ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р 71673— 2024

Системы искусственного интеллекта в лучевой диагностике

АЛГОРИТМЫ АНАЛИЗА МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ

Методы испытаний на определение точности измерений

Издание официальное

Москва
Российский институт стандартизации
2024

gost.e-ecolog.ru

Предисловие

1 РАЗРАБОТАН Государственным бюджетным учреждением здравоохранения города Москвы «Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы» (ГБУЗ «НПКЦ ДиТ ДЗМ»)

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 164 «Искусственный интеллект»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 7 октября 2024 г. № 1386-ст

4 ВВЕДЕН ВПЕРВЫЕ

*Правила применения настоящего стандарта установлены в статье 26 Федерального закона от 29 июня 2015 г. № 162-ФЗ «О стандартизации в Российской Федерации». Информация об изменениях к настоящему стандарту публикуется в ежегодном (по состоянию на 1 января текущего года) информационном указателе «Национальные стандарты», а официальный текст изменений и поправок — в ежемесячном информационном указателе «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ближайшем выпуске ежемесячного информационного указателя «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет (*gost.e-ecolog.ru

© Оформление. ФГБУ «Институт стандартизации», 2024

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

Содержание

1 Область применения 1
2 Нормативные ссылки 1
3 Термины и определения 2
4 Общие положения 2
5 Методы оценки точности измерений 3
Библиография 7

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

Системы искусственного интеллекта в лучевой диагностике

АЛГОРИТМЫ АНАЛИЗА МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ

Методы испытаний на определение точности измерений

Artificial intelligence in radiology. Algorithms for medical image analysis. Trials to estimate diagnostic accuracy

Дата введения — 2025—01—01

1 Область применения

Настоящий стандарт определяет требования к системам искусственного интеллекта (СИИ) в лучевой диагностике.

Настоящий стандарт представляет описание методов испытаний на определение точности измерений СИИ при использовании алгоритмов анализа медицинских изображений.

2 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты:

ГОСТ 34100.3/ISO/IEC Guide 98-3:2008 Неопределенность измерения. Часть 3. Руководство по выражению неопределенности измерения

ГОСТ ISO/IEC 17025 Общие требования к компетентности испытательных и калибровочных лабораторий

ГОСТ Р 8.736—2011 Государственная система обеспечения единства измерений. Измерения прямые многократные. Методы обработки результатов измерений. Основные положения

ГОСТ Р 59898—2021 Оценка качества систем искусственного интеллекта. Общие положения

ГОСТ Р 59921.0 Системы искусственного интеллекта в клинической медицине. Основные положения

ГОСТ Р 59921.5 Системы искусственного интеллекта в клинической медицине. Часть 5. Требования к структуре и порядку применения набора данных для обучения и тестирования алгоритмов

ГОСТ Р 59921.7—2022 Системы искусственного интеллекта в клинической медицине. Алгоритмы анализа медицинских изображений. Методы испытаний. Общие требования

ГОСТ Р ИСО 5725-1—2002 Точность (правильность и прецизионность) методов и результатов измерений. Часть 1. Основные положения и определения

Примечание — При пользовании настоящим стандартом целесообразно проверить действие ссылочных стандартов в информационной системе общего пользования на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет или по ежегодному информационному указателю «Национальные стандарты», который опубликован по состоянию на 1 января текущего года, и по выпускам ежемесячного информационного указателя «Национальные стандарты» за текущий год. Если заменен ссылочный стандарт, на который дана недатированная ссылка, то рекомендуется использовать действующую версию этого стандарта с учетом всех внесенных в данную версию изменений. Если заменен ссылочный стандарт, на который дана датированная ссылка, то рекомендуется использовать версию этого стандарта с указанным выше годом утверждения (принятия). Если после утверждения настоящего стандарта в ссылочный стандарт, на который дана датированная ссылка, внесено изменение, затрагивающее положение, на которое дана ссылка, то это положение рекомендуется применять без учета данного изменения. Если ссылочный стандарт отменен без замены, то положение, в котором дана ссылка на него, рекомендуется применять в части, не затрагивающей эту ссылку.

Издание официальное

3 Термины и определения

В настоящем стандарте применены термины и определения по ГОСТ Р 59921.0, а также следующие термины с соответствующими определениями:

3.1

медицинское изображение: Информация, получаемая с использованием средств визуализации внутренних структур и функций человеческого тела, представленная в виде изображения, удобного для медицинской диагностики.

Примечания

1 Медицинское изображение может быть получено радиологическими или нерадиологическими методами.

2 Радиологические методы получения медицинского изображения основаны на использовании электромагнитного поля (например, методы рентгеновской диагностики, компьютерной томографии, магнитно-резонансной томографии и т. д.).

3 Нерадиологическими методами получают медицинские изображения, которые отсняты видеокамерой (эндоскопия), или сфотографированные (микроскопические изображения в гистологии, патологии, дерматологические изображения и т. п.).

[ГОСТ Р 59921.7—2022, пункт 3.7]

3.2 **модальность медицинского оборудования для лучевой диагностики:** Тип метода лучевой диагностики, предназначенный для получения медицинского изображения с использованием различных способов (например, компьютерной томографии, магнитно-резонансной томографии, рентгенографии, маммографии).

3.3

принятое опорное значение: Значение, которое служит в качестве согласованного для сравнения и получено как:

a) теоретическое или установленное значение, базирующееся на научных принципах;

b) приписанное или аттестованное значение, базирующееся на экспериментальных работах какой-либо национальной или международной организации;

c) согласованное или аттестованное значение, базирующееся на совместных экспериментальных работах под руководством научной или инженерной группы;

d) математическое ожидание измеряемой характеристики, то есть среднее значение заданной совокупности результатов измерений — лишь в случае, когда a), b) и c) недоступны.

[ГОСТ Р ИСО 5725-1—2002, пункт 3.5]

3.4

точность: Степень близости результата измерений к принятому опорному значению.

Примечание — Термин «точность», когда он относится к серии результатов измерений (испытаний), включает сочетание случайных составляющих и общей систематической погрешности.

[ГОСТ Р ИСО 5725-1—2002, пункт 3.6]

4 Общие положения

4.1 Медицинские изображения используют для измерения характеристик, необходимых для диагностики различных заболеваний.

К таким характеристикам по [1] относят:

- линейные размеры анатомических структур в норме и при патологических состояниях (объемные образования, сосуды, анатомические структуры сердца, легких и т. д.);
- площадь и объем;
- интенсивность сигнала в области интереса, выраженная в различных единицах (например, единицы Хаунсфилда для модальностей, рентгенография или компьютерная томография);
- расстояния и углы (на медицинских изображениях измеряют расстояния между точками или углы между структурами);
- результаты динамических измерений. Данные измерения выполняют по данным серии медицинских изображений, чтобы изучить зависимость показателей от времени (например, динамическое

накопление контрастного вещества, измерение кровотока, оценка движения органов, изменений в размере или форме структур и т. д.);

- функциональные характеристики. Медицинские изображения используют для оценки функции органов, например: изображения, полученные в процессе позитронно-эмиссионной томографии, применяют для измерения метаболической активности тканей;

- результаты текстурного анализа. На основании медицинских изображений выполняют измерения, такие как текстурный анализ, в ходе которого исследуют микроструктуру тканей на изображении по [2].

4.2 Данные измерения выполняют с использованием программного обеспечения для просмотра и анализа медицинских изображений или автоматизированно с помощью СИИ в лучевой диагностике по [3], [4].

4.3 Точность измерений должна быть определена с учетом погрешности применяемого оборудования в соответствии с ГОСТ Р ИСО 5725-1, а также ГОСТ Р 8.736.

Примечание — В настоящем стандарте не рассматривают СИИ в качестве средств измерения.

4.4 Порядок проведения измерений для медицинских изображений в процессе работы СИИ должен быть указан в технической и эксплуатационной документации производителя. Правила выполнения измерений при разработке СИИ должны быть основаны на клинических рекомендациях, методических рекомендациях и принятых в клинической практике нормах. Необходимо указать точность и погрешность между результатами работы СИИ и экспертной разметкой медицинских изображений.

Примечание — Формирование набора данных с учетом его размера и состава определяют на основании ГОСТ Р 59921.5. Примеры подготовки наборов данных для тестирования СИИ в лучевой диагностике приведены в [5], [6].

4.5 Испытания в целях оценки точности измерений, выполняемых СИИ в лучевой диагностике, рекомендовано проводить в рамках функционального тестирования при оценке уровня точности в соответствии с ГОСТ Р 59921.7—2022 (раздел 7).

5 Методы оценки точности измерений

5.1 Общее описание методов

В настоящем стандарте выделены несколько методов оценки точности измерений: метод применения верифицированного набора данных (5.4.1), метод математического (5.4.2) и физического (5.4.3) моделирования. Метод применения верифицированного набора данных необходимо рассматривать в качестве основного метода.

5.2 Требования к квалификации персонала

Испытания проводят квалифицированные технические специалисты, имеющие опыт тестирования СИИ, обработки медицинских изображений и анализа данных результатов. В подготовке набора данных для испытаний задействованы медицинские специалисты при наличии у них опыта работы и надлежащей квалификации в областях, соответствующих решаемым задачам СИИ.

5.3 Требования к оборудованию

Испытание проводят с применением испытательного стенда, которым может быть как отдельное автоматизированное рабочее место, так и виртуальная тестовая среда, удовлетворяющая техническим требованиям СИИ (см. [7], статья 10).

Кроме того, возможно использование виртуальной тестовой среды, позволяющей выполнять автоматические тестирования (см. [7], статья 8.1).

Перед выполнением тестирования необходимо проверить отсутствие существенных различий между средой проведения тестирования и средой эксплуатации СИИ в соответствии с ГОСТ Р 59898—2021 (подраздел 7.2).

Оборудование, необходимое для выполнения разметки и верификации набора данных, определяют исходя из указанного по 5.4 метода испытаний.

Оборудование, применяемое в ходе выполнения испытания, должно иметь установленные показатели точности измерений.

5.4 Описание методов оценки точности измерений

5.4.1 Метод верифицированного набора данных

Данный метод подразумевает оценку точности измерений СИИ путем анализа данных, полученных на верифицированном наборе данных.

Следует определить задачу, на решение которой направлена испытуемая СИИ, и выделить выполняемые измерения, которые влияют на выполнение задачи СИИ.

Исходя из выполняемых измерений необходимо установить требования к верифицируемому набору данных согласно ГОСТ Р 59921.5, в том числе к привлекаемым для проведения разметки и верификации медицинским специалистам. Набор данных должен включать медицинские изображения, удовлетворяющие требованиям технической документации СИИ, а также все классы медицинских изображений, которые определены задачей СИИ (например, классы с признаками нормы и патологии). Объем набора данных должен быть статистически обоснованным и достаточным для обеспечения необходимой точности измерений.

Разметку и верификацию проводят с использованием специализированного программного обеспечения, предназначенного для просмотра и анализа медицинских изображений. Данное программное обеспечение должно иметь установленные показатели точности измерений. Разметку также можно выполнить путем измерения посредством доказанного более точного метода лучевой диагностики.

Для каждого вида измерений используют установленный метод, приведенный в таблице 1.

Таблица 1 — Методы измерений в зависимости от вида

Вид измерений	Метод измерений
Линейный размер	Определение границы структур с использованием метода определения половины максимального значения изменения интенсивности; измерение расстояния между границами структур
Диаметр	Определение границы структур с использованием метода определения половины максимального значения изменения интенсивности; определение центра структуры; измерение двух расстояний от центра до границ в противоположных направлениях
Угол	Установление линий; измерение угла
Площадь	Выделение области интереса внутри структуры; определение площади путем интегрирования
Объем	Выделение области изучения, определение объема выявленной области изучения
Интенсивность сигнала	Выделение области анализа внутри структуры; определение интенсивности сигнала путем измерения среднего значения интенсивности пикселей в области анализа

На вход СИИ подают верифицированный набор данных и фиксируют выходные данные СИИ, включая результаты измерений, на основании которых СИИ решает поставленную задачу. Результаты измерений предоставляют в виде оценки измеряемой величины и среднего квадратического отклонения в соответствии с ГОСТ Р 8.736—2011 (раздел 5).

Показатели точности для каждого элемента верифицированного набора данных рассчитывают путем сопоставления результатов измерений с помощью СИИ и результатов измерений, выполненных экспертами при разметке и верификации набора данных.

Показатели точности включают общую систематическую погрешность, а также случайную составляющую погрешности согласно требованиям ГОСТ Р ИСО 5725-1. Систематическую погрешность Δ рассчитывают в соответствии с ГОСТ Р ИСО 5725-1—2002 (подраздел 6.3) в виде абсолютной разности и неопределенности оценки с учетом доверительной вероятности 95 %.

Для оценки результата испытания с использованием верифицированного набора данных необходимо определить максимальное и среднее отклонения систематической погрешности всех элементов набора данных от среднего значения систематической погрешности.

Максимальное отклонение $\Delta_{макс}$ рассчитывают по данным значений модуля отклонений каждого i-го элемента набора данных Δ_i по формуле

$$\Delta_{макс} = макс\{|\Delta_i|\}. \quad (1)$$

Среднее отклонение $\Delta_{ср}$ рассчитывают по данным значений отклонений каждого элемента *i*-го набора данных Δ_i, состоящего из элементов *N*, по формуле

$$\Delta_{ср} = \frac{\Sigma\{\Delta_i\}}{N}. \quad (2)$$

Показатели неопределенности выполненных измерений рассчитывают в соответствии с требованиями ГОСТ 34100.3:

- оценку неопределенности типа А $u_А$ проводят в ходе расчета среднего квадратического отклонения случайной погрешности оценки результата измерений; неопределенность $u_А$ соответствует стандартному отклонению среднего квадратического отклонения;

- оценку неопределенности типа Б $u_Б$ осуществляют с учетом вклада всех показателей используемого испытательного оборудования (тест-объектов, программного обеспечения);

- стандартную неопределенность $u_с$ рассчитывают по формуле

$$u_с = \sqrt{u_А^2 + u_Б^2}; \quad (3)$$

- расширенную неопределенность *U* определяют путем умножения на $u_с$ коэффициента охвата *k*, равного 2, при уровне доверительной вероятности 0,95.

5.4.2 Метод математического моделирования

Метод основан на формировании набора данных с использованием математического моделирования (численные фантомы).

Необходимо определить задачу, на решение которой направлена испытуемая СИИ, и выделить те выполняемые измерения, которые влияют на решение задачи СИИ.

Необходимо сгенерировать набор данных с использованием математических и программных средств. Набор данных должен соответствовать требованиям, предъявляемым к медицинским изображениям, которые установлены в технической документации СИИ, чтобы обеспечить правильность определения целевой анатомической структуры. При выполнении моделирования должна быть обеспечена возможность установления точных характеристик целевой анатомической структуры (например, линейного размера, диаметра, значений интенсивности).

Далее испытание выполняют по аналогии с методом по 5.4.1 путем подачи набора данных на вход СИИ и анализа отклонений измеренных значений от установленных, а также расчета расширенной неопределенности.

Применение данного метода может быть сложным при его реализации в связи с тем, что СИИ разрабатывают с использованием наборов данных реальной клинической практики. Это может привести к появлению ошибок функционирования СИИ при подаче на вход наборов данных, сформированных путем математического моделирования.

5.4.3 Метод физического моделирования

Данный метод использует физическое моделирование для оценки точности измерений. Физическое моделирование включает применение стандартных тест-объектов (фантомов), например: фантомы, моделирующие органы грудной клетки и включающие модели различных анатомических структур (см. [8]).

Перед использованием фантома необходимо определить показатели точности установки параметров анатомических структур.

Далее испытание выполняют по аналогии с испытанием по методу, приведенному в 5.4.1, путем подачи набора данных на вход СИИ и анализа отклонений измеренных значений от установленных, а также расчета расширенной неопределенности.

5.5 Предоставление результатов

Результаты испытаний оформляют в виде протокола в соответствии с ГОСТ ISO/IEC 17025. В протоколе следует привести параметры, указанные в таблице 2.

Таблица 2 — Параметры измерения

Параметр	Размерность
Показатели точности испытательного оборудования (тест-объектов, программного обеспечения)	Абсолютные единицы/ относительные единицы
Результаты измерения для всех элементов набора данных	Абсолютные единицы
Значение отклонения	
Максимальное значение систематической погрешности	Абсолютные единицы
Среднее значение систематической погрешности	
Рассчитанное значение расширенной неопределенности	Абсолютные единицы

Примечания

1 Единицы измерения определяют исходя из модальности медицинского изображения и выполняемого измерения.

2 Результаты измерения могут быть представлены в виде следующих единиц измерения: мм/мм^2 /мм^3/°/единицы Хаунсфилда или др.

Помимо сведений, указанных в ГОСТ ISO/IEC 17025, рекомендовано включить в протокол также сведения о серии/срезе медицинского изображения, выбранные для измерения всех элементов набора данных.

Библиография

[1] Морозов С.П. Принципы и правила описаний результатов лучевых исследований: методические рекомендации/ С.П. Морозов, Д.В. Буренчев, А.В. Владзимирский и др.// Серия «Лучшие практики лучевой и инструментальной диагностики». — Вып. 97. — М.: ГБУЗ «НПКЦ ДиТ ДЗМ», 2021. — 36 с.

[2] Хоружая А.Н. Роль системы контроля качества лучевой диагностики онкологических заболеваний в радиомике/ А.Н. Хоружая, Е.С. Ахмад, Д.С. Семенов // Digital Diagnostics. — 2021. — Т. 2. — №2. — С. 170—184

[3] Морозов С.П. Базовые рекомендации к работе сервисов искусственного интеллекта для лучевой диагностики: методические рекомендации/сост. С.П. Морозов, Л.Р. Абуладзе, А.Е. Андрейченко и др. // Серия «Лучшие практики лучевой и инструментальной диагностики». — Вып. 119. — М.: ГБУЗ «НПКЦ ДиТ ДЗМ», 2022. — 68 с.

[4] Компьютерное зрение в лучевой диагностике: первый этап московского эксперимента: монография/ А.В. Владзимирский, Ю.А. Васильев, К.М. Арзамасов и др. — 2-е изд. перераб. и доп. — Москва: ООО «Издательские решения», 2022. — 376 с.

[5] Подготовка набора данных для обучения и тестирования программного обеспечения на основе технологии искусственного интеллекта: учебно-методическое пособие/Ю.А. Васильев, К.М. Арзамасов, А.В. Владзимирский и др. // ООО «Издательские решения». Москва, 2023 — 131 с.

[6] Бобровская Т.М., Васильев Ю.А., Никитин Н.Ю., Арзамасов К.М. Подходы к формированию наборов данных в лучевой диагностике. Врач и информационные технологии. 2023; 4: С. 14—23

[7] ISO/IEC TR 29119-11:2020(E). Software and systems engineering — Software testing — Part 11: Guidelines on the testing of AI-based system // ISO: [сайт]. — URL: gost.e-ecolog.ru

[8] Rodríguez Pérez S. Characterization and validation of the thorax phantom Lungman for dose assessment in chest radiography optimization studies /S. Rodríguez Pérez, N.W. Marshall, L. Struelens, H.J. Bosmans // Med Imaging. — 2018. Vol. 5. — № 1. — 013504

УДК 615.841:006.354 ОКС 11.040.01

Ключевые слова: система искусственного интеллекта, искусственный интеллект, лучевая диагностика, медицинские изображения, измерения, контроль качества

Редактор *Л.С. Зимилова*
Технический редактор *И.Е. Черепкова*
Корректор *Р.А. Ментова*
Компьютерная верстка *И.А. Налейкиной*

Сдано в набор 08.10.2024. Подписано в печать 11.10.2024. Формат 60×84⅛. Гарнитура Ариал.
Усл. печ. л. 1,40. Уч.-изд. л. 1,18.

Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

Создано в единичном исполнении в ФГБУ «Институт стандартизации»
для комплектования Федерального информационного фонда стандартов,
117418 Москва, Нахимовский пр-т, д. 31, к. 2.

gost.e-ecolog.ru gost.e-ecolog.ru

gost.e-ecolog.ru