

# Модельно-информированная нейросетевая классификация ОК-томограмм на основе Фурье-модели формирования изображения

Е. Н. Волков

ФИЦ ИУ РАН

envolkoff@gmail.com

**Аннотация.** Рассматривается задача классификации реконструированных изображений оптической когерентной томографии при отсутствии исходного интерференционного сигнала. Предлагается модельно-информированная регуляризация нейросетевого классификатора, основанная на дифференцируемом операторе повторной деградации скрытого представления. Оператор учитывает осевое размытие в Фурье-представлении, глубинный спад чувствительности и глубинно-зависимое поперечное размытие. В отличие от предварительного подавления спекл-шума, предложенная схема не направлена на восстановление эталонного изображения, а вводит в функцию потерь слагаемое самосогласованности между наблюдаемой томограммой и результатом повторной деградации скрытого представления. Подход не требует пар «зашумленное-эталонное» изображение и применим к реконструированным томограммам. Оценка выполнена на наборах данных OCTDL и MultiEYE. Модельно-информированная ИНС демонстрирует более высокие средние значения Masco-F1 и средней полноты по сравнению с базовой моделью, а также меньшую чувствительность к спекл-шуму и гауссовому шуму. Полученные результаты указывают на целесообразность использования эффективной модели деградации как регуляризирующего ограничения при классификации реконструированных ОК-томограмм.

**Ключевые слова:** оптическая когерентная томография; глубокое обучение; модельно-информированная регуляризация; классификация медицинских изображений; устойчивость к деградациям изображений

## I. ВВЕДЕНИЕ

Оптическая когерентная томография является интерферометрическим методом визуализации, в котором глубинный профиль объекта восстанавливается из интерференционного сигнала с использованием преобразования Фурье [1]. В офтальмологии ОКТ позволяет неинвазивно получать поперечные срезы сетчатки. При этом реконструированные ОК-томограммы содержат характерные искажения: спекл-шум, глубинный спад сигнала, а также ограничения осевого и поперечного разрешения [2].

В анализе ОК-томограмм глубокие нейронные сети применяются для классификации патологий, сегментации слоев сетчатки, подавления шума, повышения разрешения и улучшения качества изображений [3, 4, 5]. Большинство классификационных моделей обучается преимущественно по данным и не учитывает явно механизм формирования томограмм, что снижает интерпретируемость модели и может ухудшать устойчивость при изменении качества входных данных.

Физически-информированные нейронные сети включают априорные физические ограничения в архитектуру, функцию потерь или процедуру обучения [6]. Для ОКТ наиболее строгая форма такого информирования возможна при наличии исходного спектрального или интерференционного сигнала. В исследовании [2] использовались исходные интерференционные полосы и спектрограммы для повышения разрешения и снижения спекл-шума. В исследовании [7] решались задачи повышения осевого разрешения и снижения шума с использованием априорных сведений и эквивариантности. В исследовании [4] предложена модельно-информированная схема SDNet для одновременного подавления шума и повышения разрешения низко-сигнальных данных ОКТ. В исследовании [8] для снижения спекл-шума применялись состязательные сети.

В отличие от подходов, ориентированных на восстановление изображения, в данной работе рассматривается классификация реконструированных ОК-томограмм в условиях, когда исходный интерференционный сигнал недоступен. В связи с этим не выполняется полная физическая реконструкция процесса формирования ОКТ-изображения. Вместо этого в обучение классификатора вводится модельно-информированная функциональная регуляризация, основанная на согласовании скрытого представления с эффективной моделью повторной деградации.

В предлагаемой схеме скрытое представление согласуется с моделью деградации таким образом, чтобы после повторной деградации оно приближало наблюдаемую реконструированную томограмму. Данный шаг вводится как дополнительное слагаемое функции потерь и направлено на повышение устойчивости классификатора к физически релевантным искажениям качества изображения, в первую очередь к спекл-шуму.

## II. МАТЕРИАЛЫ И МЕТОДЫ

### A. Постановка задачи

Пусть задана выборка, состоящая из реконструированных ОК-томограмм:

$$\mathcal{D} = \{(y_i, c_i)\}_{i=1}^N, \quad y_i \in [0,1]^{H \times W}, \quad c_i \in \mathcal{Y}.$$

Рассматривается задача классификации на четыре класса, в которой каждому В-скану сопоставляется ровно одна метка из множества:

$$\mathcal{Y} = \{\text{NO, AMD, ERM, RVO}\}, \quad C = 4.$$

Нейросетевой классификатор (ИНС) формирует распределение вероятностей по классам:

$$f_{\theta}(y_i) = p_{\theta}(y_i) \in \Delta^{C-1}, \quad p_{\theta,c}(y_i) = \frac{\exp(s_{\theta,c}(y_i))}{\sum_{k=1}^C \exp(s_{\theta,k}(y_i))}.$$

Обучение задается как минимизация эмпирического риска:

$$\hat{\theta} = \arg \min_{\theta \in \mathcal{P}} \left\{ \frac{1}{N} \sum_{i=1}^N \ell_w(f_{\theta}(y_i), c_i) + \Omega(\theta) \right\}.$$

Для учета дисбаланса классов классификационное слагаемое задается взвешенной перекрестной энтропией:

$$\ell_w(f_{\theta}(y_i), c_i) = -w_{c_i} \log p_{\theta,c_i}(y_i),$$

где  $w_c$  – вес класса  $c$ , а  $\Omega(\theta)$  – регуляризующее слагаемое. В базовой модели ИНС  $\Omega(\theta)$  отсутствует или сводится к стандартной регуляризации параметров. В модельно-информированной ИНС критерий оптимизации дополнительно включает функциональную регуляризацию, определяемую далее.

### В. Модельно-информирование ИНС

Используемый способ модельного информирования отличается от классической постановки физически-информированных нейронных сетей, в которых физическая информация вводится через невязку дифференциального уравнения [6]. В настоящей работе такая невязка не используется, поскольку решается задача классификации реконструированных ОК-томограмм, а исходный интерференционный сигнал недоступен. При этом подход согласуется с более широкой трактовкой физически-информированного машинного обучения, где данные совмещаются с математической моделью через дифференцируемое слагаемое функции потерь [9]. В данной работе таким слагаемым является модельно-информированная функциональная регуляризация, основанная на операторе эффективной повторной деградации, реконструированной ОК-томограммы. Поэтому априорная информация о формировании изображения используется не как полная физическая реконструкция измерительного процесса и не как жесткое архитектурное ограничение, а как слагаемое самосогласованности скрытого представления с эффективной Фурье-моделью деградации томограммы

В архитектуре используется вспомогательный модуль  $R_{\psi}: y \mapsto \hat{x}$ , который по входной ОК-томограмме  $y$  формирует скрытое представление:  $\hat{x} = R_{\psi}(y)$ . Далее к этому представлению применяется дифференцируемый оператор повторной деградации  $P_{\alpha}$  моделирующий основные искажения реконструированного изображения:  $P_{\alpha}: \hat{x} \mapsto \tilde{y}$ , где  $\alpha$  – обучаемые параметры эффективной модели деградации.

На вход классификатора подаются исходное изображение, скрытое представление, результат повторной деградации и остаточное изображение:

$$z = \text{concat}(y, \hat{x}, \tilde{y}, y - \tilde{y}).$$

Итоговое распределение вероятностей по классам определяется выражением:

$$p_{\theta}(c | y) = \text{softmax}(F_{\phi}(z)),$$

где  $F_{\phi}$  – сверточная нейронная сеть, а полный набор обучаемых параметров модели имеет вид:  $\theta = (\psi, \alpha, \phi)$ .

### С. Фурье-модель повторной деградации

Оператор  $P_{\alpha}$  задается как композиция трех дифференцируемых преобразований, отражающих основные факторы деградации, реконструированной ОК-томограммы: осевое размытие, глубокий спад чувствительности и поперечное размытие:

$$P_{\alpha}(\hat{x}) = B_x^{\alpha}(G^{\alpha}(B_z^{\alpha}(\hat{x}))).$$

Осевой компонент  $B_z^{\alpha}$  реализуется через одномерное преобразование Фурье по глубинной координате:

$$B_z^{\alpha}(\hat{x}) = \mathcal{F}_z^{-1}[H_{\alpha}(k_z) \odot \mathcal{F}_z(\hat{x})],$$

где  $\mathcal{F}_z$  – дискретное преобразование Фурье по оси глубины,  $H_{\alpha}(k_z)$  – обучаемая неотрицательная передаточная функция. Такая формализация согласуется с тем, что глубинная структура получаемых изображений связана с Фурье-преобразованием интерференционного сигнала [1, 2].

Глубинный спад чувствительности моделируется экспоненциальной огибающей по нормированной глубинной координате  $z \in [0,1]$ :

$$G^{\alpha}(u)(z, x) = \exp(-\beta z)u(z, x), \quad \beta > 0.$$

Поперечное размытие задается глубинно-зависимой сверткой по координате  $x$ :

$$B_x^{\alpha}(u)(z, \cdot) = K_{\sigma(z)} *_{x} u(z, \cdot),$$

где  $K_{\sigma(z)}$  – нормированное гауссово ядро, а  $\sigma(z) > 0$  – обучаемая положительная функция ширины размытия по глубине.

### Д. Информированная функция потерь

Комплексная функция потерь задается как взвешенная сумма классификационного слагаемого, модельно-информированного слагаемого самосогласованности и слагаемого полной вариации:

$$\mathcal{L} = \mathcal{L}_{cls} + \lambda_{phys}(t)\mathcal{L}_{phys} + \lambda_{tv}\mathcal{L}_{tv}.$$

При этом, модельно-информированное слагаемое самосогласованности определяется выражением:

$$\mathcal{L}_{phys} = \frac{1}{B} \sum_{i=1}^B \|P_{\alpha}(R_{\psi}(y_i)) - y_i\|_1.$$

Таким образом, к критерию оптимизации добавляется слагаемое самосогласованности, минимизация которого способствует формированию скрытых представлений, согласованных с наблюдаемой реконструированной ОК-томограммой после применения оператора повторной деградации.

Дополнительно используется слагаемое полной вариации, далее обозначаемое как  $\mathcal{L}_{tv}$ , зависящее от суммы модулей разностей между соседними элементами скрытого представления по глубинной и поперечной координатам:

$$\mathcal{L}_{tv} = \frac{1}{B} \sum_{i=1}^B \text{TV}(\hat{x}_i),$$

где

$$TV(\hat{x}) = \sum_{z,x} |\hat{x}_{z+1,x} - \hat{x}_{z,x}| + |\hat{x}_{z,x+1} - \hat{x}_{z,x}|.$$

При положительном весовом коэффициенте данного слагаемого его включение данного слагаемого в комплексную функцию потерь способствует уменьшению избыточной пространственной изменчивости скрытого представления.

Для стабилизации обучения применяется поэтапное включение физического штрафа:

$$\lambda_{phys}(t) = \lambda_{max} \min\left(1, \frac{t}{T_w}\right),$$

где  $t$  – номер эпохи,  $T_w$  – число эпох включения. Такая схема позволяет на начальных эпохах преимущественно формировать признаки, значимые для классификации, а затем постепенно увеличивать вклад модельно-информированного слагаемого.

#### Е. Набор данных и протокол оценки

Для обучения и валидации ИНС использовались совпадающие классы биомаркеров из наборов данных OCTDL [3] и MultiEYE [10]: NO (норма), AMD (возрастная макулярная дегенерация), ERM (эпиретинальная мембрана) и RVO (окклюзия вен сетчатки). Сведения о характеристиках, использованных подмножеств наборов данных, приведены в табл. I.

ТАБЛИЦА I. ХАРАКТЕРИСТИКА ИСПОЛЬЗОВАННЫХ НАБОРОВ ДАННЫХ ОК-ТОМОГРАММ

Класс	Наборы данных	
	OCTDL [3]	MultiEYE [11]
NO	332	27 854
AMD	1 231	961
ERM	155	1 585
RVO	101	1 053
Всего	1 819	31 453
Использованные модели томографов	Optovue Avanti RTVue XR	Topcon DRI Triton Swept Source OCT, Velite C3000 Spectral Domain OCT

Качество классификации оценивается по показателям, представленным в табл. 2 и на рис. 1: F1-мере и полноте по каждому классу, макроусредненной F1-мере и сбалансированной точности. Вводятся точность класса  $P_c$ , полнота класса  $R_c$  и F1-мера:

$$P_c = \frac{TP_c}{TP_c + FP_c}, \quad R_c = \frac{TP_c}{TP_c + FN_c}, \quad F1_c = \frac{2P_c R_c}{P_c + R_c},$$

$$MacroF1 = \frac{1}{C} \sum_{c=1}^C F1_c, \quad BalancedAcc = \frac{1}{C} \sum_{c=1}^C R_c.$$

Основной метрикой является Macro-F1, поскольку она усредняет значения F1-меры по классам без учета их частоты в выборке. Сбалансированная точность в рассматриваемой многоклассовой задаче соответствует средней полноте по классам. Нормированная матрица ошибок на рис. 1 строится по строкам, соответствующим истинным классам:

$$Q_{ij} = \frac{n_{ij}}{\sum_{k=1}^C n_{ik}},$$

где  $n_{ij}$  – число объектов истинного класса  $i$ , отнесенных классификатором к классу  $j$ , а  $C$  – число классов. Для проверки устойчивости вводится семейство деградаций:

$$\mathcal{T}_{d,\rho}: y \mapsto y^{(d,\rho)},$$

где  $d$  – тип деградации, а  $\rho$  – ее уровень. Рассматриваются спекл-шум, гауссов шум, гауссово размытие и снижение контраста. На рис. 2 анализируется зависимость Macro-F1 от уровня деградации, а также снижение этой метрики относительно валидационной выборки без искусственно внесенных деградаций:

$$\Delta M(d,\rho) = M(\mathcal{D}_{val}) - M(\mathcal{T}_{d,\rho}(\mathcal{D}_{val})), \quad M = Macro - F1.$$

Для результатов, приведенных в табл. 2, используются серии независимых запусков. Пусть  $M_r$  – значение метрики  $M$  в  $r$ -м повторе,  $r = 1, \dots, R$ , где  $R \geq 2$ . В табл. 2 приводятся выборочное среднее и выборочное стандартное отклонение (mean  $\pm$  std):

$$\bar{M} = \frac{1}{R} \sum_{r=1}^R M_r, \quad s_M = \sqrt{\frac{1}{R-1} \sum_{r=1}^R (M_r - \bar{M})^2}.$$

### III. РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

По данным табл. 2, модельно-информированная ИНС имеет более высокие средние значения F1-меры и полноты по сравнению с базовой ИНС на всех четырех классах. Для классов NO и AMD значения F1-меры увеличивается с  $0,84 \pm 0,04$  до  $0,92 \pm 0,02$  и с  $0,86 \pm 0,03$  до  $0,93 \pm 0,02$  соответственно. Малочисленные классы ERM и RVO также имеют тенденцию к росту значений F1-меры: с  $0,64 \pm 0,08$  до  $0,81 \pm 0,06$  и с  $0,80 \pm 0,09$  до  $0,89 \pm 0,06$  соответственно. Macro-F1 увеличивается с  $0,785 \pm 0,060$  до  $0,888 \pm 0,040$ , а средняя полнота по классам – с  $0,825 \pm 0,058$  до  $0,908 \pm 0,035$ . Таким образом, улучшение наблюдается не только для классов с большим числом изображений, но и для менее представленных патологических категорий ERM и RVO.

ТАБЛИЦА II. ОЦЕНКА КАЧЕСТВА КЛАССИФИКАЦИИ НА ВАЛИДАЦИОННОЙ ВЫБОРКЕ OCTDL: СРЕДНЕЕ  $\pm$  СТАНДАРТНОЕ ОТКЛОНЕНИЕ

Класс	Базовая ИНС, F1	Мод.-инф. ИНС, F1	Базовая ИНС, Recall	Мод.-инф. ИНС, Recall
NO	$0.84 \pm 0.04$	$0.92 \pm 0.02$	$0.85 \pm 0.03$	$0.93 \pm 0.02$
AMD	$0.86 \pm 0.03$	$0.93 \pm 0.02$	$0.86 \pm 0.03$	$0.94 \pm 0.02$
ERM	$0.64 \pm 0.08$	$0.81 \pm 0.06$	$0.75 \pm 0.07$	$0.85 \pm 0.06$
RVO	$0.80 \pm 0.09$	$0.89 \pm 0.06$	$0.84 \pm 0.10$	$0.91 \pm 0.06$
Macro	$0.785 \pm 0.060$	$0.888 \pm 0.040$	$0.825 \pm 0.058$	$0.908 \pm 0.035$

В табл. 2 показано увеличение средних значений Macro-F1 и сбалансированной точности у модельно-информированной ИНС относительно базовой модели. По рис. 2 видно, что в выбранном протоколе деградаций снижение Макро-F1 при спекл-шуме и гауссовом шуме на OCTDL менее выражено у модельно-информированной ИНС, чем у базовой ИНС. Дополнительно результаты получают внешнее подтверждение при валидации на наборе данных MultiEYE. Так на рис. 2 показано, что модельно-информированная ИНС сохраняет сопоставимый характер изменения Macro-F1 при рассмотренных деградациях и на этой выборке.

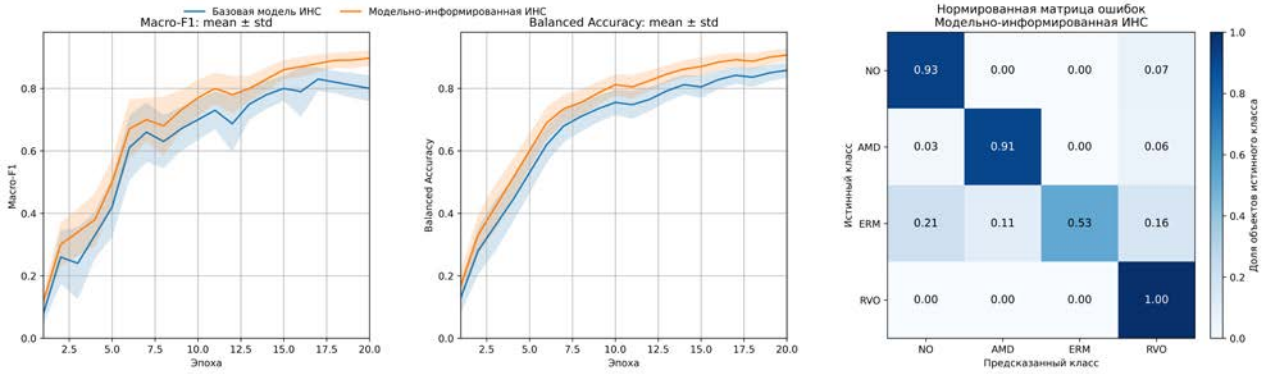


Рис. 1. Значения метрик Macro-F1, Balanced Accuracy и нормированная по строкам матрица ошибок модельно-информированной ИНС для одного представленного запуска на валидационной выборке набора данных OCTDL [3].

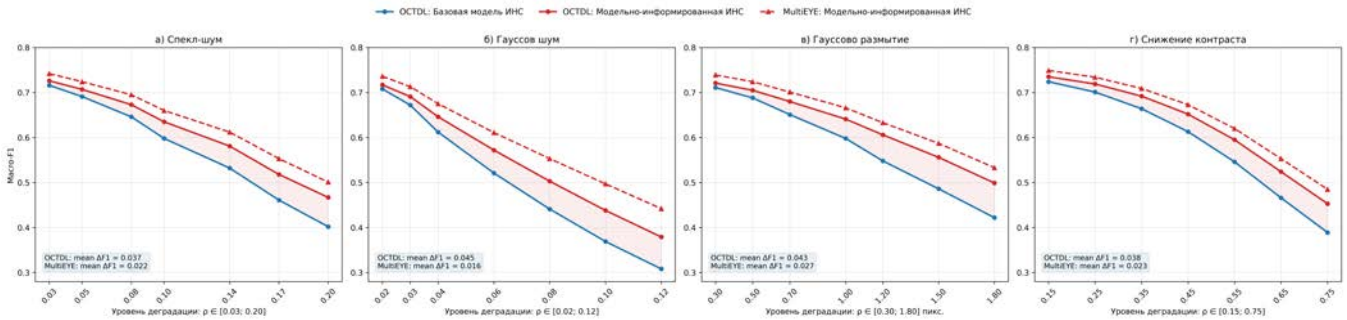


Рис. 2. Зависимость значений метрики Macro-F1 от уровня деградации ОК-томограмм для базовой и модельно-информированной ИНС на OCTDL [3] и для модельно-информированной ИНС на MultiEYE [10]. Деградации: а) спекл-шум; б) гауссов шум; в) гауссово размытие; г) снижение контраста.

#### IV. ЗАКЛЮЧЕНИЕ

Полученные результаты показывают, что модельно-информированная регуляризация повышает качество классификации ОК-томограмм по Macro-F1 и средней полноте. Рост этих метрик указывает на улучшение не только для классов с большим числом изображений, но и для менее представленных патологических категорий в использованных подмножествах данных. При нескольких типах шумовых деградаций предложенная схема медленнее теряет качество на OCTDL, а результаты на MultiEYE дополнительно показывают сохранение сопоставимого характера изменения Macro-F1 при рассмотренных деградациях, что подтверждает целесообразность внедрения модельно-информированного слагаемого самосоогласованности при отсутствии исходного интерференционного сигнала.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Huang D., Swanson E. A., Lin C. P. et al. Optical coherence tomography // *Science*. 1991. Vol. 254, no. 5035. P. 1178-1181. DOI: 10.1126/science.1957169.
- [2] Lee W., Nam H. S. Deep learning-based image enhancement in optical coherence tomography by exploiting interference fringe // *Communications Biology*. 2023. Vol. 6. Article 464. DOI: 10.1038/s42003-023-04846-7.
- [3] Kulyabin M., Zhdanov A., Nikiforova A. et al. OCTDL: Optical Coherence Tomography Dataset for Image-Based Deep Learning Methods // *Scientific Data*. 2024. Vol. 11. Article 365. DOI: 10.1038/s41597-024-03182-7.
- [4] Rakhshani S., Arbab A., Habibi A. et al. Self-supervised model-informed deep learning for low-SNR SS-OCT domain transformation // *Scientific Reports*. 2025. Vol. 15. Article 17791. DOI: 10.1038/s41598-025-02375-3.
- [5] Ge C., Yu X., Yuan M. et al. Self-supervised Self2Self denoising strategy for OCT speckle reduction with a single noisy image // *Biomedical Optics Express*. 2024. Vol. 15, no. 2. P. 1233-1252. DOI: 10.1364/BOE.515520.
- [6] Raissi M., Perdikaris P., Karniadakis G. E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations // *Journal of Computational Physics*. 2019. Vol. 378. P. 686-707. DOI: 10.1016/j.jcp.2018.10.045.
- [7] Li K., Yang J., Liang W. et al. O-PRESS: Boosting OCT axial resolution with Prior guidance, Recurrence, and Equivariant Self-Supervision // *Medical Image Analysis*. 2025. Vol. 99. Article 103319. DOI: 10.1016/j.media.2024.103319.
- [8] Ma Y., Chen X., Zhu W. et al. Speckle noise reduction in optical coherence tomography images based on edge-sensitive cGAN // *Biomedical Optics Express*. 2018. Vol. 9, no. 11. P. 5129-5146. DOI: 10.1364/BOE.9.005129.
- [9] Karniadakis G. E., Kevrekidis I. G., Lu L., Perdikaris P., Wang S., Yang L. Physics-informed machine learning // *Nature Reviews Physics*. 2021. Vol. 3, no. 6. P. 422-440. DOI: 10.1038/s42254-021-00314-5.
- [10] Wang L., Qi C., Ou C., An L., Jin M., Kong X., Li X. MultiEYE: Dataset and Benchmark for OCT-Enhanced Retinal Disease Recognition from Fundus Images // *IEEE Transactions on Medical Imaging*. 2025. Vol. 44, no. 4. P. 1711-1722. DOI: 10.1109/TMI.2024.3518067.