



Модель нейронной сети для выявления и классификации стенозов пояснично-крестцового отдела позвоночника на МР-томограммах

© Русаков А.С.¹, Тумко В.В.¹, Сарбаев Р.С.², Успенская Н.А.¹, Нуднов Н.В.³, Кармазановский Г.Г.^{4, 5}, Коробов А.В.⁶, Титова Л.А.⁷, Скачков А.А.⁸, Кульнева Т.В.⁶, Измалков Д.В.⁶, Андриенко Е.А.⁶, Иванников М.Е.^{3*}

¹ Remedy Logic; 6 St Johns Ln, New York, NY 10013, США

² ООО "Системы поддержки принятия решений"; 428022 Чебоксары, ул. Декабристов, зд.33, Чувашская Республика, Российская Федерация

³ ФГБУ "Российский научный центр рентгенодиагностики" Минздрава России; 117997 Москва, ул. Профсоюзная, д. 86, Российская Федерация

⁴ ФГБУ "Национальный медицинский исследовательский центр хирургии им. А.В. Вишневского" Минздрава России; 117997 Москва, ул. Большая Серпуховская, д. 27, Российская Федерация

⁵ ФГАОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России; 117997 Москва, ул. Островитянова, д. 1, Российская Федерация

⁶ Автономная некоммерческая организация дополнительного профессионального образования "Институт Эксперт"; 394018 Воронеж, ул. Фридриха Энгельса, д. 58А, Российская Федерация

⁷ ФГБОУ ВО "Воронежский государственный медицинский университет имени Н.Н. Бурденко" Минздрава России; 394036 Воронеж, ул. Студенческая, д. 10, Российская Федерация

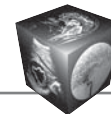
⁸ ООО "Объединенное IT пространство"; 398017 Липецк, ул. Metallургов, владение 1, литер А, помещение 18, Российская Федерация

Цель исследования: разработать модель свёрточной нейронной сети (convolutional neural network, CNN) для классификации наличия и тяжести стенозов поясничного отдела позвоночника (lumbar spinal stenosis, LSS) в ходе МРТ-исследования и продемонстрировать ее эффективность в качестве точного и последовательного диагностического инструмента.

Материал и методы. Морфологические классификации стенозов поясничного отдела позвоночника и количественные измерения ключевых анатомических структур были объединены с помощью разных моделей свёрточных нейронных сетей. Для классификации стенозов модели были обучены на 1635 размеченных исследованиях МР-томограммах поясничного отдела позвоночника. Данные исследования состояли из изображений, взвешенных по T2 в сагиттальной и аксиальной плоскостях на уровне каждого позвонка. Точность модели была оценена с помощью внешнего валидационного набора из 150 МРТ-исследований, оцененных группой из 7 рентгенологов по градации в виде: отсутствия стеноза, незначительной, умеренной или выраженной степени стенозов позвоночного канала. Эталонное значение для всех типов стенозов было определено большинством голосов и в случае разногласий решение выносилось опытным врачом-рентгенологом, не участвовавшим в подсчете голосов. После этого выводы врачей были сопоставлены с выводами обученной модели.

Результаты. Модель продемонстрировала сопоставимую производительность со средними показателями рентгенолога как с точки зрения выявления наличия/отсутствия центральных стенозов поясничного отдела позвоночника, так и классификации степени тяжести для всех трех типов стеноза. В случае стеноза центрального канала чувствительность и специфичность свёрточной нейронной сети составили 0,93 и 0,85 для бинарной классификации (наличие/отсутствие) по сравнению со средним показателем рентгенолога (0,86; 0,86). При стенозе латерального кармана чувствительность и специфичность свёрточной нейронной сети составили 0,92 и 0,80 по сравнению со средним показателем рентгенолога 0,83 и 0,94. При foraminalном стенозе чувствительность и специфичность свёрточной нейронной сети составили 0,89 и 0,86 по сравнению со средним показателем рентгенолога 0,81 и 0,91. Многоклассовая классификация степени тяжести стенозов показала сходные статистические данные.

Заключение. Свёрточная нейронная сеть продемонстрировала сопоставимые со специалистами-рентгенологами результаты в выявлении и классификации центральных стенозов поясничного отдела позвоночника. Интеграция моделей нейронных сетей при обнаружении патологии могла бы обеспечить более высокую точность, эффективность, системность и возможность последующей интерпретации в диагностической практике.



Ключевые слова: стеноз поясничного отдела позвоночника; магнитно-резонансная томография; МРТ; искусственный интеллект; нейронная сеть; машинное обучение

Авторы подтверждают отсутствие конфликтов интересов.

Для цитирования: Русаков А.С., Тумко В.В., Сарбаев Р.С., Успенская Н.А., Нуднов Н.В., Кармазановский Г.Г., Коробов А.В., Титова Л.А., Скачков А.А., Кульнева Т.В., Измалков Д.В., Андриенко Е.А., Иванников М.Е. Модель нейронной сети для выявления и классификации стенозов пояснично-крестцового отдела позвоночника на МР-томограммах. *Медицинская визуализация*. 2025; 29 (1): 102-112. <https://doi.org/10.24835/1607-0763-1436>

Поступила в редакцию: 22.01.2024.

Принята к печати: 19.11.2024.

Опубликована online: 20.01.2025.

A neural network model for detection and classification of central lumbosacral spinal stenosis on MRI scans

© Andrey S. Rusakov¹, Vladislav V. Tumko¹, Ruslan S. Sarbaev²,
Natalia A. Uspenskaya¹, Nikolay V. Nudnov³, Grigory G. Karmazanovsky^{4, 5},
Andrey V. Korobov⁶, Lilia A. Titova⁷, Artur A. Skachkov⁸, Taisiya V. Kulneva⁶,
Dmitry V. Izmailkov⁶, Elizaveta A. Andrienko⁶, Mikhail E. Ivannikov^{3*}

¹ Remedy Logic; 6 St Johns Ln, New York, NY 10013, USA

² LLC "Decision Support Systems"; 33, Dekabristov str., Cheboksary 428022, Chuvash Republic, Russian Federation

³ Russian Scientific Center of Roentgenoradiology of the Ministry of Healthcare of the Russian Federation; 86, Profsoyusnaya str., Moscow 117997, Russian Federation

⁴ A.V. Vishnevsky National Medical Research Center of Surgery of the Ministry of Healthcare of the Russian Federation; 27, Bol'shaya Serpukhovskaya str., Moscow 117997, Russian Federation

⁵ Pirogov Russian National Research Medical University; 1, Ostrivtynova str., Moscow 117997, Russian Federation

⁶ Autonomous non-profit organization of additional professional education "Expert Institute"; 58A, Friedrich Engels str., Voronezh 394018, Russian Federation

⁷ Voronezh State Medical University named after N.N. Burdenko; 10, Studencheskaya str., Voronezh 394036, Russian Federation

⁸ LLC "United IT Space"; 1, liter A, 18, Metallurgov str., Lipetsk 398017, Russian Federation

Objective: to develop a convolutional neural network (CNN) model to classify the presence and severity of central lumbar spinal stenosis (LSS) during MRI examination and to demonstrate its effectiveness as an accurate and consistent diagnostic tool.

Methods. Morphological classifications of LSS and quantitative measurements of key anatomical structures were combined using two CNNs. To classify central stenosis, models were trained on 1635 labeled lumbar spine MRI studies consisting of T2-weighted sagittal and axial planes at the level of each vertebra. The accuracy of the model was assessed using an external validation set of 150 MRI studies graded by a panel of 7 radiologists as: no stenosis, mild, moderate or severe spinal canal stenosis. The reference value for all types of stenosis was determined by majority vote, and in the event of disagreement, a decision was made by an external radiologist. The radiologists' interpretations were then compared with those of the trained model.

Results. The model demonstrated comparable performance to the average radiologist both in identifying the presence/absence of LSS and in classifying severity for all 3 types of stenosis. For central canal stenosis, the sensitivity and specificity of the CNN were (0.93; 0.85) for binary classification (presence/absence) compared to the average radiologist (0.86; 0.86). For lateral pocket stenosis, the sensitivity and specificity of CNN were (0.92; 0.80) compared to the radiologist's mean (0.83; 0.94). For foraminal stenosis, the sensitivity and specificity of CNN were (0.89; 0.86) compared to the radiologist's mean (0.81; 0.91). Multiclass classification of stenosis severity showed similar statistics.

Conclusions. CNNs showed comparable performance to radiologists in detecting and classifying LSS. The integration of neural network models in pathology detection could provide higher accuracy, efficiency, systematicity, and the possibility of subsequent interpretation in diagnostic practice.

Keywords: lumbar spinal stenosis; magnetic resonance imaging; MRI; artificial intelligence; neural network; machine learning

Conflict of interest. The authors declare no conflict of interest. The study had no sponsorship.



For citation: Rusakov A.S., Tumko V.V., Sarbaev R.S., Uspenskaya N.A., Nudnov N.V., Karmazanovsky G.G., Korobov A.V., Titova L.A., Skachkov A.A., Kulneva T.V., Izmailkov D.V., Andrienko E.A., Ivannikov M.E. A neural network model for detection and classification of central lumbosacral spinal stenosis on MRI scans. *Medical Visualization*. 2025; 29 (1): 102–112. <https://doi.org/10.24835/1607-0763-1436>

Received: 22.01.2024.

Accepted for publication: 19.11.2024.

Published online: 20.01.2025.

Введение

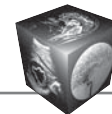
По данным Всемирной организации здравоохранения, боль в спине является основным экономическим бременем, а также одной из основных причин инвалидизации и обращения за медицинской помощью. Распространенной причиной боли являются дегенеративные заболевания позвоночника [1, 2]. Стенозы позвоночного канала являются распространенным дегенеративным заболеванием. Частота выявляемого поясничного стеноза составляет 5 случаев на 1000 обследованных и увеличивается с возрастом: до 40 лет распространенность стеноза составляет 24%, а в возрасте 60–69 лет – более 60% [3]. Клинические симптомы стенозов поясничного отдела позвоночника могут включать боль в пояснице, нижних конечностях и ягодицах, онемение, значительную потерю подвижности и серьезные ограничения активной социальной жизни пациентов [4]. Среди дегенеративных заболеваний позвоночника стенозы представляют собой одну из наиболее распространенных причин хирургического вмешательства на позвоночнике у пациентов старше 65 лет, следовательно, точная и надежная диагностика и последующее лечение имеют важное экономическое и социальное значение. Стенозы поясничного отдела позвоночника характеризуются сужением переднезаднего размера позвоночного канала, сужением латеральных/субарткулярных карманов и фораминальных окон, что, в свою очередь, вызывает сдавливание соответствующих нервных корешков [5].

Стенозы в поясничном отделе позвоночника определяются как клинический синдром и диагностируется как таковой. Диагностическая медицинская визуализация часто используется в качестве подтверждающего метода перед определением тактики лечения. Магнитно-резонансная томография (МРТ) является основным методом детальной анатомической оценки позвоночника с высококачественным дифференцированием мягких тканей и используется для подтверждения диагнозов, основанных на клинических симптомах, и определения оптимального курса лечения. В ряде исследований была предпринята попытка определить основные количественные радиологические кри-

терии для диагностики стенозов поясничного отдела позвоночника [6–8]. Некоторые измерения, такие как площадь поперечного сечения дурального мешка, широко признаны диагностически значимыми, однако окончательная взаимосвязь между проявлением клинических симптомов и конкретными рентгенологическими измерениями остается неясной.

Мнение экспертов-практиков сочетает в себе как качественную оценку морфологических признаков, так и количественные измерения, однако отсутствие единого мнения относительно критериев принятия клинических решений поднимает серьезные вопросы о надежности данной оценки. Следовательно, для повышения точности и достоверности диагностики необходимы методы независимой от врача оценки, сочетающие качественные и количественные характеристики. Модели машинного обучения (ML), включая глубокие свёрточные нейронные сети (CNN), уже успешно применяются для оценки стенозов в поясничном отделе позвоночника и других дегенеративных изменений с высокой точностью в различных подходах [9, 10]. Однако большинство алгоритмов свёрточных нейронных сетей основаны на однокомпонентных моделях и бинарной классификации (присутствует/отсутствует) центрального стеноза поясничного отдела позвоночника. В работе [11] применялась двухкомпонентная свёрточная нейронная сеть для обнаружения стенозов позвоночного канала, латеральных карманов и фораминальных окон с последующей градацией стеноза.

Цель исследования: разработка двухэтапной модели с использованием алгоритмов свёрточной нейронной сети, которая сочетает в себе как количественные измерения ключевых анатомических структур, так и качественные морфологические особенности для обнаружения стенозов поясничного отдела позвоночника на аксиальных и сагиттальных МРТ-изображениях. Классификация, выполненная свёрточной нейронной сетью, охватывает выявление всех трех типов стеноза – центрального канала, латеральных карманов и фораминальных окон. Впоследствии производительность модели сравнивали с работой врачей-рентгенологов, чтобы проверить ее надежность и точность.



Материал и методы

Набор данных и аннотация

Первоначальный набор данных представлен дата-сетом, собранным авторами, состоящим из 1635 МРТ-исследований, выполненных на МРТ-аппаратах Philips Achieva 1,5 Тл, Philips Intera 1,5 Тл и Siemens Magnetom Symphony 1,5 Тл. В наборе данных находились DICOM-изображения взрослых пациентов (старше 18 лет), направленных на МРТ поясничного отдела позвоночника по поводу болевого синдрома в области поясницы. Пациенты с имплантатами, тяжелым сколиозом и плохим качеством изображения были исключены. Каждое МРТ-исследование проводилось с использованием стандартного протокола исследования поясничного отдела позвоночника, включая T1- и T2-взвешенные сагиттальные изображения без и с режимом жироподавления, а также T2-взвешенные аксиальные импульсные последовательности со сбалансированными метками. В качестве этапа предварительной обработки изучаемые изображения были обрезаны до области поясничных позвонков. Группой врачей-рентгенологов был составлен алгоритм разметки изображений с определением его порядка, количества размеченных анатомических структур и патологических изменений, границ между вариантом нормы и патологией, допустимой точности разметки для каждой анатомической структуры и патологического процесса. На изображениях в сагиттальной плоскости были сегментированы межпозвонковые диски, поясничные и крестцовые позвонки, позвоночный канал, спинной мозг с нервными корешками, задняя продольная связка и остистые отростки. Для аксиальных изображений были размечены мышечная ткань, межпозвонковые диски, позвоночный канал, фораминальные окна, нервные корешки, латеральные карманы, фасеточные суставы, остистый отросток, суставной отросток, желтая связка, протрузии, экструзии дисков, аорта, нижняя полая вена и почки.

В дополнение к сегментации ключевых анатомических структур врачи-рентгенологи, прошедшие подготовку в области опорно-двигательного аппарата, также оценили исследования по шкале: 0 (отсутствует), 1 (легкая степень), 2 (умеренная), 3 (тяжелая) в зависимости от степени центрального стеноза. Основание для разделения по степеням – измерение соотношения размера деформации дурального мешка относительно его нормальных размеров (классификация Lurie). Эталонный стандарт был определен по правилу голосования большинства и в случае несогласия утвержден другим рентгенологом. В ходе первого обучения модель сегментации выявляла фасеточные суста-

вы и позвоночный канал. Затем изображения были обрезаны по самым латеральным точкам дугоотростчатых суставов, а также по самым краниальным и каудальным точкам позвоночного канала с использованием дополнительных буферов ширины и высоты изображения, которые затем изменялись по размеру и увеличивались (горизонтальные перевороты и вращения). Из 1635 исследований 1390 были использованы для тренировки весов свёрточной нейронной сети, а 245 – в качестве валидационного набора при настройке гиперпараметров.

Кроме того, для окончательной оценки точности модели был использован набор отдельных данных, состоящий из 150 исследований, собранный из тех же источников, что и обучающий набор. Критерии включения/исключения были идентичны критериям набора обучающих данных. Контрольный набор данных для валидации был оценен по шкале отсутствия, легкой, умеренной или тяжелой степени центрального стеноза позвоночного канала группой из 7 радиологов-специалистов. Затем интерпретации рентгенолога были сопоставлены с интерпретацией модели.

Модель нейронной сети

Чтобы включить как количественные измерения, так и качественные модели формы в комплексную систему диагностики изображений, были созданы две параллельные последовательности из свёрточных нейронных сетей. Первая свёрточная нейронная сеть сегментировала ключевые анатомические объекты и выполнила измерения размера, формы, угла, контраста и площади в обеих (аксиальной и сагиттальной) плоскостях (рис. 1). Архитектура сегментации соответствует модифицированной архитектуре U-Net, которая структурирована из свёрток с повышающей выборкой, затем ректифицированных линейных единиц, свёрток с понижающей выборкой и окончательной карты свёртки [12]. На выходе получается 17 каналов бинарных масок присутствующих объектов – диск, дуральный мешок, нервный корешок, фораминальное окно, позвоночный канал, грыжа, почка, мышца, вена, артерия, фасета, латеральное углубление, остистый отросток, фасеточный сустав, суставной отросток, желтая связка. После сегментации объектов были проведены расчеты пространственных координат для количественных показателей, включая, помимо прочего, выпуклость диска, вентрально-дорсальный диаметр костного позвоночного канала, ширину, высоту и площадь поперечного сечения позвоночного канала и дурального мешка, соотношение диаметра и площади поперечного сечения, асимметрию

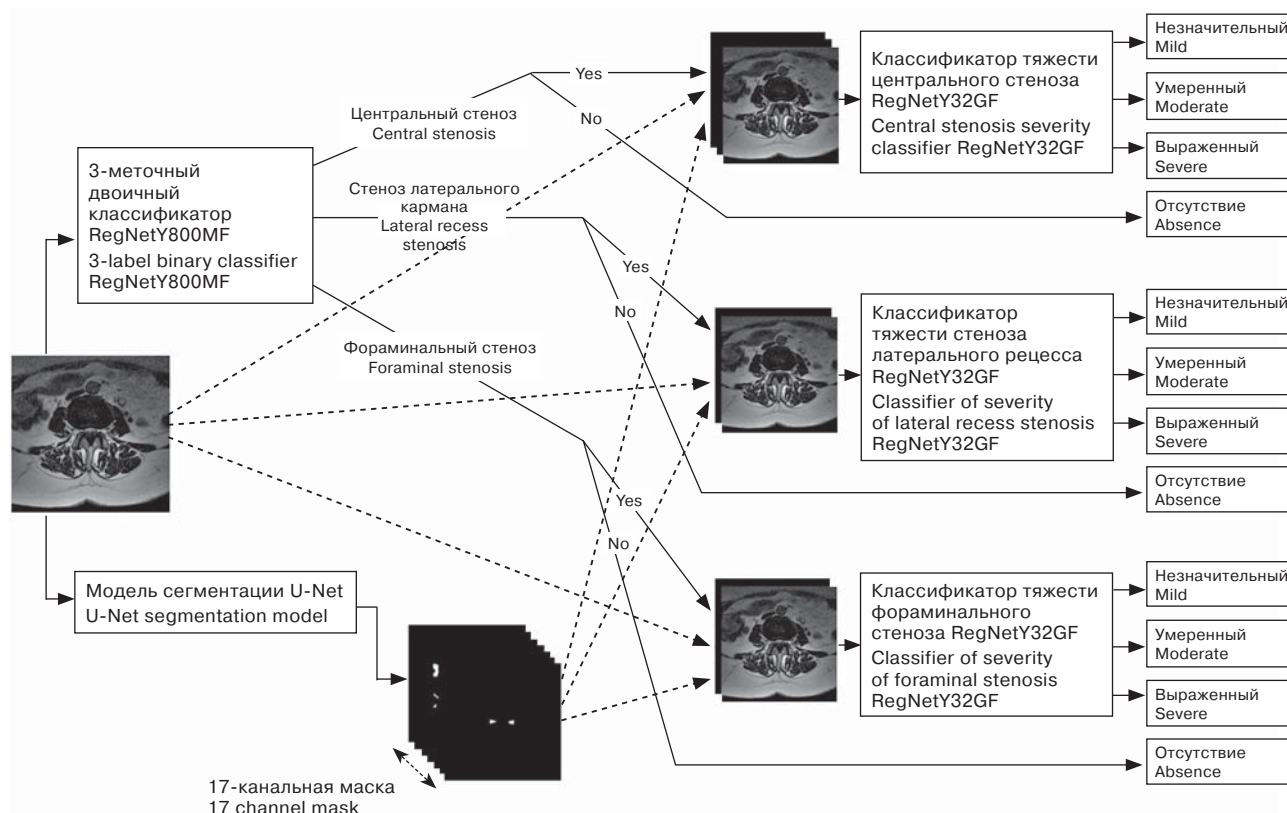


Рис. 1. Репрезентативные T2-взвешенные МРТ-изображения поясничного отдела позвоночника, демонстрирующие автоматизированный процесс сегментации ключевых анатомических объектов с помощью первой сверхточной нейронной сети в аксиальной плоскости.

Fig. 1. Representative T2-weighted MR images of the lumbar spine demonstrating the automated process of segmenting key anatomical landmarks using the first super-accurate neural network in the axial plane.

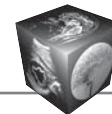
и пространственные размеры грыжи при ее наличии. Эти показатели легли в основу определения наличия и выраженности стеноза центрального канала в одной из частей общей модели.

Для второй модели мы построили последовательность свёрточных нейронных сетей непосредственно в качестве классификаторов. Все модели классификации основаны на архитектурах RegNet, глубоких остаточных сетях с серийным нормированием, идентификационными соединениями и уровнями узких слоев [13]. Первый этап представляет собой бинарный многометочный классификатор для определения наличия/отсутствия стенозов по локализации – центрального, латерального кармана и фораминального окна. Этот классификатор соответствует архитектуре regnet_y_800mf. Второй этап состоит из условных классификаторов степени тяжести. Для каждого типа стеноза (центрального, латерального, фораминального) существует один трехклассовый классификатор степени тяжести: легкой, умеренной и тяжелой. Он активен только при условии, что первый классификатор

определяет наличие стеноза. Эти классификаторы следуют архитектуре regnet_y_32_gf. Обучающий набор аксиальных МРТ-изображений для определения степени стеноза был размечен врачами-рентгенологами. Эта модель способна улавливать нюансы морфологических особенностей, указывающих на наличие центрального, латерального и фораминального стенозов. Окончательная модель представляет собой ансамбль моделей сегментации и классификации, так что количественные и качественные характеристики дополняют друг друга для повышения точности прогноза.

Статистический анализ

Результаты модели и результаты рентгенологов были сопоставлены с эталонным стандартом для выявления и классификации стенозов поясничного отдела позвоночника. Для бинарной классификации были рассчитаны чувствительность, специфичность и площадь под кривой ошибок (AUC-ROC). Для многоклассовой классификации эти показатели рассчитывались в сравнении



с остальными, где средние значения взвешивались по количеству наблюдений.

Межрейтинговое согласие между эталоном, моделью и радиологами оценивалось с помощью каппа-коэффициента Коэна и каппа-коэффициента Флейсса для двух- и множественных соглашений с оценщиками соответственно. Уровень согласия определялся следующим образом: <0 – плохой; $0,0-0,2$ – незначительный; $0,21-0,4$ – удовлетворительный; $0,41-0,6$ – умеренный; $0,61-0,8$ – значительный; $0,81-1,0$ – почти идеальный. В многоклассовой среде также была вычислена альфа Кrippендорфа для учета ординарности в оценках тяжести. Все расчеты выполнялись на языке Python (версия 3.7.4) с использованием scikit-learn (версия 1.0.2).

Результаты исследования

Бинарная классификация

Для сравнения результатов бинарной классификации были объединены оценки 0 и 1 – как отрицательные; 2 и 3 – как положительные. Для оценки каждый тип стеноза рассматривался отдельно. Показатели чувствительности и специфичности приведены на рис. 2.

В случае стеноза центрального канала средние значения чувствительности, специфичности и AUC-ROC для рентгенологов составили 0,86; 0,866; 0,82 соответственно, в то время как показатели для модели были 0,90; 0,88; 0,89. 95% доверительные интервалы (ДИ) для показателей чувствительности и специфичности были следующими: 0,837; 0,944 и 0,866; 0,884 соответственно. Таким образом, модель продемонстрировала сопоста-

вимую производительность по сравнению со среднестатистическим врачом-рентгенологом, участвовавшим в исследовании. Для бинарного взаиморасчетного согласия средний показатель каппа Коэна был значительным между радиологами и эталонным стандартом со значением – 0,72. Среднее значение для модели свёрточных нейронных сетей составило – 0,77. Коэффициент множественной оценки каппа Флейсса среди оценщиков составил – 0,65 по сравнению с 0,76 между моделью и эталонным стандартом.

Для стеноза латерального кармана средние значения чувствительности, специфичности и AUC-ROC для рентгенологов составили 0,83; 0,94; 0,76 соответственно, в то время как показатели для модели были 0,92; 0,80; 0,87. 95% ДИ для показателей чувствительности и специфичности были следующими: 0,865; 0,954 и 0,790; 0,812. Для бинарной оценки среднее значение коэффициента каппы Коэна соотносилось между рентгенологами и эталонным стандартом со значением – 0,37. Среднее значение модели свёрточных нейронных сетей составило – 0,21.

Для фораминального стеноза средние значения чувствительности, специфичности и AUC-ROC для рентгенологов составили 0,81; 0,91; 0,89 соответственно, в то время как показатели для модели составили 0,89; 0,86; 0,87. 95% ДИ для показателей чувствительности и специфичности были следующими: 0,871; 0,916 и 0,847; 0,868 соответственно. Для бинарной оценки средний показатель каппа Коэна был значительным между радиологами и эталонным стандартом со значением – 0,63. Среднее значение для модели составило 0,54.

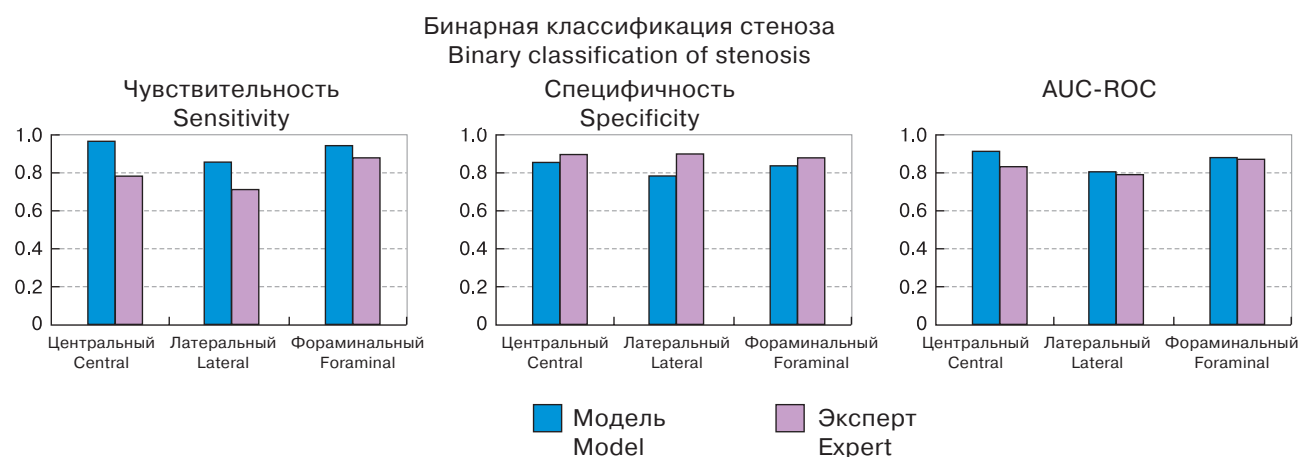


Рис. 2. Сравнение чувствительности, специфичности и AUC-ROC группы рентгенологов и модели классификации наличия/отсутствия стеноза поясничного отдела позвоночника.

Fig. 2. Comparison of sensitivity, specificity, and AUC-ROC of the radiologist group and the model for classifying the presence/absence of lumbar spinal stenosis.



Мультиклассовая классификация

Проводилось также сравнение показателей по полной шкале оценки центрального стеноза поясничного отдела позвоночника: отсутствие, легкая, среднетяжелая и тяжелая степень (рис. 3). Показатели эффективности для классификации по степени тяжести приведены на рис. 4.

Чтобы учесть классовый дисбаланс, показатели были рассчитаны с помощью весовых коэффициентов наблюдений. Для центрального стеноза средние показатели чувствительности, специфичности и AUC-ROC для рентгенологов составили 0,77; 0,86; 0,86 соответственно, тогда как для CNN эти показатели были 0,88; 0,90; 0,89. 95% ДИ для показателей чувствительности и специфичности были следующими: 0,866; 0,884 и 0,837; 0,944 соответственно. Среднее значение коэффициента Коэна между радиологами и эталонным стандартом составило 0,62 и 0,68 между моделью и эталонным стандартом. Порядковый показатель альфа Криппендорфа составил 0,70 у рентгенологов и 0,80 у модели и эталонного стандарта.

При стенозе латерального углубления средние показатели чувствительности, специфичности и AUC-ROC для рентгенологов составили 0,93; 0,59; 0,76 соответственно, тогда как для модели эти показатели были 0,80; 0,92; 0,87. 95% ДИ для показателей чувствительности и специфичности были следующими: 0,790; 0,812 и 0,865; 0,954 соответственно. Для взаимного согласования среднее значение коэффициента Коэна между радиологами и эталонным стандартом составило 0,32 и 0,19 между моделью и эталонным стандартом.

Для фораминального стеноза средняя чувствительность и специфичность для рентгенологов составили 0,87; 0,89; 0,89 соответственно, тогда как для CNN эти показатели были 0,86; 0,90; 0,87. 95% ДИ для показателей чувствительности и специфичности были следующими: 0,847; 0,868 и 0,871; 0,916 соответственно. Для взаимного согласования среднее значение коэффициента Коэна между радиологами и эталонным стандартом составило 0,56 и 0,48 между моделью и эталонным стандартом.

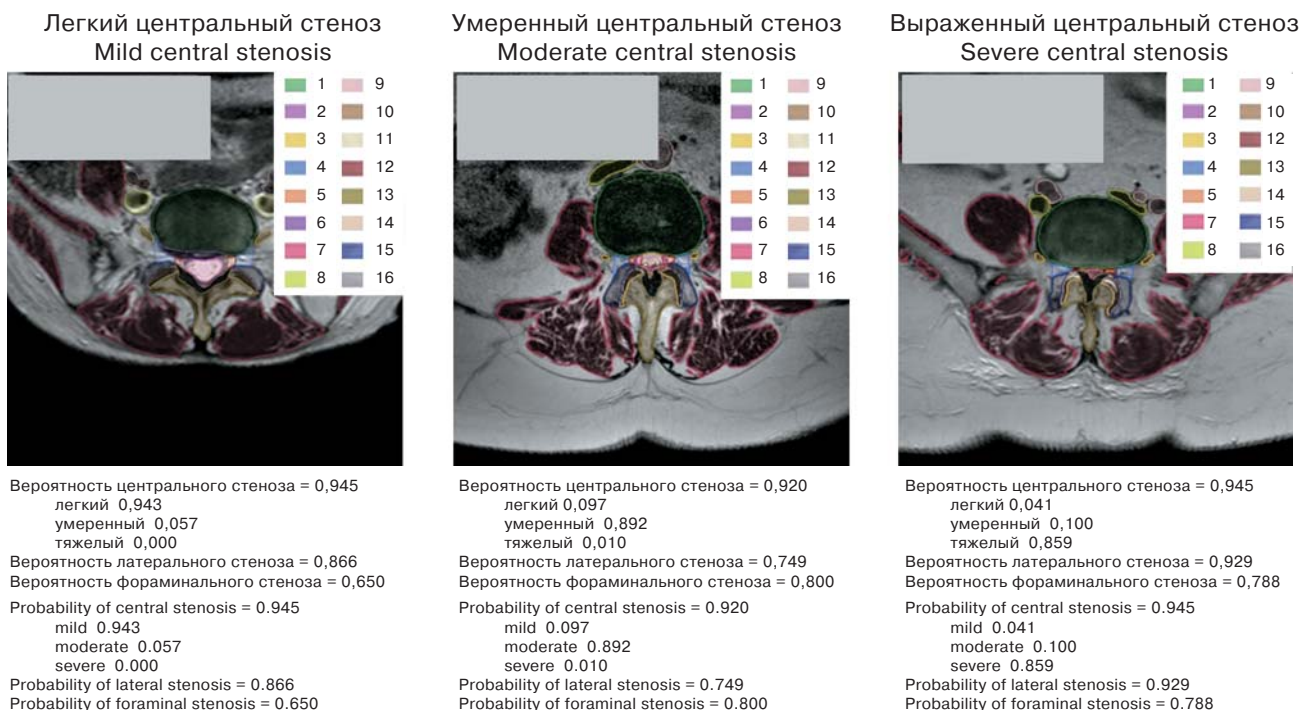


Рис. 3. Репрезентативные аксиальные T2-взвешенные МР-томограммы поясничного отдела позвоночника пациентов, которые были классифицированы нейронной сетью как легкий, умеренный и выраженный центральный стеноз. 1 – диск; 2 – дуральный мешок; 3 – нерв; 4 – нервные отверстия; 5 – спинной канал; 6 – грыжа; 7 – мышцы; 8 – вена; 9 – артерия; 10 – позвоночная дуга; 11 – нервные корни; 12 – боковая выемка; 13 – остистый отросток; 14 – фасеточные суставы; 15 – суставной процесс; 16 – желтая связка.

Fig. 3. Representative axial T2-weighted MRIs of the lumbar spine of patients classified by the neural network as having mild, moderate, and severe central stenosis. 1 – disc; 2 – dural sac; 3 – nerve; 4 – foraminal windows; 5 – spinal canal; 6 – hernia; 7 – muscles; 8 – vein; 9 – artery; 10 – vertebral arch; 11 – nerve roots; 12 – lateral pocket; 13 – spinous process; 14 – facet joints; 15 – articular process; 16 – ligamentum flavum.

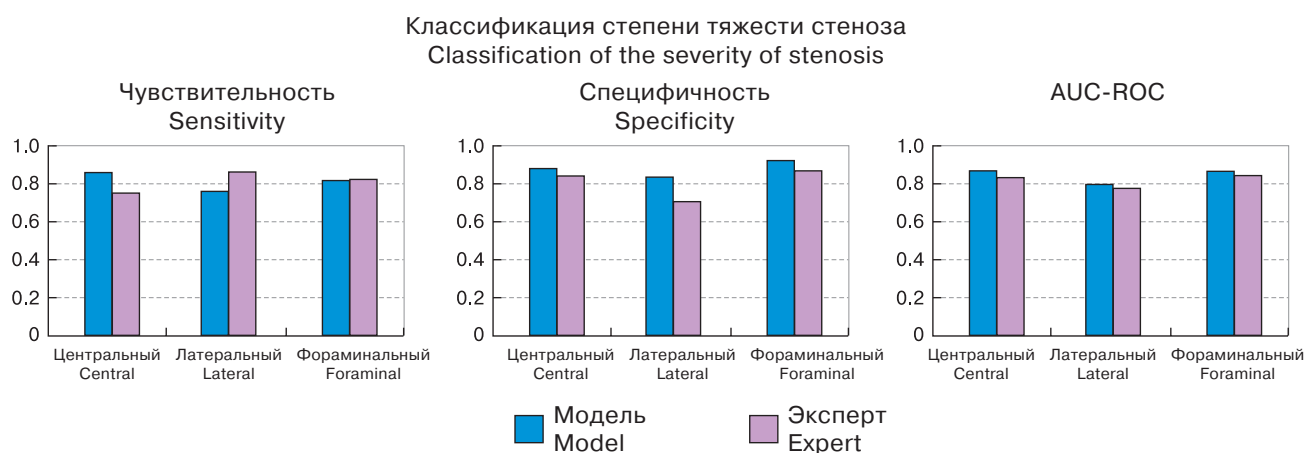
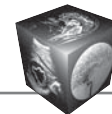


Рис. 4. Сравнение чувствительности, специфичности и AUC-ROC группы рентгенологов и модели для классификации тяжести стеноза поясничного отдела позвоночника.

Fig. 4. Comparison of sensitivity, specificity, and AUC-ROC of the radiology group and the lumbar spinal stenosis severity classification model.

В очередной раз модель продемонстрировала сопоставимую точность по сравнению со средне-статистическим специалистом-рентгенологом на контрольной панели при всех типах стеноза.

Обсуждение

Стеноз поясничного отдела позвоночника (LSS) – одна из самых распространенных патологий при визуализации и одно из наиболее частых показаний к операции на позвоночнике [14]. МРТ поясничного отдела позвоночника имеет важное значение для оценки стеноза поясничного отдела позвоночника и для подтверждения клинических данных, в значительной степени заменяя исследования КТ-миелографии из-за ее высокой контрастности мягких тканей. Тем не менее интерпретация МР-томограмм поясничного отдела позвоночника занимает много времени и зависит от знаний и опыта отдельного рентгенолога, поэтому желательна поддержка объективных и стандартизированных методов диагностики и принятия решений.

Целью данного исследования стала разработка модели из свёрточных нейронных сетей для классификации наличия и тяжести стеноза центрального канала с демонстрацией его эффективности в качестве точного и последовательного диагностического инструмента. Нами разработана ансамблевая модель свёрточных нейронных сетей, которая объединила количественные показатели сегментированных анатомических структур и морфологические особенности для прямой оценки стеноза позвоночного канала. Оценка по внешнему набору данных продемонстрировала аналогич-

ные результаты в сравнении с группой из 7 врачей-рентгенологов по выявлению и классификации тяжести стеноза позвоночного канала. Например, для бинарной диагностики стеноза центрального канала точность модели была выше по сравнению со средним показателем рентгенолога, достигая значений чувствительности, специфичности и AUC-ROC 0,88; 0,89; 0,89 по сравнению со средним показателем рентгенолога 0,86; 0,866; 0,82 соответственно. Для многоклассовой степени тяжести модели продемонстрировала результаты 0,81; 0,90; 0,85 по сравнению со средним показателем рентгенолога 0,77; 0,86; 0,86. По результатам внешней проверки соответствие между эталонным стандартом и рентгенологами было существенным для бинарной классификации (0,65) и многоклассовой классификации (0,62). Межэкспертное согласие между эталонным стандартом и моделью было несколько выше как для бинарной классификации (0,76), так и для многоклассовой классификации (0,68) по сравнению с данными рентгенологов.

Предыдущие исследования рассматривали эффективность различных моделей машинного обучения при обнаружении и классификации стенозов на МР-томограммах поясничного отдела позвоночника. В большинстве исследований для оценки эффективности моделей использовались два эксперта, в то время как при нашем подходе экспертная группа состояла из 7–8 радиологов. J.T.P.D. Hallinan и соавт. разработали двухэтапную модель глубокого обучения, используя первую модель для размещения интересующей области в центральном канале, латеральных карманах и фо-



раминальных окнах с последующей оценкой по второй модели [11]. В отличие от наших результатов для классификации центрального канала, их модель показала более высокие межклассовые соответствия, но лучшая точность фиксировалась у рентгенологов как для бинарной (0,98), так и для многоклассовой (0,89) классификации по сравнению с моделью глубокого обучения. На более крупном наборе данных из 7108 МР-томограмм поясничного отдела позвоночника на аксиальных и сагиттальных срезах J.T. Lu и соавт. разработали архитектуру U-Net, основанную на слабоконтролируемых метках обработки естественного языка, полученных из отчетов радиологов. Точность оценки потери по нескольким классам (нормальная–легкая–среднетяжелая–тяжелая) достигла в среднем 70,6% [15]. В другой работе двухэтапная свёрточная нейронная сеть была обучена локализовать дуральный мешок при аксиальном T2-взвешенном МРТ-исследовании в одном двигательном сегменте (L_{IV-V}) с вторичной оценкой локализованной области [16]. Согласованность оценок между двумя экспертными метками составила 77,5%, в то время как модель достигла точности 77,9–83,0% по сравнению с экспертными метками. Относительно недавно Z.H. Su и соавт. создали многозадачную классификационную сеть на основе большого набора данных аксиальной МРТ поясничного отдела позвоночника с T2-взвешиванием для оценки грыжи межпозвонкового диска, фораминального стеноза и стеноза позвоночного канала, которая была оценена двумя клиницистами в качестве эталонного стандарта [17]. Точность моделей варьировала от 79,7 до 87,0% для внешнего и внутреннего набора данных соответственно.

Несмотря на то что рентгенологическая оценка степени поясничного стеноза является основной частью диагностики и составления тактики лечения в клинической практике, она не всегда может коррелировать с тяжестью боли или функциональными нарушениями у некоторых пациентов [18]. Следовательно, модели нейронных сетей могут также включать клинические симптомы в дополнение к данным визуализации. Недавно была предложена гибридная модель свёрточной нейронной сети, сочетающая в себе возможности как клинической, так и визуализационной диагностики, для определения кандидатуры на хирургическое вмешательство у пациентов со стенозами поясничного отдела позвоночника [19]. В этом исследовании гибридная модель продемонстрировала высокую точность прогнозирования кандидатов на оперативное вмешательство с коэффициентом AUC-ROC 0,93 в сравнении с 0,84 по индивидуальной

рекомендации врача. Однако если гибридные модели, сочетающие в себе признаки как клинических факторов, так и данных визуализации, превосходят те, которые сочетают признаки, основанные исключительно на данных визуализации, то это должно быть подтверждено в будущих исследованиях.

Наше исследование имеет ограничения, большая часть которых является теоретическими и связанными с характером данного проекта – отсутствием “золотого стандарта” для сравнительного анализа текущей модели, то есть отсутствие твердо установленного консенсуса в отношении определений шкалы оценки МР-томограмм. В данном исследовании эталонные значения для каждого случая были установлены с помощью обобщенной радиологической оценки. Трудоемкий процесс ручного описания ограничивает количество МРТ-исследований, которые можно было бы использовать для обучения модели нейронной сети. Кроме того, в некоторых случаях не было достигнуто консенсуса между экспертами в отношении отдельных нюансов разметки. Заслуживает внимания функционал искусственного интеллекта, который будет учитывать аспекты как анализа изображений, так и клинических заключений, например, одновременный анализ дегенеративных изменений, деформации позвоночника и клинических симптомов. Однако разработка таких алгоритмов выходила за рамки данного исследования и будет изучаться в дальнейшем.

Заключение

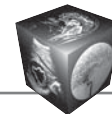
Таким образом мы продемонстрировали, что интеграция моделей свёрточных нейронных сетей способна диагностировать и оценивать стенозы на МР-томограммах поясничного отдела позвоночника с высокой точностью, надежностью и возможностью интерпретации post-hoc (найти закономерности после завершения исследования и обнаружить результаты, которые не были основной целью), сравнимой с результатами опытных рентгенологов, и может функционировать в качестве вспомогательного диагностического инструмента. В будущих работах модель свёрточных нейронных сетей должна быть развита для оценки более широкого спектра дегенеративных изменений, которые коррелируют с наличием у пациентов стенозов в поясничном отделе позвоночника.

Участие авторов

Русаков А.С. – концепция и дизайн исследования.

Тумко В.В. – проведение исследования.

Сарбаев Р.С. – анализ и интерпретация полученных данных.



Успенская Н.А. – статистическая обработка данных.
Нуднов Н.В. – утверждение окончательного варианта статьи.

Кармазановский Г.Г. – утверждение окончательного варианта статьи.

Коробов А.В. – подготовка, создание опубликованной работы.

Титова Л.А. – подготовка и редактирование текста.

Скачков А.А. – сбор и обработка данных.

Кульнева Т.В. – обзор публикаций по теме статьи.

Измалков Д.В. – написание текста.

Андрюенко Е.А. – ответственность за целостность всех частей статьи.

Иванников М.Е. – участие в научном дизайне.

Authors' participation

Rusakov A.S. – concept and design of the study.

Tumko V.V. – conducting research.

Sarbaev R.S. – analysis and interpretation of the obtained data.

Uspenskaya N.A. – statistical analysis.

Nudnov N.V. – approval of the final version of the article.

Karmazanovsky G.G. – approval of the final version of the article.

Korobov A.V. – preparation and creation of the published work.

Titova L.A. – text preparation and editing.

Skachkov A.A. – collection and analysis of data.

Kulneva T.V. – review of publications.

Izmalkov D.V. – writing text.

Andrienko E.A. – responsibility for the integrity of all parts of the article.

Ivannikov M.E. – participation in scientific design.

Список литературы [References]

- Kalichman L., Cole R., Kim D.H. et al. Spinal stenosis prevalence and association with symptoms: the Framingham Study. *Spine J.* 2009; 9 (7): 545–550. <https://doi.org/10.1016/j.spinee.2009.03.005>
- Wu A.M., Zou F., Cao Y. et al. Lumbar spinal stenosis: an update on the epidemiology, diagnosis and treatment. *Am. Med. J.* 2017; 2 (5): 63–63. <https://doi.org/10.21037/amj.2017.04.13>
- Katz J.N., Harris M.B. Lumbar Spinal Stenosis. *New Engl. J. Med.* 2008; 358 (8): 818–825. <https://doi.org/10.1056/nejmcp0708097>
- Lurie J., Tomkins-Lane C. Management of lumbar spinal stenosis. *BMJ.* 2016; 352: h6234. <https://doi.org/10.1136/bmj.h6234>
- Kreiner D.S., Shaffer W.O., Baisden J.L. et al. An evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis (update). *Spine J.* 2013; 13 (7): 734–743. <https://doi.org/10.1016/j.spinee.2012.11.059>
- Majidi H., Shafizad M., Niksolt F. et al. Relationship Between Magnetic Resonance Imaging Findings and Clinical Symptoms in Patients with Suspected Lumbar Spinal Canal Stenosis: a Case-control Study. *Acta Informatica Medica.* 2019; 27 (4): 229–233. <https://doi.org/10.5455/aim.2019.27.229-233>
- Steurer J., Roner S., Gnannt R. et al. Quantitative radiologic criteria for the diagnosis of lumbar spinal stenosis: a systematic literature review. *BMC Musculoskelet. Disord.* 2011; 12 (1): 175. <https://doi.org/10.1186/1471-2474-12-175>
- Andreisek G., Deyo R.A., Jarvik J.G. et al. Consensus conference on core radiological parameters to describe lumbar stenosis – an initiative for structured reporting. *Eur. Radiol.* 2014; 24 (12): 3224–3232. <https://doi.org/10.1007/s00330-014-3346-z>
- Lehnen N.C., 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>
- Jamaludin A., Kadir T., Zisserman A. SpineNet: Automated classification and evidence visualization in spinal MRIs. *Med. Image Anal.* 2017; 41: 63–73. <https://doi.org/10.1016/j.media.2017.07.002>
- Hallinan J.T.P.D., 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): 130–138. <https://doi.org/10.1148/radiol.2021204289>
- Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18 (p. 234–241). Springer International Publishing.
- Radosavovic I., Kosaraju R.P., Girshick R. et al. Designing network design spaces. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 10428–10436.
- Pannell W.C., Savin D.D., Scott T.P. et al. Trends in the surgical treatment of lumbar spine disease in the United States. *Spine J.* 2015; 15 (8): 1719–1727. <https://doi.org/10.1016/j.spinee.2013.10.014>
- Lu J.T., Pedemonte S., Bizzo B. et al. DeepSPINE: Automated Lumbar Vertebral Segmentation, Disc-level Designation, and Spinal Stenosis Grading Using Deep Learning. Arxiv. Published online 2018. <https://doi.org/10.48550/arxiv.1807.10215>
- Won D., Lee H.J., Lee S.J., Park S.H. Spinal Stenosis Grading in Magnetic Resonance Imaging Using Deep Convolutional Neural Networks. *Spine.* 2020; 45 (12): 804–812. <https://doi.org/10.1097/brs.0000000000003377>
- Su Z.H., Liu J., Yang M.S. et al. Automatic Grading of Disc Herniation, Central Canal Stenosis and Nerve Roots Compression in Lumbar Magnetic Resonance Image Diagnosis. *Front. Endocrinol.* 2022; 13: 890371. <https://doi.org/10.3389/fendo.2022.890371>
- Andrasinova T., Adamova B., Buskova J. et al. Is there a Correlation Between Degree of Radiologic Lumbar Spinal Stenosis and its Clinical Manifestation? *Clin. Spine Surg.* 2018; 31 (8): E403–E408. <https://doi.org/10.1097/bsd.0000000000000681>
- Mourad R., Kolisnyk S., Baiun Y. et al. Performance of hybrid artificial intelligence in determining candidacy for lumbar stenosis surgery. *Eur. Spine J.* 2022; 31 (8): 2149–2155. <https://doi.org/10.1007/s00586-022-07307-7>



Для корреспонденции*: Иванников Михаил Евгеньевич – e-mail: ivannikovmichail@gmail.com

Русаков Андрей Сергеевич – основатель и главный исполнительный директор Remedy Logic, New York.

Тумко Владислав Владимирович – главный операционный директор Remedy Logic, New York.

Сарбаев Руслан Сергеевич – генеральный директор ООО “Системы поддержки принятий решений”, Чебоксары.

Успенская Наталья Александровна – руководитель команды компьютерного зрения Remedy Logic, New York.

Нуднов Николай Васильевич – доктор мед. наук, профессор заместитель директора по научной работе, заведующий научно-исследовательским отделом комплексной диагностики заболеваний и радиотерапии ФГБУ “Российский научный центр рентгено-радиологии” Минздрава России, Москва. <https://orcid.org/0000-0001-5994-0468>

Кармазановский Григорий Григорьевич – академик РАН, доктор мед. наук, профессор, заведующий отделом лучевых методов диагностики ФГБУ “НМИЦ хирургии им. А.В. Вишневского” Минздрава России; профессор кафедры лучевой диагностики и терапии медико-биологического факультета ФГБОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России, Москва. <https://orcid.org/0000-0002-9357-0998>

Коробов Андрей Владимирович – директор АНО ДПО “Институт Эксперт”, Воронеж.

Титова Лилия Александровна – доктор мед. наук, доцент, заведующая кафедрой инструментальной диагностики, ФГБОУ ВО “ВГМУ им Н.Н. Бурденко” Минздрава России, Воронеж. <https://orcid.org/0000-0002-8421-3411>

Скачков Артур Андреевич – специалист по машинному обучению ООО “Объединенное IT пространство”, Липецк. <https://orcid.org/0009-0002-6072-8143>

Кульнева Таисия Владимировна – заместитель директора по экспертной работе в области медицинской визуализации АНО ДПО “Институт Эксперт”, Воронеж.

Измалков Дмитрий Витальевич – заведующий кафедрой лучевой диагностики АНО ДПО “Институт Эксперт”, Воронеж.

Андрienко Елизавета Андреевна – специалист общего отдела АНО ДПО “Институт Эксперт”, Воронеж.

Иванников Михаил Евгеньевич – ординатор по специальности “рентгенология” ФГБУ “Российский научный центр рентгено-радиологии” Минздрава России, Москва. <https://orcid.org/0009-0007-0407-0953>

Contact*: Mikhail E. Ivannikov – e-mail: ivannikovmichail@gmail.com

Andrey S. Rusakov – Founder and CEO of Remedy Logic, New York.

Vladislav V. Tumko – Chief Operating Officer of Remedy Logic, New York.

Ruslan S. Sarbaev – General Director of Decision Support Systems LLC, Cheboksary.

Natalia A. Uspenskaya – Computer Vision Team Leader at Remedy Logic, New York.

Nikolay V. Nudnov – Doct. of Sci. (Med.), Professor, Deputy Director for Scientific Work, Head of the Research Department for Complex Diagnostics of Diseases and Radiotherapy, Russian Scientific Center of Roentgenoradiology of the Ministry of Healthcare of the Russian Federation, Moscow. <https://orcid.org/0000-0001-5994-0468>

Grigory G. Karmazanovsky – Russian Academy of Sciences (RAS) Full Member, Doct. of Sci. (Med.), Professor, Head of the Diagnostic Radiology Department at the A.V. Vishnevsky National Medical Research Center of Surgery; Professor of radiology department, Pirogov Russian national research medical university, Moscow. <https://orcid.org/0000-0002-9357-0998>

Andrey V. Korobov – Director of Autonomous non-profit organization of additional professional education “Expert Institute”, Voronezh.

Lilia A. Titova – Doct. of Sci. (Med.), Head of the Department of Instrumental Diagnostics, Voronezh State Medical University named after N.N. Burdenko, Voronezh. <https://orcid.org/0000-0002-8421-3411>

Artur A. Skachkov – machine learning expert of LLC “United IT Space”, Lipetsk. <https://orcid.org/0009-0002-6072-8143>

Taisiya V. Kulneva – Deputy Director for Expertise in Medical Imaging of Autonomous non-profit organization of additional professional education “Expert Institute”, Voronezh.

Dmitry V. Izmalkov – Head of the Department of Radiologic Diagnostics of Autonomous non-profit organization of additional professional education “Expert Institute”, Voronezh.

Elizaveta A. Andrienko – general specialist of Autonomous non-profit organization of additional professional education “Expert Institute”, Voronezh.

Mikhail E. Ivannikov – resident in the specialty of “radiology”, Russian Scientific Center of Roentgenoradiology of the Ministry of Healthcare of the Russian Federation, Moscow. <https://orcid.org/0009-0007-0407-0953>