

Идентификация увеличения желудочковой системы головного мозга при помощи машинного обучения

Аннотация

Глубокое обучение – это активно развивающаяся технология, являющаяся частью более широкой области – машинного обучения, использование которого в медицине с каждым годом все интенсивнее исследуется во всем мире. Оценка работы алгоритма по выявлению расширения желудочковой системы головного мозга проводилась на примере 200 серий цифровых изображений МРТ головного мозга, выполненных в Т2-взвешенном режиме в аксиальной плоскости. Имеющиеся цифровые данные делились на три части: 1) тренировочную – для обучения; 2) валидационную – для остановки обучения и выбора модели по наилучшему параметру; 3) тестовую – для подсчета метрик качества модели. Показатели точности предсказания искомого отклонения от нормы – увеличения желудочковой системы – составили 97,5 %. Показатели чувствительности составили 96,3 % и специфичности – 98,1 % соответственно.

Введение

Гидроцефалия – это широкое понятие, включающее в себя ряд патологических состояний, которые характеризуются избыточным скоплением спинномозговой жидкости в ликвороносных путях головного мозга [1].

Клиническая картина обусловлена видом и стадией гидроцефалии. Специфические симптомы, такие как увеличение окружности головы, выбухание родничка, расширение вен скальпа, отставание в развитии от сверстников, характерны для детей раннего возраста и чаще всего ассоциированы с врожденной гидроцефалией. У детей старше двух лет, подростков и взрослых клинические симптомы вариативны и неспецифичны. Нормотензивная гидроцефалия, диагностируемая у лиц старшей (> 65 лет) возрастной группы, имеет характерную триаду симптомов, описанных S. Nakim и R.D. Adams в 1965 году: нарушение походки, недержание мочи, снижение когнитивных функций [2]. Перечисленные обстоятельства не позволяют диагностировать гидроцефалию, основываясь исключительно на клинической картине.

Совокупные данные по встречаемости синдрома гидроцефалии отмечены в единичных когортных исследованиях. S.C. Birg с соавторами приводят показатель заболеваемости гидроцефалией у взрослых – 17 случаев на 100 000 населения в год, что значительно ниже, чем заболеваемость гидроцефалией в детском возрасте: 82 случая на 100 000 новорожденных [3]. Гидроцефалия детского возраста – одна из самых распространенных патологий в педиатрической нейрохирургической практике; заболеваемость варьирует от 68 случаев на 100 000 новорожденных в Северной Америке до 145 и 316 случаев на 100 000 новорожденных в странах Африки и Латинской Америки соответственно [4].

Чаще в литературе встречаются эпидемиологические данные по отдельным нозологическим формам. Так, нормотензивная гидроцефалия (НТГ) встречается в 10...22 случаях на 100 000 населения старше 50 лет [5], [6]; по данным разных авторов, НТГ выявляется у 6...10 % больных с деменцией.

Вторичная гидроцефалия, которая до 62 % случаев сопутствует различным новообразованиям головного мозга, может быть обусловлена как прямым блоком тока ликвора (обструктивная), так и являться следствием гиперпродукции ликвора в ответ на изменение его химического состава на фоне роста опухоли. При опухолях задней черепной ямки у детей ее встречаемость составляет 70...90 % [7]-[9]. При вестибулярной шванноме у взрослых показатель варьирует от 3,7 до 42 % [10]; при глиомах ствола головного мозга у взрослых наиболее частым первым симптомом является головная боль, которая чаще всего обусловлена наличием гидроцефалии [11].

Таким образом, гидроцефалия может являться как самостоятельным, порой трудно диагностируемым заболеванием, так и сопутствовать ряду патологических состояний, и является актуальной проблемой в современной клинической практике.

Даже при условии наличия «очевидных» симптомов диагноз гидроцефалии устанавливается по данным нейровизуализационных методов: МРТ и/или КТ (МСКТ) головного мозга. Основываясь на диагностических критериях, возможно определить внутреннюю, наружную и смешанные формы гидроцефалии. Однако анализ и интерпретация этих критериев требуют временных затрат специалистов в области лучевой диагностики.

Нередко клинические проявления отстают от морфологических изменений, выявляемых при томографических исследованиях, в связи с чем проведение МРТ-скрининга позволит выявить случаи отклонений в показателях желудочковой системы мозга, что, в свою очередь, позволит заподозрить и диагностировать гидроцефалию на ранних стадиях с минимальными клиническими проявлениями. Для выполнения таких исследований и одновременного снижения чрезмерной нагрузки на специалистов лучевой диагностики потребуются новые технологические разработки, позволяющие дифференцировать норму и патологию.

Машинное обучение сегодня активно используется в различных практических направлениях и отраслях науки. Машинное обучение определяется как набор методов, способных автоматически выделять полезные паттерны в данных, чтобы в последующем использовать эти вновь открытые паттерны для предсказаний на новых данных или принятия решений в условиях неопределенности [12].

С тех пор как в 2012 году при помощи глубокого обучения (раздел машинного обучения) был достигнут существенный прогресс в области компьютерного зрения, глубокие нейронные сети, и в частности сверточные нейронные сети, стали основным методом решения задач детекции, сегментации и классификации изображений [13]. Возможности применения этих методов в рамках решения медицинских диагностических задач активно изучаются в настоящее время. Недавно опубликованные исследования о применении алгоритмов глубокого обучения в области лучевой диагностики демонстрируют многообещающие результаты в приложении к заданию определения и классификации патологии [14]-[19].

Наряду с отличными результатами, методы глубокого обучения имеют некоторые ограничения. Существенным препятствием для активного распространения технологии в медицинской диагностике является зависимость от наличия большого объема качественно подготовленных данных и сложности интерпретации глубоких нейронных сетей. Тем не менее для решения проблемы малых наборов данных может быть использована техника «переноса обучения» [20]: нейронная сеть сначала обучается на большом наборе данных, не связанном со специфической задачей (например, набор из нескольких миллионов обычных фотографий ImageNet) [21], а далее эта сеть более тонко настраивается под специальную задачу, например классификацию опухоли на медицинских изображениях.

Целью данного исследования явилась оценка возможностей применения глубокого обучения для определения расшире-

ния супратенториальной части желудочковой системы мозга на основе аксиальных проекций Т2 – взвешенных МРТ-изображений.

Материалы и методы

Постановка задачи

Диагностика расширения желудочковой системы головного мозга по МРТ-изображениям является задачей бинарной классификации. В данном случае МРТ-сканы были разделены на две группы: группа с характерными для расширения желудочковой системы головного мозга признаками (класс – «патология») и группа без таковых признаков (класс – «норма»). Для тренировки классификатора через алгоритм машинного обучения многократно пропускается набор обучающих данных, после чего качество модели тестируется на отложенной выборке данных. Итогом этого процесса в идеальном случае должна стать модель, способная к обобщению, т. е. предсказывающая правильные классы для изображений, не использовавшихся в процессе обучения. В рамках данного исследования задача бинарной классификации по МРТ-изображениям решалась с применением методики машинного обучения «с учителем».

Сбор, разметка и предварительная обработка данных

Сбор данных осуществлялся врачом лучевой диагностики с субспециализацией в МРТ-диагностике и опытом работы не менее 5 лет. Данные были собраны ретроспективно, в рамках одного медицинского центра. Было отобрано 200 серий DICOM МРТ-снимков головного мозга. МРТ-исследования проводились на аппарате «Toshiba Vantage Titan 1,5 T» («Toshiba Corporation», Япония). Анализируемые серии были представлены аксиальными проекциями цифровых изображений, выполненных в Т2-режиме сканирования в формате DICOM. Далее врачом-специалистом извлекались аксиальные срезы на уровне боковых желудочков и на уровне третьего желудочка. Эти изображения были анонимизированы, конвертированы в формат JPEG и далее использованы при проведении экспериментов.

Изображения были разделены на две группы: «норма» – 100 серий и «патология» – 100 серий. В связи с тем, что внутренняя гидроцефалия с клиническими позиций является более значимым состоянием, так как может представлять потенциальный риск для жизни больного, в группу «патология» отбирались изображения согласно МРТ-критериям расширения супратенториальной части желудочковой системы мозга: увеличение третьего желудочка (> 6 мм), увеличение размеров боковых желудочков на уровне их тел (> 18 мм). В исследование включались МРТ-снимки пациентов в возрасте от 18 до 80 лет, не имевших какой-либо дополнительной патологии головного мозга. Эпидемиологические характеристики приведены в табл. 1, исследования были проведены в период с ноября 2016 года по декабрь 2018 года.

Таблица 1

Основные характеристики исследуемых групп

Параметр	Группа	
	Норма	Патология
Возраст, лет	$33,8 \pm 12,9$	$66,3 \pm 11,5$
Мужчины	21	36
Женщины	79	64

Для обучения и оценки обобщающей способности модели выборки были разделены на тренировочный, валидационный и тестовый наборы. Такое разделение позволяет обучать модель на тренировочном наборе данных, проверять качество модели и изменять различные настройки для улучшения моде-

ли, используя валидационный набор, а затем оценивать качество после завершения всех экспериментов на тестовом наборе данных, который не использовался ни в процессе обучения, ни в процессе настройки модели.

Тренировка модели

Учитывая успешные результаты использования методологии «переноса обучения» в различных задачах области компьютерного зрения в применении к обычным изображениям и, в ряде случаев, к медицинским изображениям, в качестве предобученной сети в исследовании была использована модель архитектуры ResNet34 из набора предобученных моделей фреймворка fastai (<https://github.com/fastai/fastai/tree/master/fastai>). Кроме качества модели, важное значение имела также ее интерпретируемость, поэтому также была исследована область изображения, которая имела наибольшее значение при обучении модели. Для этой цели использовалась методика Gradient-weighted Class Activation Mapping (Grad-CAM) [22]. Данный метод использует карты активации последнего сверточного слоя, взвешивая их при помощи градиента предсываемого класса.

Техническое оснащение

Работа проводилась на сервере, оснащенном процессором «Intel Core i7-7740X» 4,30 ГГц («Intel Corporation», США), двумя видеокартами «GeForce GTX 1080» («Nvidia Corporation», США) по 8 ГБ видеопамяти и оперативной памятью 32 ГБ под управлением операционной системы «Linux» («Ubuntu 16.04.5 LTS»). В написании программного кода использовался язык программирования Python 3.7, а также фреймворки Pytorch 1.0 и fastai v1.0.39.

Статистическую обработку проводили посредством программы «IBM SPSS Statistics 19.0» («IBM Corporation», США). Описательные характеристики показателей желудочковой системы представлены в виде средней и одного стандартного отклонения средней ($M \pm SD$). Точность, чувствительность и специфичность метода, метрики качества и доверительные интервалы (указанные в скобках) рассчитывались с использованием статистических методов бутстрэпинга.

Результаты и обсуждение

Анализ показателей размеров желудочковой системы в группах продемонстрировал следующие результаты. В группе «норма» ширина боковых желудочков варьировалась от 7,8 до 17,4 мм, ширина третьего желудочка – от 1,2 до 4,8 мм. В группе «патология» ширина боковых желудочков варьировалась от 18,2 до 109,6 мм, ширина третьего желудочка – от 6,0 до 38,3 мм. Значения средних величин приведены в табл. 2.

Таблица 2

Основные показатели исследуемых групп

Параметр	Группа		Достоверность различий
	Норма	Патология	
Возраст, лет	$33,8 \pm 12,9$	$66,3 \pm 11,5$	$p < 0,05$
Ширина боковых желудочков, мм	$12,5 \pm 2,2$	$25,4 \pm 10,9$	$p < 0,05$
Ширина третьего желудочка, мм	$2,9 \pm 0,8$	$9,1 \pm 3,8$	$p < 0,05$

* Параметры указаны в виде средней величины и стандартного отклонения средней ($M \pm SD$).

Средние показатели точности предсказания искомого отклонения от нормы – увеличения желудочковой системы – составили 97,5 % (96,4...98,2 %), показатели чувствительности – 96,3 % (94,8...97,7 %) и специфичности – 98,1 % (96,2...99,1 %). Следует отметить, что в текущей работе использовали метод ранней остановки, выбор наилучшей модели по критерию наименьшего показателя на валидации. Важным фактом является то, что модель в качестве предсказания дает не само значение

класса («норма» или «патология»), а вероятность этого класса. Зная, что для задач диагностики в медицине в одних случаях наиболее важным является показатель чувствительности, а в других – специфичности и, учитывая предсказания в виде вероятностей, есть возможность более тонкой настройки порога, по которому предсказывается класс; таким образом, классификатор может быть настроен под нужды конкретной диагностической задачи.

Для интерпретации результатов модели была использована методика Grad-CAM. *Рис. 1* демонстрирует зоны считываения информации алгоритмом с изображения. Как видно, фокусировка «компьютерного зрения» направлена на желудочковую систему мозга.

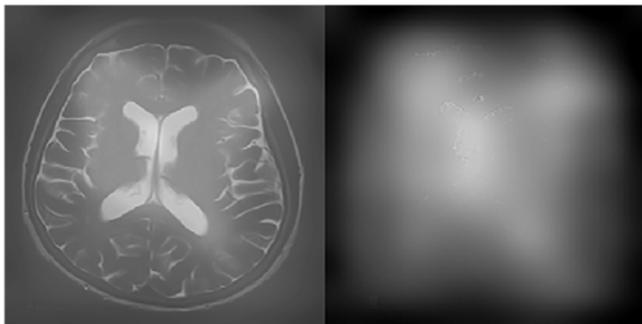


Рис. 1. Фокусировка компьютерного зрения при оценке МРТ-изображения

Мнения относительно сплошных скрининговых нейровизуализационных исследований головного мозга остаются противоречивыми. Авторы указывают на сомнительную экономическую выгоду: высокие затраты на проводимые исследования и низкое количество «случайных находок» (СН), которые требуют лечения при сплошном исследовании головного мозга здоровых людей [23]. Стоит отметить, что за десятилетний период количество СН увеличилось: так, в исследовании 2007 года M.W. Vernooy с соавторами [24] указывают на показатель 7,2 % среди 2 000 здоровых людей в возрасте от 45,7 до 96,7 лет. В 2017 году A. Brugulat-Serrat с соавторами [25] указывает на значение 27 % среди 575 здоровых волонтеров в возрасте от 45 до 75 лет. Не исключено, что данный факт был связан с мощностью магнитно-резонансного томографа: в первом случае – 1,5 Тл, во втором – 3 Тл. Несмотря на возросшее количество СН, количество лиц, которым требуется лечение, остается одинаково крайне низким (0,1...0,3 %). Однако в указанных исследованиях сканирование проводили среди добровольцев, не имеющих каких-либо жалоб. По нашему мнению, проведение скрининговых томографических исследований может быть оправдано при осмотре пациентов с неспецифичными жалобами, под маской которых могут протекать отдельные патологические состояния головного мозга с развитием увеличения желудочковой системы головного мозга, к примеру, возрастные больные с деменцией или дети и молодые люди с частыми головными болями.

Поскольку при указанных исследованиях получается довольно объемный массив данных, требующий обязательной интерпретации специалистом, такие подходы пока не получают широкого распространения. Внедрение в практику применения алгоритмов машинного обучения с возможностью определения различных патологических паттернов позволит выявлять отклонения от нормы с минимальным участием человека. В нашей серии мы сфокусировались на выявлении внутренней гидроцефалии. Как было указано во введении, гидроцефалия – достаточно широко встречаемое явление в неврологической и нейрохирургической практике, которое может являться как самостоятельным заболеванием, так и иметь вторичный характер. Подходы ранней верификации увеличения желудочковой системы мозга на фоне неспецифических жалоб актуальны во всех возрастных группах. В детской практике это имеет значение при опухолях головного мозга; в среднем воз-

расте МРТ-скрининг позволяет выявлять разнообразные отклонения в размерах желудочковой системы как опухолевой, так и неопухолевой природы. У старшей возрастной группы можно будет выявить категорию лиц, находящихся в потенциальной группе риска развития нормотензивной гидроцефалии.

Методика скринингового исследования, дополненная применением методов машинного обучения, позволит снизить нагрузку на специалистов лучевой диагностики и повысить доступность МРТ-диагностики патологии головного мозга.

Как указывалось выше, количество работ, посвященных изучению применения машинного обучения в медицине, с каждым годом растет [26]. Гидроцефалию в подобных исследованиях рассматривают с позиции предсказания исходов хирургического лечения, основываясь на анализе «больших массивов данных» о пациенте [27], [28].

В завершение следует отметить, что данное исследование в качестве основной своей задачи рассматривало возможность использования подходов машинного обучения в анализе данных МРТ-изображений. В качестве анализируемого отклонения от нормы был выбран достаточно простой для специалиста лучевой диагностики параметр – наличие или отсутствие расширения желудочковой системы мозга. Однако для начала оценки работы алгоритма и требовалась относительно простая задача, которая позволила сделать выводы об уместности такого подхода в условиях имеющихся диагностических и вычислительных мощностей. Практическая значимость алгоритма указана в качестве одной из возможных сфер его применения на текущем этапе. Увеличение обучающей выборки с разнообразной неврологической и нейрохирургической патологией позволит усложнить задачи, решаемые при помощи указанного подхода.

Заключение

В ходе работы был собран набор данных МРТ-изображений головного мозга с наличием и без признаков увеличения супратенториальной части желудочковой системы головного мозга. Эти данные были собраны и размечены врачами-специалистами, и таким образом было получено высокое качество разметки. Учитывая небольшой размер выборки, исследование было нацелено на изучение возможностей методологии «переноса обучения» в решении задачи классификации на «норму» и «патологию» по МРТ-изображениям. В работе была использована предобученная на изображениях из набора данных ImageNet сверточная нейронная сеть архитектуры ResNet34. После «тонкой настройки» и дообучения этой нейронной сети с использованием набора МРТ-изображений головного мозга были получены обнадеживающие результаты. Достигнуты следующие показатели качества модели при предсказании: точность – 97,5 %, чувствительность – 96,3 % и специфичность – 98,1 %. В проведенной работе впервые демонстрируются возможности использования глубокого обучения и методики «переноса обучения» в решении задачи диагностики синдрома внутренней гидроцефалии по данным МРТ головного мозга; при этом высокое качество модели достигается при относительно небольшом размере обучающих данных. Дальнейшее расширение обучающих выборок и набора патологий в перспективе позволит решать задачи быстрого поиска патологических состояний по данным нейровизуализационных методик обследования.

Список литературы:

1. Rekate H.L. The definition and classification of hydrocephalus: A personal recommendation to stimulate debate // Cerebrospinal Fluid Research. 2008. Vol. 5. № 1. PP. 1-7.
2. Adams R., Fisher C., Hakim S. et al. Symptomatic occult hydrocephalus with normal cerebrospinal fluid pressure: A treatable syndrome // New England Journal of Medicine. 1965. Vol. 273. № 3. PP. 117-126.
3. Bir S.C., Patra D.P., Maiti T.K. et al. Epidemiology of adult-onset hydrocephalus: Institutional experience with 2001 patients // Neurosurgical Focus. 2016. Vol. 41. № 3. E5.

4. Dewan M.C., Rattani A., Mekary R. et al. Global hydrocephalus epidemiology and incidence: Systematic review and meta-analysis // Journal of Neurosurgery. 2018. Vol. 130. № 4. PP. 1065-1079.
5. Martin-Laez R., Caballero-Arzapalo H., Lopez-Menendez L.A. et al. Epidemiology of idiopathic normal pressure hydrocephalus: A systematic review of the literature // World Neurosurgery. 2015. Vol. 84. № 6. PP. 2002-2009.
6. Klassen B.T., Ahlskog J.E. Normal pressure hydrocephalus: How often does the diagnosis hold water? // Neurology. 2011. Vol. 77. № 12. PP. 1119-1125.
7. Lam S., Reddy G.D., Lin Y. et al. Management of hydrocephalus in children with posterior fossa tumors // Surgical Neurology International. 2015. Vol. 6. Suppl. 11. PP. 346-348.
8. Dorner L., Fritsch M.J., Stark A.M. et al. Posterior fossa tumors in children: How long does it take to establish the diagnosis? // Child's Nervous System. 2007. Vol. 23. № 8. PP. 887-890.
9. Prasad K.S.V., Ravi D., Pallikonda V. et al. Clinicopathological study of pediatric posterior fossa tumors // Journal of Pediatric Neurosciences. 2017. Vol. 12. № 3. PP. 245-250.
10. Prabhuraj A., Sadashiva N., Kumar S. et al. Hydrocephalus associated with large vestibular schwannoma: Management options and factors predicting requirement of cerebrospinal fluid diversion after primary surgery // Journal of Neurosciences in Rural Practice. 2017. Vol. 8. Suppl. 1. PP. 27-32.
11. Hu J., Western S., Kesari S. Brainstem glioma in adults // Frontiers in Oncology. 2016. Vol. 6. P. 180.
12. Murphy K.P. Machine learning: A probabilistic perspective. – The MIT Press; Illustrated edition, 2012. 1104 p.
13. Krizhevsky A., Sutskever I., Hinton G.E. Imagenet classification with deep convolutional neural networks // Advances in neural information processing systems. 2012. Vol. 25. PP. 1097-1105.
14. Kohli M., Prevedello L.M., Filice R.W. et al. Implementing machine learning in radiology practice and research // American Journal of Roentgenology. 2017. Vol. 208. № 4. PP. 754-760.
15. Erickson B.J., Korfiatis P., Akkus Z. et al. Machine learning for medical imaging // RadioGraphics. 2017. Vol. 37. № 2. PP. 505-515.
16. Lakhani P., Sundaram B. Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks // Radiology. 2017. Vol. 284. № 2. PP. 574-582.
17. Zhang N., Yang G., Gao Z. et al. Deep learning for diagnosis of chronic myocardial infarction on nonenhanced cardiac cine MRI // Radiology. 2019. Vol. 291. № 3. PP. 606-617.
18. Soffer S., Ben-Cohen A., Shimon O. et al. Convolutional neural networks for radiologic images: A radiologist's guide // Radiology. 2019. Vol. 290. № 3. PP. 590-606.
19. Cicero M., Bilbily A., Colak E. et al. Training and validating a deep convolutional neural network for computer-aided detection and classification of abnormalities on frontal chest radiographs // Investigative Radiology. 2017. Vol. 52. № 5. PP. 281-287.
20. Yosinski J., Clune J., Bengio Y. et al. How transferable are features in deep neural networks? / NIPS'14: Proceedings of the 27th International Conference on Neural Information Processing Systems. 2014. Vol. 2. PP. 3320-3328.
21. Russakovsky O., Deng J., Su H. et al. Imagenet large scale visual recognition challenge // International Journal of Computer Vision. 2015. Vol. 115. № 3. PP. 211-252.
22. Selvaraju R.R., Cogswell M., Das A. et al. Grad-cam: Visual explanations from deep networks via gradient-based localization / Proceedings of the IEEE International Conference on Computer Vision. 2017. PP. 618-626.
23. Komotar R.J., Starke R.M., Connolly E.S. Brain magnetic resonance imaging scans for asymptomatic patients: Role in medical screening // Mayo Clinic Proceedings. 2008. Vol. 83. № 5. PP. 563-565.
24. Vernooij M.W., Ikram M.A., Tanghe H.L. et al. Incidental findings on brain MRI in the general population // New England Journal of Medicine. 2007. Vol. 357. № 18. PP. 1821-1828.
25. Brugulat-Serrat A., Rojas S., Bargallo N. et al. Incidental findings on brain MRI of cognitively normal first-degree descendants of patients with Alzheimer's disease: A cross-sectional analysis from the alfa (Alzheimer and families) project // BMJ open. 2017. Vol. 7. № 3. e013215.
26. Celikci E. A Systematic review on machine learning in neurosurgery: The future of decision-making in patient care // Turkish Neurosurgery. 2018. Vol. 28. № 2. PP. 167-173.
27. Azimi P., Mohammadi H.R. Predicting endoscopic third ventriculostomy success in childhood hydrocephalus: An artificial neural network analysis // Journal of Neurosurgery Pediatrics. 2014. Vol. 13. PP. 426-432.
28. Habibi Z., Ertiae A., Nikdad M.S. et al. Predicting ventriculoperitoneal shunt infection in children with hydrocephalus using artificial neural network // Child's Nervous System. 2016. Vol. 32. PP. 2143-2151.

Сергей Валерьевич Мишинов,
канд. мед. наук, ст. научный сотрудник,
врач-нейрохирург,
отделение нейрохирургии № 1,
ФГБУ «НИИ Травматологии
и ортопедии им. Я.Л. Цивяни» МЗ РФ,
г. Новосибирск,
Алексей Иванович Демянчук,
научный сотрудник,
Hannover Medical School, MHH.
Clinic of Laryngology, Rhinology and Otology,
Hannover, Germany,
Екатерина Владимировна Пушкина,
врач лучевой диагностики,
АНО «Клиника НИИТО»,
Вячеслав Владимирович Ступак,
д-р мед. наук, профессор, начальник,
научно-исследовательское
отделение нейрохирургии,
ФГБУ «НИИ Травматологии
и ортопедии им. Я.Л. Цивяни» МЗ РФ,
Тимур Маратович Фатыхов,
старший дата-аналитик,
ЗАО «Золотая корона»,
Николай Евгеньевич Русских,
начальник,
отдел машинного обучения,
Дмитрий Николаевич Штокало,
директор,
ООО «АкадемДжин»,
г. Новосибирск,
e-mail: smishinov@yandex.ru