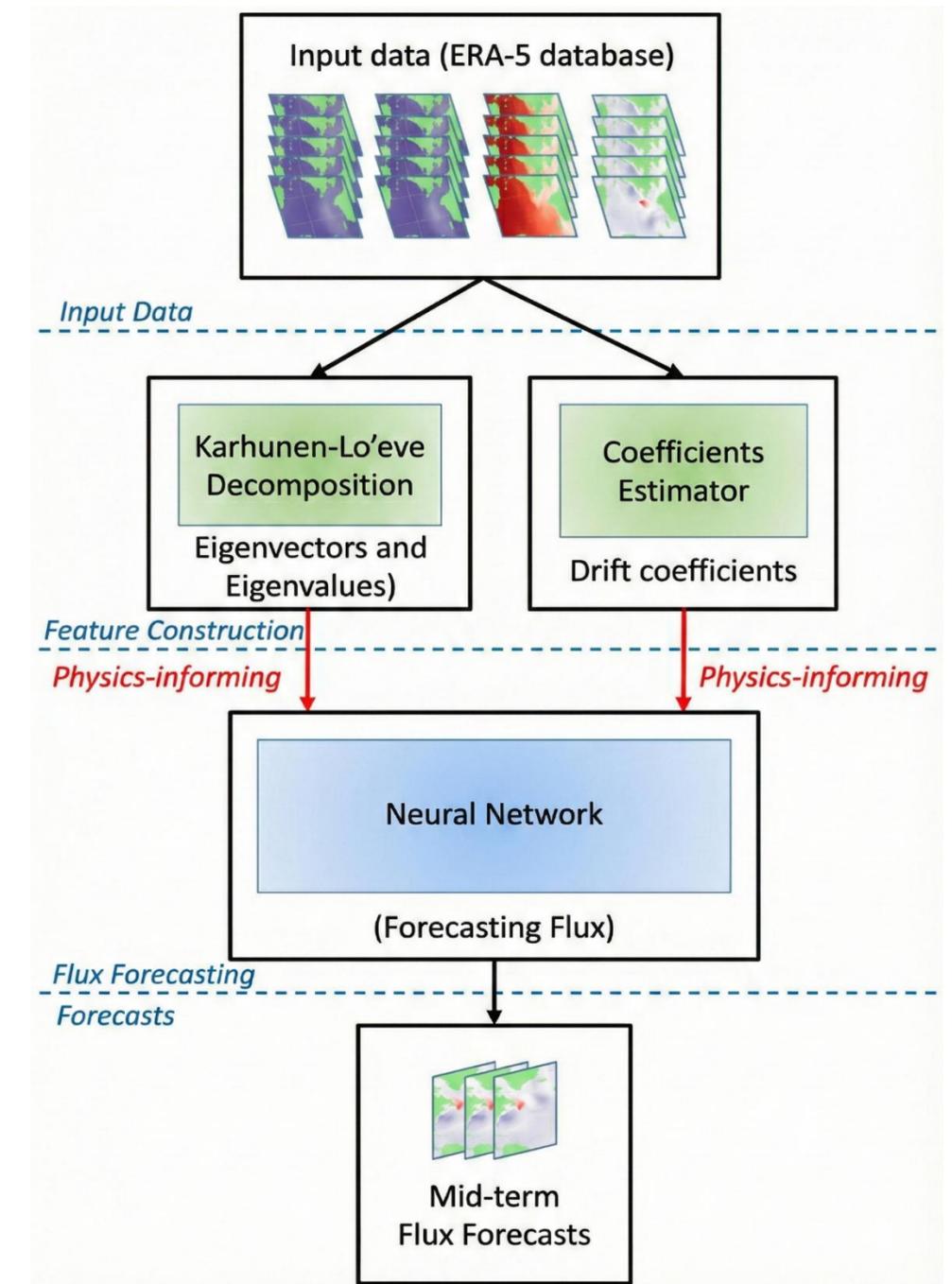
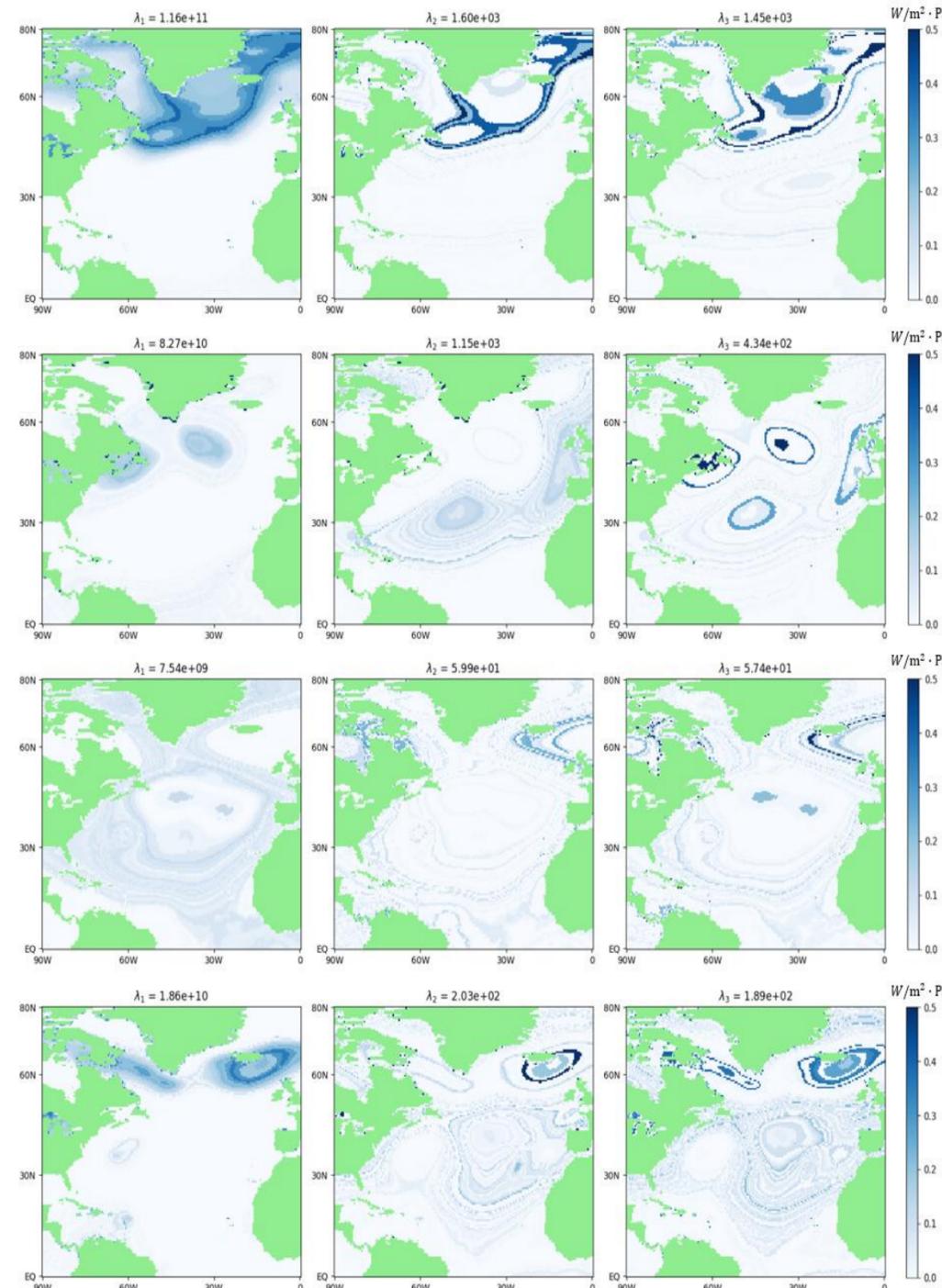
➤ Информирование признаками.

➤ Информирование на уровне функции потерь:

➤ +10% точности в базовом варианте;

➤ разностные схемы типа Эйлера-Маруямы.

➤ Возможность превзойти горизонт прогноза в 10-14 дней.



ГЛУБОКИЕ СМЕШАННЫЕ ГАУССОВСКИЕ МОДЕЛИ: АРХИТЕКТУРНОЕ ИНФОРМИРОВАНИЕ ДЛЯ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ



- Прогнозирование астрофизических характеристик с использованием методов полуконтролируемого обучения для малых датасетов с возможностью применения для **автоматизированного обнаружения космических объектов**.
- Новый метод формирования псевдо-меток данных для эффективной обработки нелинейных взаимосвязей признаков на основе **глубоких гауссовских смешанных моделей**:

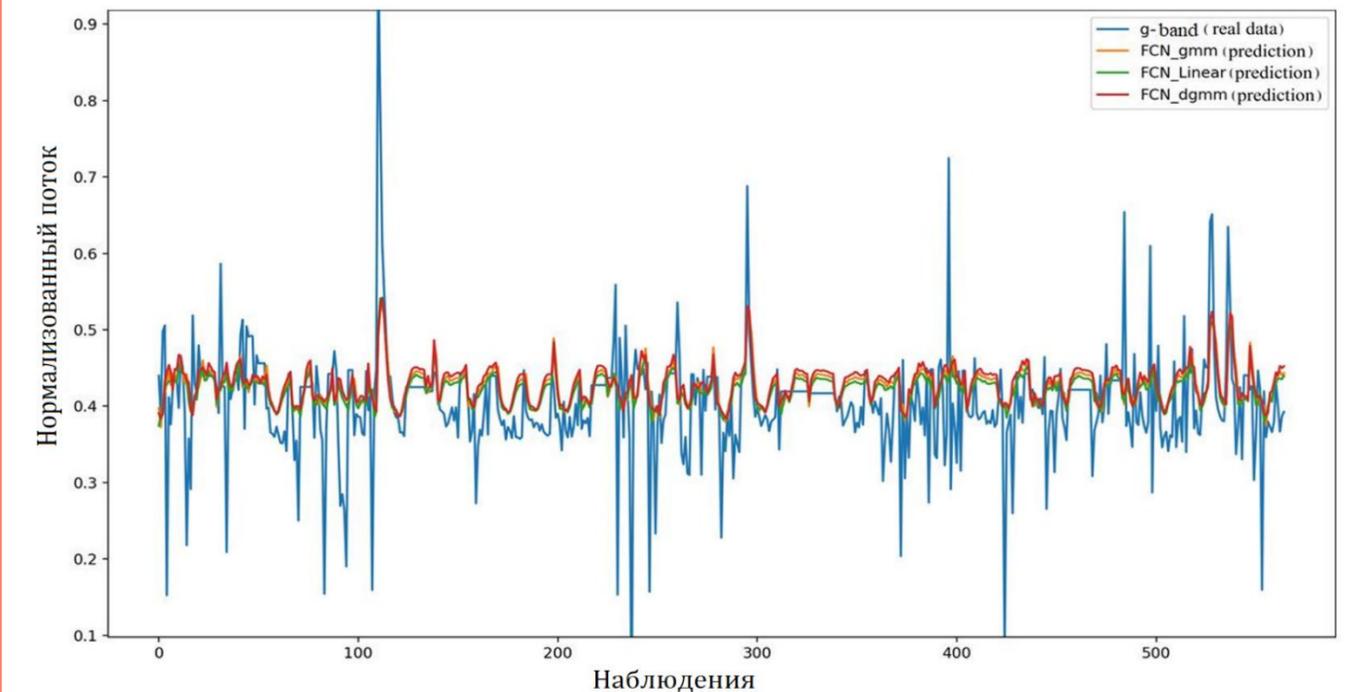
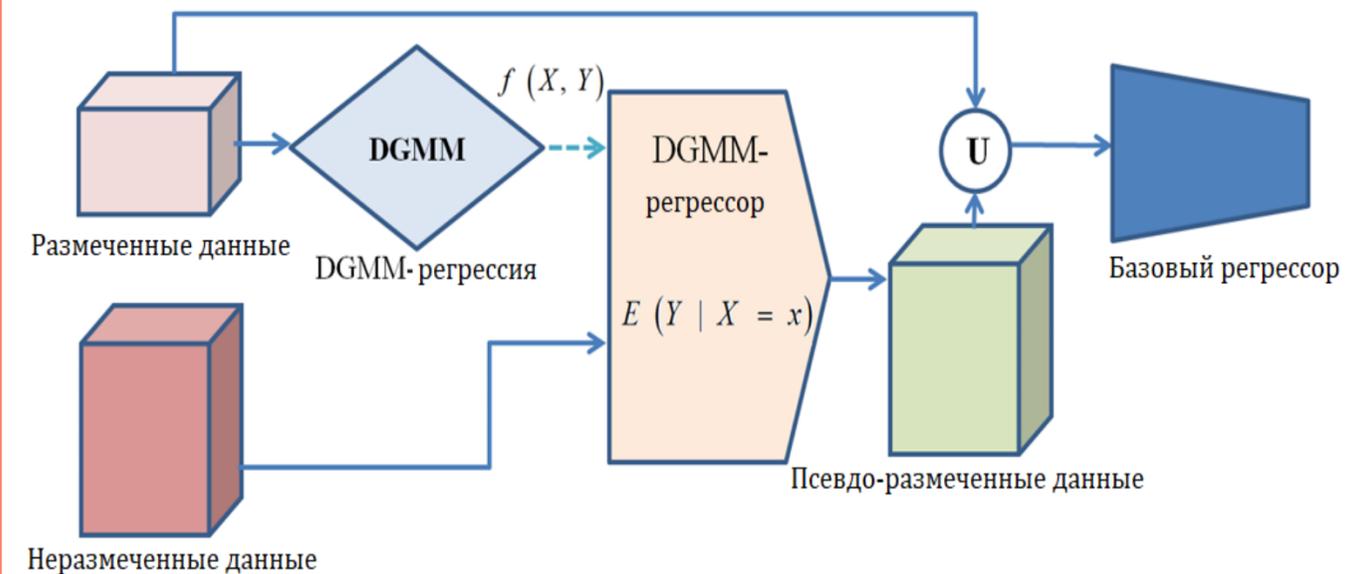
$$(1) x_i = \eta_{s_1}^{(1)} + \Lambda_{s_1}^{(1)} z_i^{(1)} + u_i^{(1)} \text{ с вероятностью } \pi_{s_1}^{(1)}, s_1 = \overline{1, k_1},$$

...

$$(h) z_i^{(h-1)} = \eta_{s_h}^{(h)} + \Lambda_{s_h}^{(h)} z_i^{(h)} + u_i^{(h)} \text{ с вероятностью } \pi_{s_h}^{(h)}, s_h = \overline{1, k_h},$$

где $i = 1, \dots, n$, $z^{(h)} \sim N(0, I_p)$ – **скрытые переменные**, $u^{(l)} \sim N(0, \Psi_{s_l}^{(l)})$, $l = 1, \dots, h$ – **случайные ошибки**, $\eta_{s_1}^{(1)}, \dots, \eta_{s_h}^{(h)}$ – **векторы математических ожиданий** длины p , $\Lambda_{s_1}^{(1)}, \dots, \Lambda_{s_h}^{(h)}$ – **квадратные матрицы факторных нагрузок** размерности p .

- На открытых данных значений **потока светового излучения сверхновых звезд** типа Ia (система парных звезд, одна из которых – белый карлик) репозитория Multimodal Universe продемонстрирован **прирост точности прогнозирования по метрике RMSE до 21,58%** относительно результатов других методов формирования псевдометок (доступен лишь **небольшой объем** наблюдений для обучения).
- **Точность сопоставима с трансформерами**, при этом модели имеют **меньшее число параметров и более простую архитектуру**.



□ Gorshenin A.K. et al.: «Computer Networks», 2024. DOI: [10.1016/j.comnet.2024.110433](https://doi.org/10.1016/j.comnet.2024.110433)

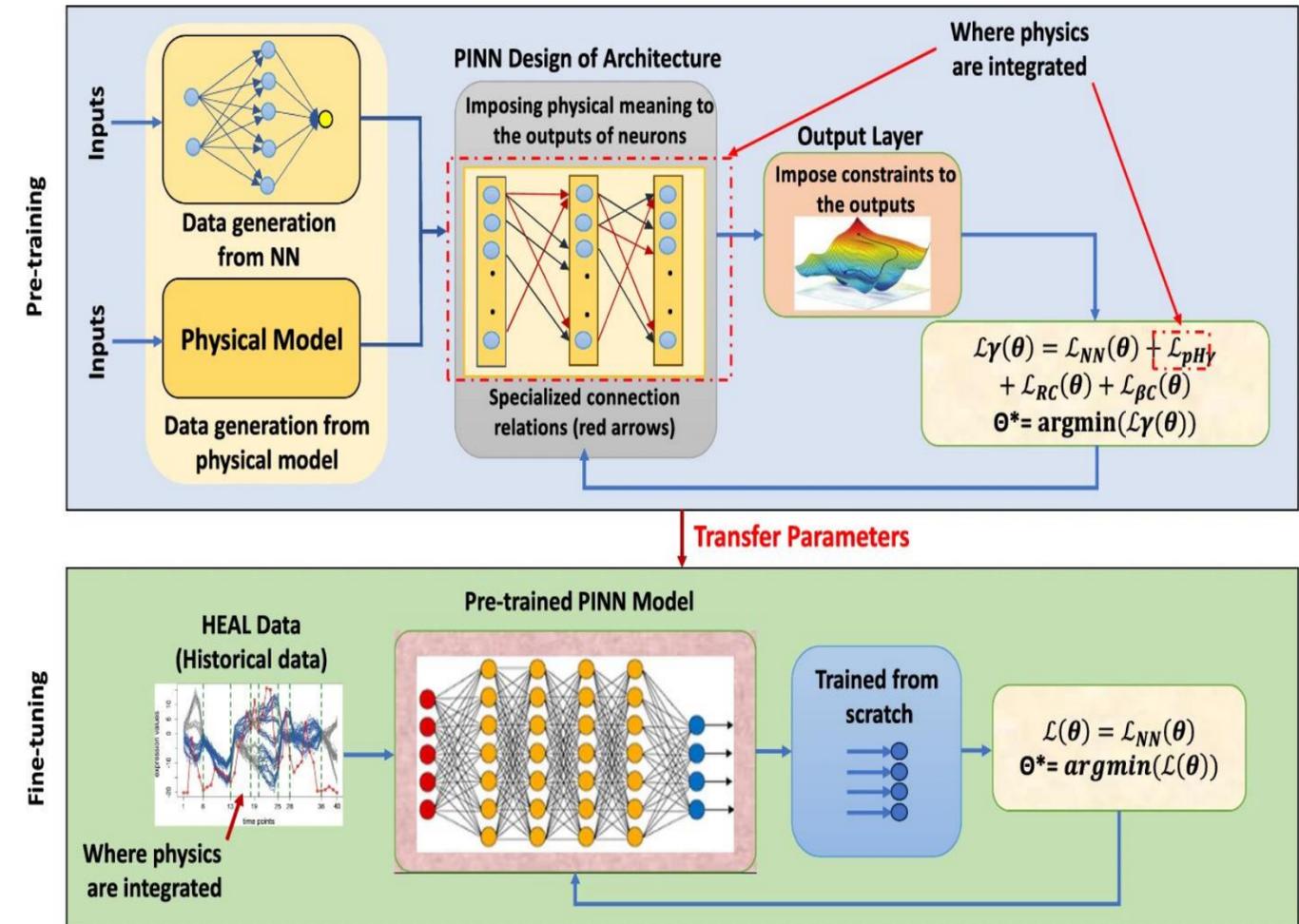
□ Gorshenin A.: «Lecture Notes in Computer Science», 2025. DOI: [10.1007/978-3-031-80853-1_26](https://doi.org/10.1007/978-3-031-80853-1_26)

□ Gorshenin A.K., Dostovalova A.M.: «Pattern Recognition and Image Analysis», 2025. DOI: [10.1134/S1054661825700014](https://doi.org/10.1134/S1054661825700014)

ИНФОРМИРОВАННЫЕ МЕТОДЫ ОБРАБОТКИ ИЗОБРАЖЕНИЙ (МЕДИЦИНСКИЕ ПРИЛОЖЕНИЯ)



Задача	Физические принципы/уравнения	Практическое применение	Способ информирования
Реконструкция изображений	Модели ослабления рентгеновского излучения	Неврология, кардиология, онкология	Функция потерь, входные признаки
Сегментация, детекция	Анатомические приоры, контрастные свойства тканей (МРТ), геометрические ограничения, физические приоры формы (КТ)	Выделение опухолей мозга, кардиальных структур, сосудистых аномалий	Функция потерь, входные признаки
Шумоподавление	Модели радиоактивного распада, кинетические модели, транспорт фотонов (PET/SPECT), уравнения Навье-Стокса, гидродинамика (УЗИ)	Онкология, неврология, гемодинамика, оценка стенозов	Функция потерь



АРХИТЕКТУРНОЕ ИНФОРМИРОВАНИЕ МОДЕЛЯМИ ФАКТОРНЫХ АНАЛИЗАТОРОВ



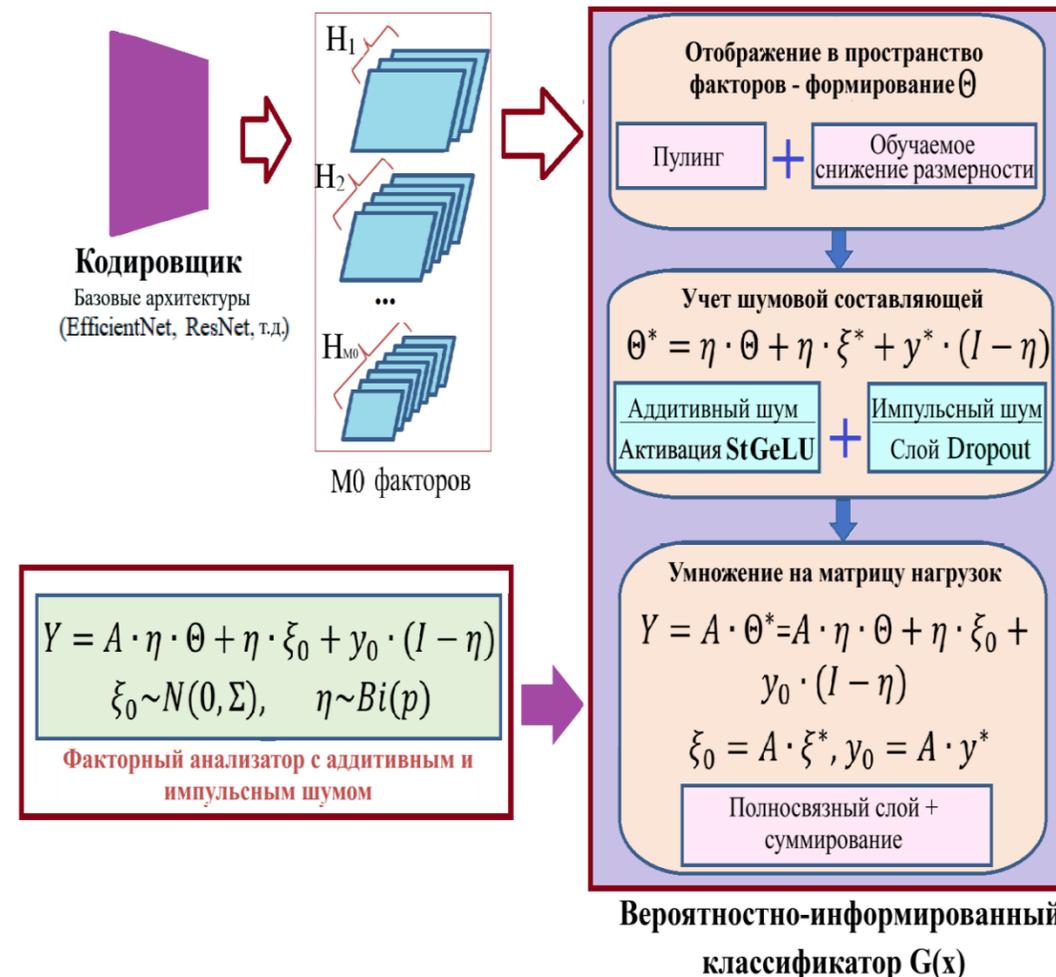
Пусть задан малый набор зашумленных изображений $X = \{X^i\}_{i=1\dots N}$ размера $H_X \times W_X$, а $ns(\cdot)$ - функция, описывающая искажение. Для повышения информативности X используются M_0 дополнительных признаков изображения $\{X_l^i\}_{l=1\dots M_0}$, $X_l^i \in \mathbb{R}^{H_l \times W_l}$, $H_l < H_X$, $W_l < W_X$, полученных, например, из скрытых слоев кодировщика $Enc(\cdot)$.

Ставится задача разработать нейросетевую модель $G(\cdot)$ слияния признаков $\{X_l^i\}_{l=1\dots M_0}$ с учетом шума: $G: \mathbb{R}^{X_1} \times \dots \times \mathbb{R}^{X_{M_0}} \rightarrow \mathbb{R}^K$ для повышения вероятности верной классификации элементов набора X сетью $f_X(\cdot) = H(Enc(\cdot))$: $\mathbb{P}(f_X(G(ns(Enc(X^i)))) = Y^i) > \mathbb{P}(f_X(X^i) = Y^i)$.

- Новая нейросетевая модель, использующая механизм взаимодействия факторов в анализаторе с аддитивным и импульсным шумами для повышения качества объединения признаков.
- Эффективность основывается на доказанных теоремах об устойчивости оценок параметров модели к искажениям и о значениях гиперпараметров, гарантирующих построение более вычислительно простого классификатора.

Теорема.

Пусть $scale$ – размерность фактора, $M \in \mathbb{N}$ – их количество, f – исходный размер фактора, K – число классов и $M \cdot scale \cdot K \in \mathbb{N}$, а для уменьшения размерности выходного слоя кодировщика $N_{M_0} \times N_{M_0} \times H_{M_0}$ используется пулинг по среднему значению. Тогда число операций в классификаторе $FtFNN$ меньше, чем в базовом полносвязном классификаторе, если $M < M^*$, где $M^* \propto \frac{1}{K \cdot scale}$, при условии, что снижение размерности фактора реализовано полносвязным слоем, и при $scale < scale^*$, $scale^* \propto \frac{1}{(f^2 + K)K}$, для сверточных слоев.



- Классификация объектов и поверхностей на оптических спутниковых изображениях в условиях малых обучающих наборов, которая может применяться для задач обнаружения и мониторинга.
- Во всех случаях превосходит аналоги без информирования: на открытых данных со спутника прирост значений метрики Top-1 Ассигасу достиг 19,8% (с базового среднего уровня в 85%).
- Архитектура содержит в 1,57 раз меньше параметров, чем сопоставимые по результатам модели.

АРХИТЕКТУРНОЕ ИНФОРМИРОВАНИЕ ДЛЯ УЧЕТА МУЛЬТИМАСШТАБНЫХ ВЗАИМОСВЯЗЕЙ

- Пусть $X = \{X^k\}_{k=1\dots N}$ - набор **сегментируемых** на K классов изображений. Предположим, что разделяемые классы **не сбалансированы** по числу элементов:

$$N_{(1)} \ll \dots \ll N_{(K)}.$$

- Пусть $X_{cat}^k = (F(X_{inner}^k), F(X^k), X^k)$, где $F(X^k) \in \mathbb{R}^{H_X \times W_X \times K}$ - сегментация снимка $F(\cdot)$, а $F(X_{inner}^k)$ - признаки X^k , полученные из **внутренних слоев** сети, где $N_{ch} = ch + K + 3$, ch - число каналов изображения.

- Для повышения вероятности правильной сегментации **несбалансированных классов** ставится задача разработать нейросетевую модель $G(\cdot)$ для более точной обработки **локальных пространственных связей между внутренними признаками** $F(X_{inner}^k)$ за счет информирования **вероятностной моделью** (Y^k реальные метки классов пикселей X^k):

$$\mathbb{P}(G(X_{cat}^k, A) = Y^k) > \mathbb{P}(F(X^k) = Y^k).$$

Теорема.

Пусть заданы нейронные сети $G_Q^*(\cdot)$ – двухветочная графовая архитектура, сверточная $F(x) = X * w$ и графовая $G(x) = A_G \cdot x^{(sp)} \cdot B_G$, причем $G_Q^*(\cdot)$ и $G(\cdot)$ обрабатывают изображение по суперпикселям размера M . Пусть $L_t(\cdot, y)$ – дифференцируемая функция потерь. Тогда:

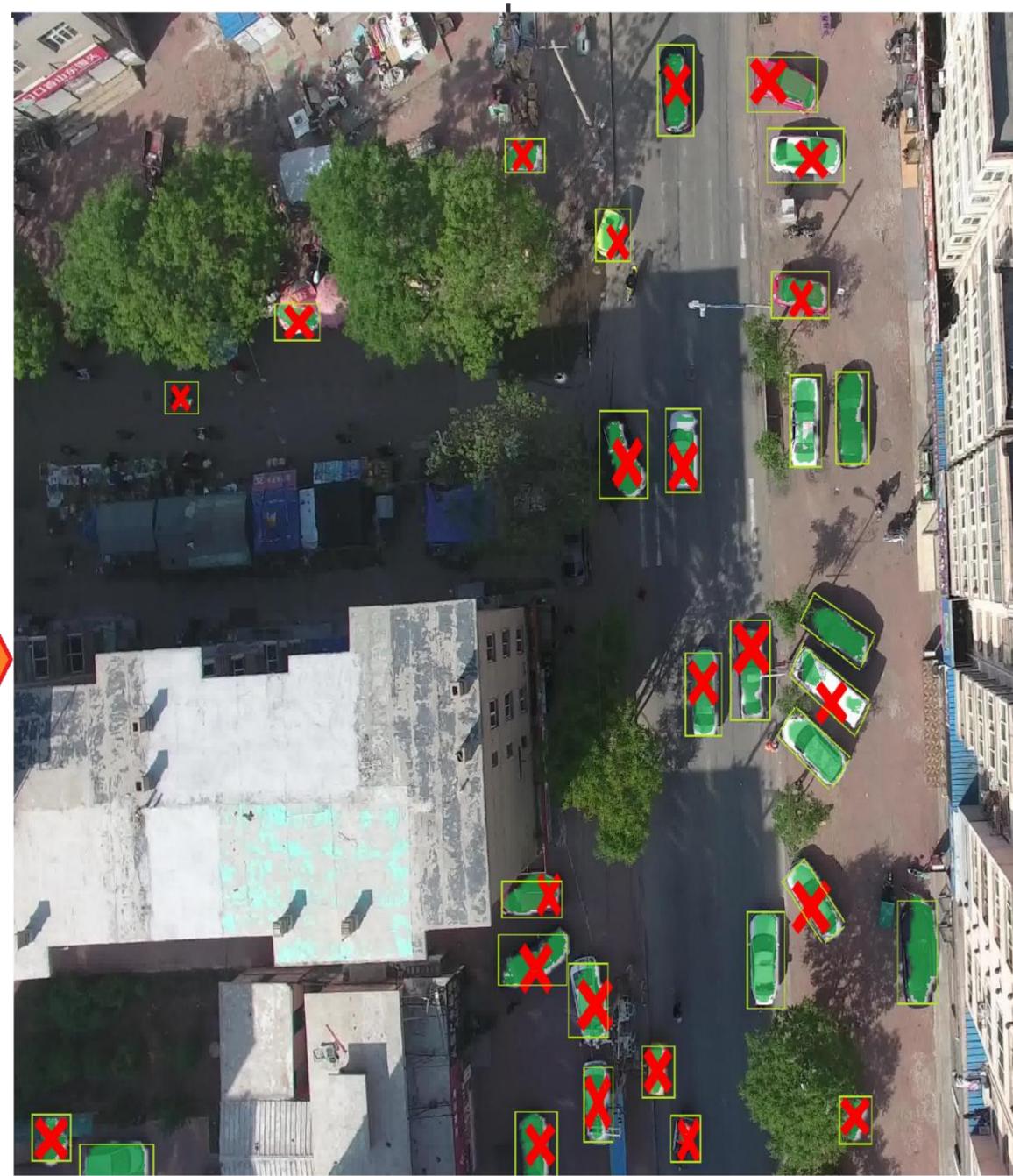
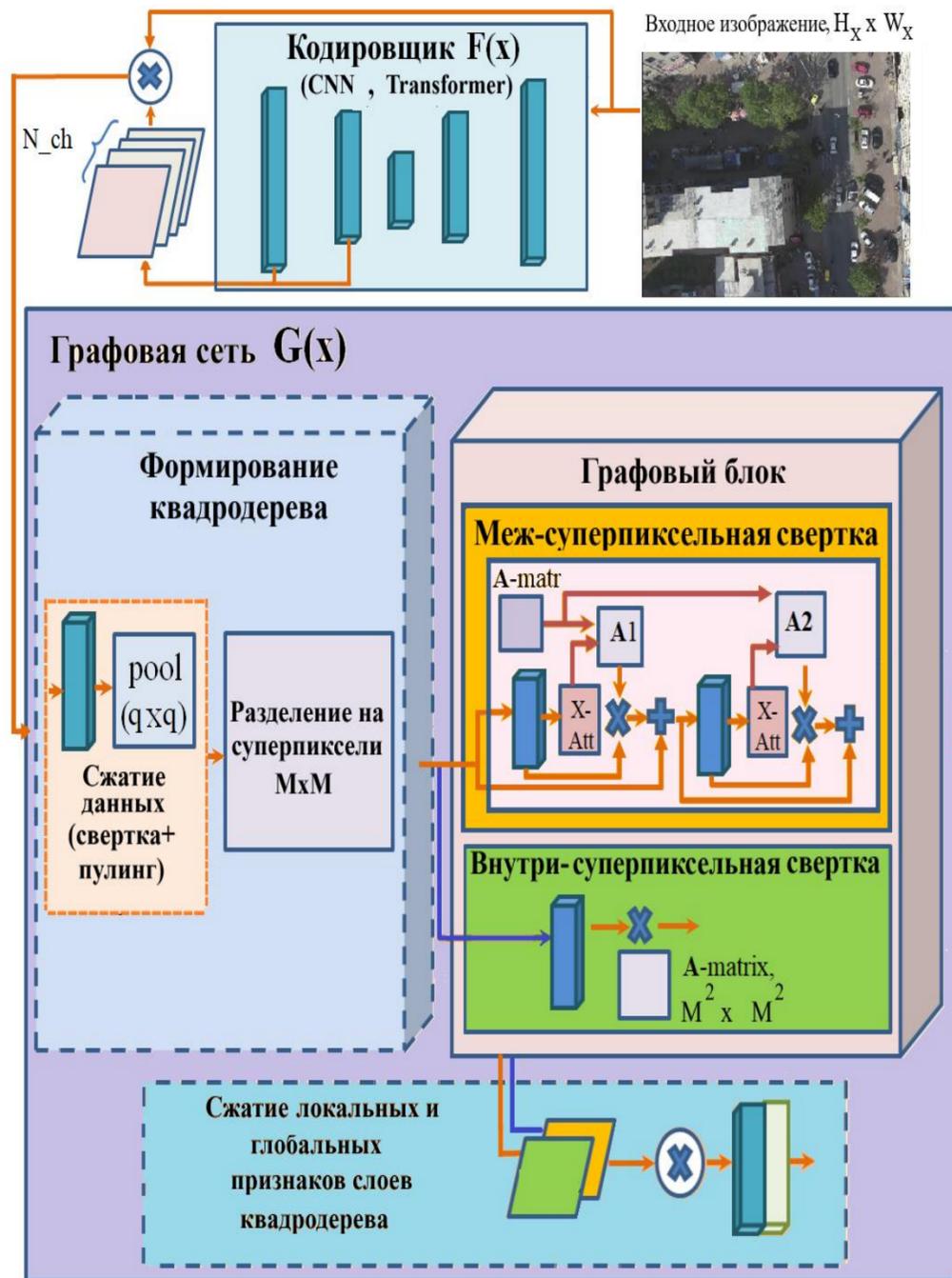
$$\Delta_t(L_t(G(x), y)) < \Delta_t(L_t(G_Q^*(x), y)),$$

$$\Delta_t(L_t(F(x), y)) < \Delta_t(L_t(G_Q^*(x), y)),$$

где y – вектор истинных значений, а t задает шаг обучения сети, то есть $G_Q^*(\cdot)$ обучается быстрее, чем сопоставимые по размеру нейросетевые решения.

- Семантическая сегментация изображений высокого разрешения **Ultra HD 4K (3840 × 2160)**, которая может применяться при анализе сложных сцен в рамках **автономной навигации для беспилотных систем**.
- Предложена **новая ансамблевая графовая нейросетевая модель** для преодоления ключевой проблемы известных методов – качественного восстановления взаимосвязей между элементами несбалансированных классов на изображениях высокого разрешения, содержащих **множество разномасштабных объектов**.
- Решение основано на интеграции в архитектуру математической модели изображения на основе **Марковских подходов**, а его эффективность базируется на доказанной теореме о более высокой скорости убывания функции потерь.

СЕМАНТИЧЕСКАЯ СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЙ ВЫСОКОГО РАЗРЕШЕНИЯ



- Предложенная архитектура превосходит по точности трансформерную SOTA-модель 2025 года для обработки аэрокосмических изображений LWGANet на открытых датасетах изображений БПЛА (UAVid, UDD) на **15,11%** по F_1 -мере.
- Реализован принцип «небольшая ИИ-модель для большой задачи»: предложенная архитектура содержит в **1,78 раз меньше параметров**, чем сопоставимые по результатам модели, применяемые в данной предметной области.
- Прирост точности для **крупных объектов** (дороги, дома, растительность) – до **29,83%** по F_1 -мере.
- Прирост точности для **малоразмерных объектов** (автомобили, корабли, причем их размер и форма отличаются) – до **66,24%** по F_1 -мере.

□ Gorshenin A., Dostovalova A.: «Big Data and Cognitive Computing», 2025. DOI: [10.3390/bdcc9090239](https://doi.org/10.3390/bdcc9090239)

□ Горшенин А.К., Достовалова А.М.: «Доклады РАН. Математика, информатика, процессы управления», 2025. DOI: [10.1134/S1064562425700255](https://doi.org/10.1134/S1064562425700255)

ВЕРОЯТНОСТНОЕ МОДЕЛИРОВАНИЕ МУЛЬТИМОДАЛЬНЫХ СРЕД



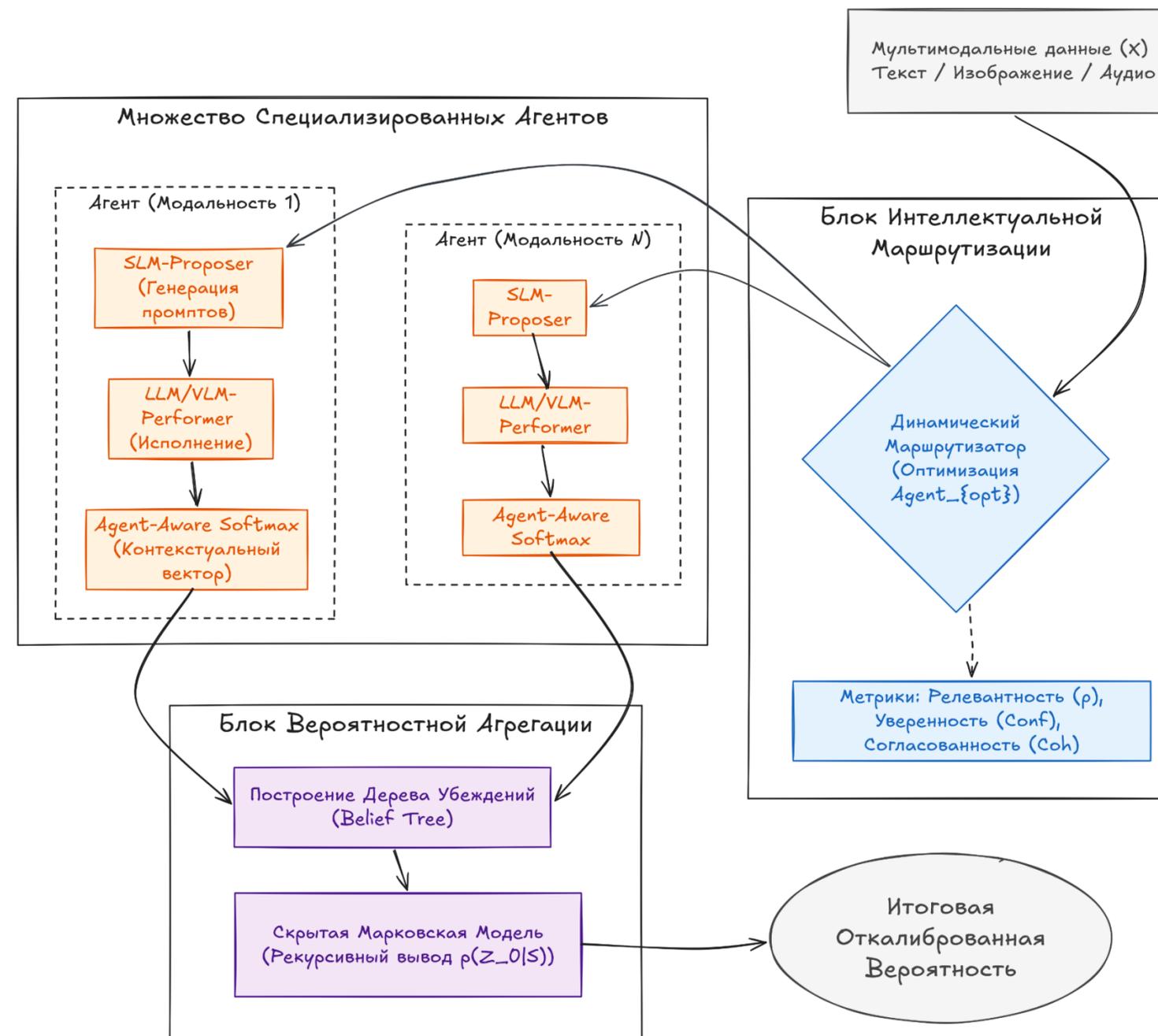
- «Умные» агенты как реализация информирования на новом уровне абстракции.
- Разработка **многоагентной мультимодальной системы**, способной осуществлять автоматизированную верификацию **разнородной информации** с учетом семантической релевантности, когнитивной уверенности (фактически, **информирование на уровне признаков**) и межагентной согласованности.

- **Стохастическая агрегация на деревьях убеждений**: интеграция мультимодальных выводов через скрытые марковские модели (как разновидность **архитектурного информирования**). Истинность гипотезы моделируется как скрытая переменная Z , а оценка агента — как наблюдаемая S . Рекурсивная схема **байесовского вывода**:

$$p(Z_0 = T | \{S_{\tau(0)}\}) = \frac{\beta(T, 0)p(Z_0 = T)}{\sum_{z \in \{T, F\}} \beta(z, 0)p(Z_0 = z)}$$

где $\beta(z, 0)$ — условная вероятность наблюдаемых оценок агентов в поддереве при скрытом состоянии z , принимающей значения «истина» или «ложь».

- На **небольшом тестовом датасете** продемонстрировано **преимущество** предложенного **агентного решения** над **одиночными LLM-моделями** (Mistral, Qwen, GPT – даже в вариантах до **175 млрд параметров**) на величину до **17% (F_1 -мера)**. Получено значительное снижение ошибки калибровки предсказаний (**Brier Score**) на величину до **30%**.



ПЕРСПЕКТИВЫ: РЕАЛИЗАЦИЯ СТОХАСТИЧЕСКИХ ПОДХОДОВ В МОДЕЛЯХ МИРА



- **Переход от моделирования одного объекта к моделированию целых взаимодействующих сред за счет комбинированного информирования.**
- **Смена парадигмы моделирования.**
 - Переход от классической максимизации правдоподобия к **обоснованному причинно-следственному выводу** (causal inference).
 - Информированная нейросеть вычисляет **эволюцию состояний** сложной системы $p_{\theta}(S_{t+1} | S_t, A_t, \mathcal{I})$, где \mathcal{I} — ограничения, в том числе и физические.
- **Преодоление барьера Out-of-Distribution.** Встраивание детерминированных и стохастических ограничений напрямую в латентное пространство предиктора для устойчивости симуляции за пределами обучающей выборки для минимизации **ошибки обобщения**.
- **Автономность и креативность.**
 - Поиск **нестандартных стратегий** (Monte Carlo Tree Search) в виртуальной среде.
 - Принятие решений в условиях **стохастической неопределенности**.
 - Моделирование сценариев для исключения галлюцинаций. **ограничено законами среды**

