



Оценка качества работы искусственного интеллекта в выявлении дегенеративных заболеваний пояснично-крестцового отдела позвоночника

Нуднов Н.В.¹, Коробов А.В.², Скачков А.А.³, Кульнева Т.В.², Шерстобоев В.В.⁴, Титова Л.А.⁵, Русаков А.С.⁶, Тумко В.В.⁶, Сарбаев Р.С.⁷, Успенская Н.А.⁶, Андриенко Е.А.², Иванников М.Е.¹

¹ ФГБУ «Российский научный центр рентгенодиагностики» Минздрава России, ул. Профсоюзная, 86, стр. 1, Москва, 117485, Российская Федерация

² АНО ДПО «Институт Эксперт», ул. Фридриха Энгельса, 58А, Воронеж, 394018, Российская Федерация

³ ООО «Объединенное IT пространство», ул. Металлургов, вл. 1, лит. А, Липецк, 398017, Российская Федерация

⁴ Группа компаний «Эксперт», ул. Металлургов, 1, Липецк, 398017, Российская Федерация

⁵ ФГБОУ ВО «Воронежский государственный медицинский университет им. Н.Н. Бурденко» Минздрава России, ул. Студенческая, 10, Воронеж, 394036, Российская Федерация

⁶ Remedy Logic, 6 St Johns Ln, New York, NY 10013, USA

⁷ ООО «Системы поддержки принятий решений», ул. Декабристов, 33, Чебоксары, 428022, Чувашская Республика, Российская Федерация

Нуднов Николай Васильевич, д. м. н., профессор, зам. директора по научной работе, заведующий научно-исследовательским отделом комплексной диагностики и радиотерапии ФГБУ «Российский научный центр рентгенодиагностики» Минздрава России; <http://orcid.org/0000-0001-5994-0468>

Коробов Андрей Владимирович, директор АНО ДПО «Институт Эксперт»

Скачков Артур Андреевич, специалист по машинному обучению ООО «Объединенное IT пространство»; <http://orcid.org/0009-0002-6072-8143>

Кульнева Таисия Владимировна, зам. директора по экспертной работе в области медицинской визуализации АНО ДПО «Институт Эксперт»

Шерстобоев Владислав Васильевич, генеральный директор Группы компаний «Эксперт»

Титова Лилия Александровна, д. м. н., доцент, заведующая кафедрой инструментальной диагностики ФГБОУ ВО «Воронежский государственный медицинский университет им. Н.Н. Бурденко» Минздрава России; <http://orcid.org/0000-0002-8421-3411>

Русаков Андрей Сергеевич, основатель и главный исполнительный директор Remedy Logic

Тумко Владислав Владимирович, главный операционный директор Remedy Logic

Сарбаев Руслан Сергеевич, генеральный директор ООО «Системы поддержки принятий решений»

Успенская Наталья Александровна, руководитель команды компьютерного зрения Remedy Logic

Андриенко Елизавета Андреевна, специалист общего отдела АНО ДПО «Институт Эксперт»

Иванников Михаил Евгеньевич, клинический ординатор по специальности «рентгенология» ФГБУ «Российский научный центр рентгенодиагностики» Минздрава России; <http://orcid.org/0009-0007-0407-0953>

Резюме

Цель: проведение сравнительной оценки выходных данных комплекса обученных моделей сверточных нейронных сетей (convolutional neural network, CNN) и интерпретации патологических изменений поясничного отдела позвоночника врачами-рентгенологами при проведении магнитно-резонансной томографии.

Материал и методы. Собрано более 12 тыс. анонимизированных архивов для формирования обучающего и тестового наборов данных нейросети среди пациентов старше 18 лет. Каждый архив состоял из набора программ в двух плоскостях, содержащих последовательности T2-TSE, T1-TSE и T2 с программой жироподавления. Далее на отобранных исследованиях выполняли разметку в два этапа, заключающуюся непосредственно в ручной разметке и ее проверке специалистами. Обучение CNN проводили отдельно для анализа нормы, качественного определения патологических изменений и количественного анализа. Проверку точности моделей путем сравнительного анализа протоколов пяти врачей-рентгенологов и выходных данных моделей CNN осуществляли в два этапа. На первом, промежуточном этапе оценивали

точность работы нейросетей в выявлении выбуханий, протрузий и экструзий дисков, стеноза позвоночного канала, латеральных стенозов, фораминальных стенозов, спондилолистеза и артроза межпозвонковых суставов. На итоговом этапе помимо патологий, рассматриваемых на промежуточном этапе, проверяли точность выявления дегенеративных изменений замыкательных пластин, синовита межпозвонковых суставов, дегенерации дисков, остеофитов, переходных позвонков, гипертрофии желтых связок и грыжи Шморля. Эталонное значение для всех рассматриваемых в данной работе патологических изменений определено большинством голосов, и в случае разногласий решение принимал внешний рентгенолог. Затем интерпретации рентгенологов были сопоставлены с интерпретацией обученной модели.

Результаты. Искусственный интеллект (ИИ) продемонстрировал сопоставимые значения чувствительности и специфичности в сравнении с эталонным результатом в группе опытных врачей-рентгенологов для бинарной классификации (наличие/отсутствие) наличия отдельных дегенеративных изменений пояснично-крестцового отдела позвоночника. Для экструзий чувствительность и специфичность результатов ИИ составили 0,88 и 0,97 соответственно, для протрузий – 0,81 и 0,94, для центрального стеноза – 0,87 и 0,98, для латерального стеноза – 0,83 и 0,85, для фораминального стеноза – 0,92 и 0,84, для артроза – 0,85 и 0,50, для дегенерации замыкательных пластин – 0,73 и 0,96, для синовита межпозвонковых суставов – 0,85 и 0,84, для дегенерации дисков – 0,91 и 0,88, для остеофитов – 0,93 и 0,72, для переходных позвонков – 1,0 и 1,0, для спондилолистезов – 0,8 и 1,0, для гипертрофии желтых связок – 0,67 и 0,99, для грыжи Шморля – 0,75 и 1,0. Точность количественных характеристик размеров протрузий и экструзий пояснично-крестцового отдела позвоночника показала неудовлетворительные результаты, однако улучшение качества определения данных параметров планируется в последующих работах.

Заключение. Модели ИИ продемонстрировали сопоставимые со специалистами-рентгенологами результаты в обнаружении дегенеративных изменений пояснично-крестцового отдела позвоночника. Последовательное улучшение моделей CNN на основе сравнительной оценки с результатами работы врачей-рентгенологов повышает чувствительность и специфичность выявления патологических изменений.

Ключевые слова: магнитно-резонансная томография, искусственный интеллект, нейронная сеть, сверточная нейронная сеть, глубокое обучение, программное обеспечение в медицине.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Финансирование. Исследование проведено при поддержке АНО ДПО «Институт Эксперт» (Воронеж, Россия).

Для цитирования: Нуднов Н.В., Коробов А.В., Скачков А.А., Кульнева Т.В., Шерстобоев В.В., Титова Л.А., Русаков А.С., Тумко В.В., Сарбаев Р.С., Успенская Н.А., Андриенко Е.А., Иванников М.Е. Оценка качества работы искусственного интеллекта в выявлении дегенеративных заболеваний пояснично-крестцового отдела позвоночника. *Вестник рентгенологии и радиологии*. 2024; 105(1): 20–8. <https://doi.org/10.20862/0042-4676-2024-105-1-20-28>

Для корреспонденции: Иванников Михаил Евгеньевич, E-mail: ivannikovmichail@gmail.com

Статья поступила 05.02.2024

После доработки 13.02.2024

Принята к печати 14.02.2024

Evaluation of Artificial Intelligence Effectiveness in Detection of Lumbosacral Spine Degenerative Diseases

Nikolay B. Nudnov¹, Andrey V. Korobov², Artur A. Skachkov³, Taisia V. Kulneva², Vladislav V. Sherstoboev⁴, Lilia A. Titova⁵, Andrey S. Rusakov⁶, Vladislav V. Tumko⁶, Ruslan S. Sarbaev⁷, Natalia A. Uspenskaya⁶, Elizaveta A. Andrienko², Mikhail E. Ivannikov¹

¹ Russian Scientific Center of Roentgenoradiology, ul. Profsoyuznaya, 86, str. 1, Moscow, 117485, Russian Federation

² Expert Institute, ul. Fridrikha Engelsa, 58A, Voronezh, 394018, Russian Federation

³ United IT Space LLC, ul. Metallurgov, 1, lit. A, Lipetsk, 398017, Russian Federation

⁴ "Expert" Group of Companies, ul. Metallurgov, 1, Lipetsk, 398017, Russian Federation

⁵ Burdenko Voronezh State Medical University, ul. Studencheskaya, 10, Voronezh, 394036, Russian Federation

⁶ Remedy Logic, 6 St Johns Ln, New York, NY 10013, USA

⁷ Decision Support Systems LLC, ul. Dekabristov, 33, Cheboksary, 428022, Chuvash Republic, Russian Federation

Nikolay B. Nudnov, Dr. Med. Sc., Professor, Deputy Director for Scientific Work, Head of Research Department of Complex Diagnostics and Radiotherapy, Russian Scientific Center of Roentgenoradiology;
<http://orcid.org/0000-0001-5994-0468>

Andrey V. Korobov, Director, Expert Institute

Artur A. Skachkov, Machine Learning Specialist, United IT Space LLC;
<http://orcid.org/0009-0002-6072-8143>

Taisia V. Kulneva, Deputy Director for Expert Work in Medical Imaging, Expert Institute

Vladislav V. Sherstoboev, Director General, "Expert" Group of Companies

Lilia A. Titova, Dr. Med. Sc., Associate Professor, Chief of Chair of Instrumental Diagnostics, Burdenko Voronezh State Medical University;
<http://orcid.org/0000-0002-8421-3411>

Andrey S. Rusakov, Founder and Chief Executive Officer, Remedy Logic

Vladislav V. Tumko, Chief Operating Officer, Remedy Logic

Ruslan S. Sarbaev, Director General, Decision Support Systems LLC

Natalia A. Uspenskaya, Head of Computer Vision Team, Remedy Logic

Elizaveta A. Andrienko, Specialist, General Department, Expert Institute

Mikhail E. Ivannikov, Clinical Resident in Radiology, Russian Scientific Center of Roentgenoradiology;
<http://orcid.org/0009-0007-0407-0953>

Abstract

Objective: comparative evaluation of output data of a set of trained convolutional neural network (CNN) models and interpretation of pathological changes in lumbar spine by radiologists during magnetic resonance imaging.

Material and methods. More than 12,000 anonymized archives were collected to generate training and test neural network datasets from patients aged over 18 years. Each archive consisted of a set of programs in two planes containing T2-TSE, T1-TSE and T2 sequences with fat suppression program. Subsequently, the selected studies were tagged in two steps, directly consisting of manual tagging and its validation by experts. CNN training was performed separately for normal analysis, qualitative detection of individual pathological changes, and quantitative analysis. The accuracy of the models was verified by comparing the protocols of five radiologists and the output of CNN models in two steps. The first, intermediate stage evaluated the accuracy of the neural networks in detecting disc bulges, protrusions and extrusions, spinal canal stenosis, lateral stenosis, foraminal stenosis, spondylolisthesis and facet joint arthrosis. In the final stage, in addition to the pathologies considered in the intermediate one, the accuracy of detecting degenerative changes of the occlusive plates, synovitis of intervertebral joints, intervertebral discs degeneration, osteophytes, transitional vertebrae, hypertrophy of yellow ligaments and Schmorl's hernia was tested. The reference value for all pathological changes considered in this paper was determined by majority vote and, in case of disagreement, by an external radiologist. The radiologists' interpretations were then compared with those of the trained model.

Results. The artificial intelligence (AI) showed comparable sensitivity and specificity values compared to the reference result in a group of experienced radiologists for binary classification (presence/absence) of individual lumbosacral spine degenerative changes. The sensitivity and specificity of AI results were 0.88 and 0.97 for extrusions, 0.81 and 0.94 for protrusions, 0.87 and 0.98 for central stenosis, 0.83 and 0.85 for lateral stenosis, 0.92 and 0.84 for foraminal stenosis, 0.85 and 0.5 for osteoarthritis, 0.73 and 0.96 for occlusive plates degeneration, 0.85 and 0.84 for intervertebral joint synovitis, 0.91 and 0.88 for osteophytes, 0.93 and 0.72 for intervertebral disc degeneration, 1.0 and 1.0 for transitional vertebrae, 0.8 and 1.0 for spondylolisthesis, 0.67 and 0.99 for yellow ligament hypertrophy, and 0.75 and 1.0 for Schmorl's hernia, respectively. The accuracy of quantitative size characterization of lumbosacral spine protrusions and extrusions showed unsatisfactory results, but improvements in the quality of determination of these parameters are planned in future work.

Conclusion. AI models showed comparable performance to expert radiologists in detecting lumbosacral spine degenerative changes. Consistent improvement of CNN models based on comparative evaluation with radiologists improves the sensitivity and specificity of pathologic change detection.

Keywords: magnetic resonance imaging; artificial intelligence; neural network; convolutional neural network; deep learning; software in medicine.

Conflict of interest. The authors declare no conflict of interest.

Funding. The study was supported by Expert Institute (Voronezh, Russia).

For citation: Nudnov NB, Korobov AV, Skachkov AA, Kulneva TV, Sherstoboev VV, Titova LA, Rusakov AS, Tumko VV, Sarbaev RS, Uspenskaya NA, Andrienko EA, Ivannikov ME. Evaluation of artificial intelligence effectiveness in detection of lumbosacral spine degenerative diseases. *Journal of Radiology and Nuclear Medicine*. 2024; 105(1): 20–8 (in Russian). <https://doi.org/10.20862/0042-4676-2024-105-1-20-28>

For corresponding: Mikhail E. Ivannikov, E-mail: ivannikovmichail@gmail.com

Received February 5, 2024

Revised February 13, 2024

Accepted February 14, 2024

Введение / Introduction

Согласно данным Всемирной организации здравоохранения основной причиной инвалидизации в мире являются болезни костно-мышечной системы, среди которых преобладают заболевания позвоночника. Люди с дегенеративными болезнями позвоночника наиболее часто нуждаются в современных высокоточных методах диагностики и составляют большую долю потока пациентов в кабинетах магнитно-резонансной томографии (МРТ).

В связи с нарастанием популярности среди населения и повышением доступности МРТ-исследований в последние годы ежедневная нагрузка на врача-рентгенолога сильно возросла, что негативно сказывается на текущем качестве работы и дальнейшей работоспособности специалиста. Также продолжается тренд по развитию алгоритмов нейронных сетей и увеличению доступных вычислительных мощностей, что позволяет реализовывать более сложные модели искусственного интеллекта (ИИ) и обрабатывать большие объемы обучающих данных. Из текущих тенденций возникают благоприятные условия для разработки и внедрения в практическую медицину сервисов на базе ИИ, помогающих врачам в приоритетных направлениях радиологии. В качестве таких систем в настоящее время применяются модели глубокого обучения нейронных сетей, которые представляют собой математическое воплощение системы взаимодействия нейронов человека. Входные данные проходят через слои нейронной сети, где значимость каждого элемента из совокупности данных меняется в зависимости от распределения весов между слоями, а затем в соответствии со степенью расхождения выходных и эталонных значений меняется распределение весов для получения более релевантных результатов в новой эпохе обучения.

Как и множество других специальностей, лучевая диагностика находится на пути к более плотному взаимодействию с системами ИИ для повышения качества и скорости выполняемой работы. Особенностью использования систем нейросетей для диагностики заболеваний позвоночника является возможность упростить оценку МРТ-снимков, что будет способствовать повышению согласованности, точности и объективности оценки исследований, а в более дальней перспективе может полностью освободить врачей-рентгенологов от интерпретации наиболее часто встречающихся патологических изменений в многократно повторяющихся однотипных исследованиях.

При помощи метода МРТ врачи-рентгенологи с достаточно высокой точностью опреде-

ляют производные дегенеративных изменений позвоночника. В то же время задача осложняется существованием нескольких подходов в дифференцировании как отдельных патологических состояний, так и степени их проявлений. Например, определение стеноза позвоночного канала на уровне поясничного отдела позвоночника отличается в зависимости от выбираемой системы оценки, критерии которой могут быть как количественными в виде размера позвоночного канала по срединной линии на уровне диска, так и качественными в виде степени деформации дурального мешка и эпидурального жира (*C. Shizas*). Также примером могут служить отличающиеся степени фораминальных стенозов в зависимости от выбранной специализированной классификации изменений (наиболее часто используются классификации по Lee и по Wildermuth) и измерения пролабирования межпозвонковых дисков в просвет позвоночного канала с учетом толщины задней продольной связки или без него.

Цель – проведение сравнительной оценки выходных данных комплекса обученных моделей сверточных нейронных сетей (convolutional neural network, CNN) и интерпретации патологических изменений поясничного отдела позвоночника врачами-рентгенологами при проведении МРТ.

Материал и методы / Material and methods

Сформирован большой объем анонимизированных архивов для использования в качестве обучающего и тестового наборов данных нейросети после ретроспективного анализа более 12 тыс. МРТ-исследований среди пациентов старше 18 лет на базе диагностических и клинико-диагностических центров: ООО «Клиника Эксперт Воронеж», ООО «Клиника Эксперт Пермь», ООО «Клиника Эксперт Курск», ООО «МРТ-Эксперт Столица», ООО «МРТ-Эксперт Мытищи» (структурное подразделение в г. Орехово-Зуево), ООО «Клиника Эксперт Липецк», ООО «Клиника Эксперт Владикавказ», ООО «Клиника Эксперт Новосибирск», ООО «МРТ-Эксперт БСК», ООО «МРТ-Эксперт-3». Все пациенты были проверены на безопасность прохождения МРТ, подписали согласие на обработку персональных данных и их передачу в ООО «Системы поддержки принятия решений».

Исследования состояли из набора программ в двух плоскостях, содержащих последовательно T2-TSE, T1-TSE, T2 с программой жироподавления (STIR) в сагиттальной плоскости и T2-TSE в аксиальной плоскости. На этапе поиска исключались исследования с выраженными артефактами от металлоконструкции и наличием изображений низкого качества, не позволяющего достаточно точно интерпретировать снимки и постконтрастные



Рис. 1. Примеры автоматической разметки с нумерацией позвонков (a–c)

Fig. 1. Examples of automatic marking with vertebrae numbering (a–c)

серии. Для дополнительного подтверждения целевых патологических изменений сбор анонимизированных архивов МРТ-исследований проводился врачами-рентгенологами. Для исключения повторной загрузки архивов при параллельном сборе разными специалистами фиксировались временные периоды и сервер, с которого скачивалась медицинская информация. Для сохранения отобранных архивов применялось облачное хранилище АНО ДПО «Институт повышения квалификации медицинских кадров» с ограниченным доступом. Возможность использовать облачное хранилище предоставлялась врачам-рентгенологам по VPN-доступу после индивидуальной настройки на компьютерах медицинских учреждений, где работали специалисты. Архивы были отсортированы в хранилище данных по целевым патологическим изменениям.

Процесс разметки анонимизированных изображений для формирования обучающего набора данных проводился в два этапа. На первом этапе разметка выполнялась сотрудниками ООО «Системы поддержки принятий решений», использующими разработанный алгоритм разметки изображений (согласованный с врачами-рентгенологами, имеющими не менее чем 5-летний опыт подготовки описаний исследований, выполненных на магнитно-резонансном томографе) методом выделения границ с применением цветовой схемы, привязанной к списку нормальных анатомических структур и рассматриваемых в работе патологических изменений. Проведена ручная разметка всех анатомических структур и патологических изменений, касающихся поясничного отдела позвоночника на аксиальных срезах, также выполнена разметка на сагиттальных срезах (в отличие от схожих

работ, использующих автоматическую разметку для обучения алгоритмов CNN) [1].

На втором этапе врачами-рентгенологами проводилась проверка разметки для исключения несовпадений между работой специалистов на первом этапе и профессиональной оценкой МРТ-изображений. При отсутствии ошибок, выявленных на втором этапе, и дополнительных комментариев по качеству выделения контуров работа с кейсом в рамках процесса разметки считалась завершенной.

В дальнейшем, после обучения нейронной сети на размеченных вручную датасетах, разметка выполнялась в автоматизированном режиме для каждого исследования с отражением нумерации позвонков (рис. 1) [2, 3].

Определение размера и структуры обучающей, валидационной и тестовой выборок, а также процесс обучения CNN для анализа нормы и патологических изменений осуществлялись сотрудниками компании Remedy Logic. Анализ отдельных патологических изменений проводился с автоматизированным выбором релевантных сагиттальных или аксиальных срезов [4, 5].

После процесса обучения нейронных сетей была организована проверка точности их работы, которая выполнялась при содействии пяти врачей-рентгенологов с опытом работы по специальности от 5 до 10 лет. Проверка осуществлялась путем анализа МРТ-изображений с последующим сравнением профессиональных выводов специалистов и результатов работы ИИ. Архивы, использованные в процессе данной проверки, ранее не применялись для формирования обучающей, валидационной и тестовой выборок, участвовавших в процессе обучения нейронных сетей.

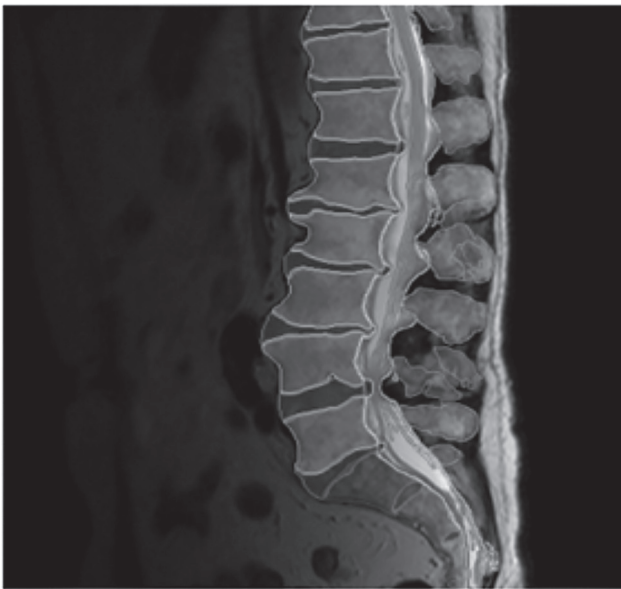
По итогу проверки формировались промежуточный сравнительный отчет (в рамках не опубликованной в медицинских журналах промежуточной научно-исследовательской работы) и итоговый сравнительный отчет, результаты которого представлены в данной статье. В промежуточном отчете проводилась первичная оценка точности работы нейросетей в выявлении выбуханий, протрузий и экструзий дисков, стеноза позвоночного канала, латеральных стенозов, фораминальных стенозов, спондилолистеза, а также артроза межпозвонковых суставов. В итоговом отчете помимо повторной оценки точности выявления патологических изменений, обозначенных в промежуточном отчете, была добавлена проверка дегенеративных изменений замыкательных пластин, синовита межпозвонковых суставов, дегенерации межпозвонковых дисков, остеофитов, переходных позвонков, гипертрофии желтых связок и грыж Шморля. Также в итоговом отчете определялись количественные измерения передне-задних размеров и высоты протрузий и экструзий межпозвонковых дисков.

Результаты и обсуждение /
Results and discussion

Осуществлен процесс автоматической отправки анонимизированных архивов исследований пациентов, прошедших МРТ-обследование, на сервер, физически расположенный на территории Российской Федерации, для обработки на обученной модели ИИ и последующей отправки результатов в радиологическую информационную систему диагностических центров. Каждому анонимизированному исследованию перед отправкой присвоен уникальный идентификационный номер, который также указан в отчете ИИ [6].

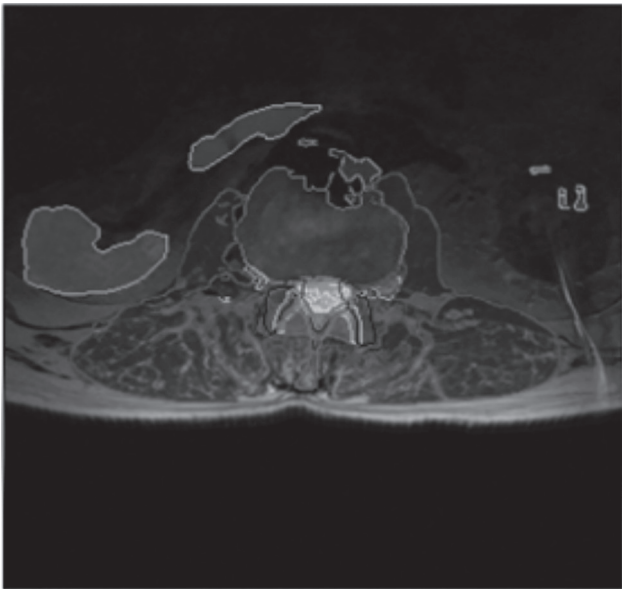
Выходные данные ИИ для каждого кейса представлены JSON-файлом с указанием уровня пояснично-крестцового отдела позвоночника, названия патологии для каждого отдела и количественных измерений или степени патологических изменений. В рамках радиологической информационной системы полученный JSON-файл преобразуется в PDF-файл, доступ к которому предоставляется в интерфейсе системы в виде интерактивного значка,

StudyUID 1.3.12.2.1107.5.2.6.23710.30000015570.2023012550930226576



Уровень: L1/L2	
экструзия	нет
протрузия	да
выбухание	да

Степени стеноза	
фораминальный стеноз	2-й степени



площадь дурального мешка	221,56
диаметр дурального мешка	14,06
площадь позвоночного канала	311,28
диаметр позвоночного канала	21,09
площадь позвоночного канала справа	21,97
площадь позвоночного канала слева	26,86
площадь фораминального окна справа	16,48
площадь фораминального окна слева	16,48

Рис. 2. Пример отчета по результатам работы искусственного интеллекта в PDF-файлах на дисковом уровне L1/L2
Fig. 2. Example of report on the results of artificial intelligence in PDF files at disk L1/L2 level

расположенного за каждой заявкой пациента, прошедшего обследование пояснично-крестцового отдела позвоночника. Время получения PDF-файла после прохождения обследования составляет менее 15 мин, далее врач может ознакомиться с полученной информацией о патологических изменениях и использовать ее в своем протоколе и заключении. В документе выведены данные о характере патологических изменений на уровне каждого диска вместе с изображениями – одним сагиттальным снимком, локализованным по срединной линии, и по одному аксиальному срезу на каждый представленный дисковый уровень. Изображения уровней без патологических изменений, а также информация об отсутствии стенозов и артроза межпозвонковых суставов в документе не отображаются во избежание перегрузки протокола лишней информацией (рис. 2).

Результаты по выявлению патологических изменений соответствуют проверенным в промежуточном отчете, а именно: выбухания, протрузии, экструзии дисков, центральный стеноз позвоночного канала, латеральные стенозы, фораминальные стенозы, спондилолистез, артроз межпозвонковых суставов. Количественные измерения, представленные в отчете, указаны в качестве справочной информации, их точность не проверялась в рамках текущего исследования точности алгоритмов. Информация о патологических изменениях из итогового отчета не выводилась в PDF-файл в рамках этапа интеграции с радиологической информационной системой, представленного на изображении.

Далее в рамках процесса последовательного улучшения работы сервиса ИИ в результаты анализа были добавлены следующие патологические изменения: дегенеративные изменения замыкательных пластин, синовит межпозвонковых суставов, дегенерация дисков, остеофиты, переходные позвонки, гипертрофия желтых связок и грыжи Шморля. Оценка работы алгоритма проводилась не только по добавленным патологическим изменениям, но и по патологическим изменениям, ранее фигурировавшим в промежуточном отчете.

В процессе проверки работы алгоритма для итогового отчета участвовали пять врачей-рентгенологов с опытом работы по интерпретации МРТ-изображений от 5 до 10 лет. Каждому врачу было предоставлено на описание по 20 одинаковых DICOM с МРТ-исследованием пояснично-крестцового отдела позвоночника. Кейсы, участвовавшие в проверке, были отобраны случайно, а также не были задействованы как часть обучающих, проверочных и тестирующих выборок в рамках работ по обучению и повышению

точности функционирования ИИ. В каждом исследовании была проведена оценка пяти дисковых уровней, что в сумме, соответственно, дает по 100 оцененных анатомических уровней на каждого специалиста.

Далее каждый врач сравнил свои результаты с результатами работы алгоритма и указал на наличие возникших расхождений. Если оценка трех врачей для конкретной патологии в одной локализации совпадала, то она устанавливалась как эталонная и относительно нее проводилось сравнение результативности алгоритма. Показатели для большинства патологических изменений имеют значения $\geq 0,8$. Предельно высокие показатели чувствительности и специфичности выявления переходных позвонков, специфичности определения спондилолистезов и грыж Шморля обусловлены малым количеством данных патологических изменений в рамках случайно отобранных исследований пояснично-крестцового отдела позвоночника.

ИИ продемонстрировал сопоставимые значения чувствительности и специфичности в сравнении с эталонным результатом в группе опытных врачей-рентгенологов (см. табл.).

Схожие результаты были продемонстрированы в других исследованиях, выполненных на базе обучения CNN для анализа дегенеративных изменений поясничного отдела позвоночника в целом [7, 8] и изменений, связанных с экструзиями и протрузиями дисков, в частности [9–11]. Значения $< 0,8$ были зафиксированы в показателях чувствительности выявления дегенерации замыкательных пластин, гипертрофии желтых связок и грыж Шморля, а также специфичности определения артроза межпозвонковых суставов и остеофитов.

Анализ количественных измерений пролабирования дисков (экструзий и протрузий) в просвет позвоночного канала проводился в рамках проверки точности выявления патологических изменений поясничного отдела позвоночника. Автоматизированная количественная оценка выбухания дисков не ставилась как задача работы алгоритмов ввиду низкой клинической значимости данных изменений. В качестве результатов работы ИИ были представлены количественные данные о передне-заднем размере и высоте сегментов дисков, пролабирующих в просвет позвоночного канала, без отдельного указания высоты миграции диска выше или ниже края замыкательной пластины тела прилежащего позвонка. Результаты проверки врачами фиксировались в виде оценок «правильно», «частично правильно» и «неправильно» в зависимости от степени их расхождения с результатами работы алгоритмов. Оценка «правильно» устанавливалась при совпадении обоих

Значения чувствительности и специфичности заключений с использованием искусственного интеллекта

Sensitivity and specificity values of conclusions using artificial intelligence

Патологические изменения / Pathological changes	Чувствительность, % / Sensitivity, %	Специфичность, % / Specificity, %
Экструзии (с измерением в мм) / Extrusion (measured in mm)	0,885	0,973
Протрузии (с измерением в мм) / Protrusions (measured in mm)	0,809	0,936
Центральный стеноз / Central stenosis	0,867	0,976
Латеральный стеноз / Lateral stenosis	0,834	0,855
Фораминальный стеноз / Foraminal stenosis	0,921	0,839
Артроз / Arthrosis	0,852	0,531
Дегенерация замыкательных пластин / Occlusive plates degeneration	0,734	0,964
Синовит межпозвонковых суставов / Intervertebral joint synovitis	0,850	0,838
Дегидратация дисков / Disc dehydration	0,911	0,879
Остеофиты / Osteophytes	0,929	0,721
Переходные позвонки / Transitional vertebrae	1,000	1,000
Спондилолистезы / Spondylolisthesis	0,800	1,000
Гипертрофия желтой связки / Yellow ligament hypertrophy	0,667	0,989
Грыжа Шморля / Schmorl's hernia	0,750	1,000

параметров, «частично правильно» – при совпадении одного из двух измеряемых параметров, «неправильно» – при полном отсутствии совпадений. Измерения врачами пролабировавших сегментов дисков осуществлялись с учетом толщины задней продольной связки.

На данном этапе точность количественного определения протрузий и экструзий оказалась низкой, расхождение между результатами врачей и алгоритма в большинстве случаев составило около 3 мм. В основном это связано с преимущественным определением экструзий и протрузий дисков только на уровне срединного сагиттального среза, что соответствует медианному сектору на аксиальных срезах. Данная ошибка будет скорректирована в последующих апдейтах ИИ.

На итоговом этапе был реализован автоматизированный обмен анонимизированной информацией между сервером, на котором хранятся исследования, сервером, где происходит работа ИИ, и серверами, на которых осуществляется работа радиологической информационной системы. Реализован способ презентации результатов работы ИИ в виде интерактивного значка в интерфейсе радиологической информационной системы, содержащего PDF-файл. В PDF-файле отображаются автоматизированно размеченные изображения в сагиттальной и аксиальной плоскостях, а также текстовая информация о патологических изменениях на каждом уровне.

В рамках итогового отчета зафиксированы хорошие результаты в определении большинства распространенных дегенеративных изме-

нений пояснично-крестцового отдела позвоночника. Также наблюдается воспроизведение высоких значений чувствительности и специфичности для патологий, демонстрировавших аналогичные высокие показатели в промежуточном отчете. Высокое качество диагностики патологических изменений позвоночника при помощи современных алгоритмов нейронных сетей было подтверждено и в других работах [12, 13]. В то же время процесс определения количественных показателей нуждается в доработке в виде оценки на каждом сагиттальном срезе для выявления наибольшего размера пролабирования диска в просвет позвоночного канала.

Сокращение временных затрат на интерпретацию и написание протокола врачами-рентгенологами составляет от 2 до 10 мин на каждое исследование в зависимости от опыта работы специалиста и выраженности дегенеративных изменений пояснично-крестцового отдела позвоночника.

Дальнейшее развитие алгоритмов нейронных сетей приведет к расширению функционала сервисов, основанных на ИИ, вплоть до полной автоматизации решения диагностических задач среди наиболее часто выявляемых патологических изменений методами лучевой диагностики. Совокупность возрастающего числа специалистов в сфере компьютерного зрения и увеличения доступных вычислительных мощностей создают условия для реализации на практике более сложных и точных моделей ИИ, способных в короткие сроки справиться с большим объемом и большей сложностью диагностических задач.

Заключение / Conclusion

Точность метода МРТ в оценке дегенеративных изменений пояснично-крестцового отдела позвоночника является самой высокой среди всех неинвазивных методов диагностики, при этом рост доступности методов лучевой диагностики прямо пропорционален увеличению количества проводимых исследований.

Значительный объем работы при анализе пояснично-крестцового отдела позвоночника в большинстве случаев складывается из повторяющегося процесса нахождения и измерения типичных проявлений дегенеративных изменений в каждом анатомическом сегменте. Как и множество других специальностей, лучевая

диагностика находится на пути к более плотному взаимодействию с системами ИИ для повышения качества и скорости выполняемой работы. С помощью современной генерации CNN возможно автоматизированное выявление целого ряда патологий на МРТ-изображениях пояснично-крестцового отдела позвоночника, что способствует снижению нагрузки на опытных врачей-рентгенологов и вероятности ошибки специалистов с малым опытом работы.

В дальнейшем вследствие повышения доступности высокоэффективных сервисов ИИ будет происходить все более плотное внедрение их в медицинскую практику, что повысит скорость и качество оказания медицинских услуг.

Литература / References

1. Li H, Luo H, Huan W, et al. Automatic lumbar spinal MRI image segmentation with a multi-scale attention network. *Neural Comput Appl*. 2021; 33(18): 11589–602. <https://doi.org/10.1007/s00521-021-05856-4>.
2. Forsberg D, Sjöblom E, Sunshine JL. Detection and labeling of vertebrae in mr images using deep learning with clinical annotations as training data. *J Digit Imaging*. 2017; 30(4): 406–12. <https://doi.org/10.1007/s10278-017-9945-x>.
3. Zhou Y, Liu Y, Chen Q, et al. Automatic lumbar MRI detection and identification based on deep learning. *J Digit Imaging*. 2019; 32(3): 513–20. <https://doi.org/10.1007/s10278-018-0130-7>.
4. Natalia F, Young JC, Afriliana N, et al. Automated selection of mid-height intervertebral disc slice in traverse lumbar spine MRI using a combination of deep learning feature and machine learning classifier. *PLoS One*. 2022; 17(1): e0261659. <https://doi.org/10.1371/journal.pone.0261659>.
5. Wang X, Zhai S, Niu Y. Automatic vertebrae localization and identification by combining deep SSAE contextual features and structured regression forest. *J Digit Imaging*. 2019; 32(2): 336–48. <https://doi.org/10.1007/s10278-018-0140-5>.
6. Lewandrowski KU, Muraleedharan N, Eddy SA, et al. Feasibility of deep learning algorithms for reporting in routine spine magnetic resonance imaging. *Int J Spine Surg*. 2020; 14(s3): S86–97. <https://doi.org/10.14444/7131>.
7. Hallinan JTPD, Zhu L, Yang K, et al. Deep learning model for automated detection and classification of central canal, lateral recess, and neural foraminal stenosis at lumbar spine MRI. *Radiology*. 2021; 300(1): 130–8. <https://doi.org/10.1148/radiol.2021204289>.
8. Jamaludin A, Lootus M, Kadir T, et al. ISSLS Prize in Bioengineering Science 2017: Automation of reading of radiological features from magnetic resonance images (MRIs) of the lumbar spine without human intervention is comparable with an expert radiologist. *Eur Spine J*. 2017; 26(5): 1374–83. <https://doi.org/10.1007/s00586-017-4956-3>.
9. Lehnen NC, Haase R, Faber J, et al. Detection of degenerative changes on mr images of the lumbar spine with a convolutional neural network: a feasibility study. *Diagnostics*. 2021; 11(5): 902. <https://doi.org/10.3390/diagnostics11050902>.
10. Tsai JY, Hung IYJ, Guo YL, et al. Lumbar disc herniation automatic detection in magnetic resonance imaging based on deep learning. *Front Bioeng Biotechnol*. 2021; 9: 708137. <https://doi.org/10.3389/fbioe.2021.708137>.
11. Han Z, Wei B, Leung S, et al. Automated pathogenesis-based diagnosis of lumbar neural foraminal stenosis via deep multiscale multitask learning. *Neuroinformatics*. 2018; 16(3–4): 325–37. <https://doi.org/10.1007/s12021-018-9365-1>.
12. Galbusera F, Casaroli G, Bassani T. Artificial intelligence and machine learning in spine research. *JOR Spine*. 2019; 2(1): e1044. <https://doi.org/10.1002/jsp2.1044>.
13. Azimi P, Yazdani T, Benzel EC, et al. A review on the use of artificial intelligence in spinal diseases. *Asian Spine J*. 2020; 14(4): 543–71. <https://doi.org/10.31616/asj.2020.0147>.