

УДК 004.588

ПОДХОДЫ К ПОСТРОЕНИЮ НЕЙРОСЕТИ ДЛЯ БИНАРНОЙ КЛАССИФИКАЦИИ РЕНТГЕНОГРАММ

© 2020 г. Р. Ш. Минязев¹, А. А. Румянцев¹, А. А. Баев^{2, *}, Т. Д. Баева²

¹Федеральное государственное бюджетное образовательное учреждение высшего образования Казанский национальный исследовательский технический университет имени А.Н. Туполева-КАИ, Казань, Россия

²Федеральное государственное бюджетное образовательное учреждение высшего образования “Поволжский государственный технологический университет”, Йошкар-Ола, Россия

*E-mail: kruzenshteyn@yandex.ru

Поступила в редакцию 15.07.2020 г.

После доработки 10.08.2020 г.

Принята к публикации 26.08.2020 г.

Предлагается использовать машинный анализ рентгеновских снимков с помощью нейросети, выполняющей бинарную классификацию. Рассматриваются архитектура и параметры разработанной глубокой сверточной нейросети, предлагаются подходы для повышения качества ее работы за счет вычисления энтропии фрагментов снимка и использования ансамбля обученных сетей.

DOI: 10.31857/S0367676520120285

ВВЕДЕНИЕ

В современной медицинской диагностике одним из основных источников информации являются рентгеновские изображения, в частности, флюорографические снимки легких. При этом одна из основных проблем быстрая и качественная обработка получаемых снимков. Ручная обработка требует привлечения высококвалифицированных врачей-рентгенологов, дефицит которых ощущается повсеместно [1]. Для сокращения объема работ, предлагается разработать веб-сервис [2] в основе работы которого лежит обученная глубокая сверточная нейросеть [3], которая осуществляет машинную бинарную классификацию рентгеновских изображений. При этом главная сложность — обеспечить приемлемый уровень качества работы нейросети с минимальным количеством ошибок первого и второго рода [4]. Для оценки качества работы обученной нейросети вычисляется общепринятая метрика AUC — площадь, ограниченная ROC-кривой (англ. receiver operating characteristic — рабочая характеристика приёмника) и осью доли ложных положительных классификаций [5]. Также оценивались такие критерии качества работы нейросети как количество ложно положительных решений, принятых сетью (TF) и ложно отрицательных (FN). Оба параметра оценивались по классу патология, так как для рассматриваемой задачи наиболее важно не пропустить больного пациента как здорового.

В последние несколько лет появилось много исследовательских проектов по машинному анализу медицинских рентгеновских снимков с использованием нейросетей. К числу наиболее известных можно отнести: проект АО Радиоконпаний Вектор “ФтизисБиоМед” совместно с выпускниками МФТИ [6], максимальный AUC = 0.736, выделяет патологии разного вида; проект “CheXNet” Стенфордского университета совместно с инженерами компании Google, который способен ставить диагнозы пневмонии по рентгенограммам легких, максимальный AUC = 0.937, выделяет 14 видов заболеваний в легких; проект по использованию нейросети для диагностики туберкулеза, максимальный AUC = 0.99, выделяет только заболевания туберкулезом [7]. Бурное развитие получили информационные платформы для поддержки цифровой медицинской диагностики с использованием нейросетей: iPavlov — международный проект по созданию искусственного интеллекта, реализуемый лабораторией нейронных систем и глубокого обучения МФТИ; платформа “Третье мнение” для распознавания медицинских изображений разного типа: рентгенограммы легких, снимки крови и костного мозга, глазного дна, гистологические изображения, маммограммы, ОПТГ, КТ, МРТ, дерматологические изображения.

Суть задачи определения патологий на рентгеновских снимках легких состоит в определении факта наличия патологий, локализации расположения на снимке и дальнейшем уточнении вида

Таблица 1. Параметры архитектуры и обучения

| Параметр | Значения |
|----------------------------------------|----------------------------------------------------|
| Архитектура | 1A, 1B, 1C, 1A2B, 1A2B1C, 2A10B, 8A20B4C, 10A20B9C |
| Размер батча | 4, 8, 16, 32, 64 |
| Функция пулинга | MaxPooling, AvgPooling |
| Функция нормализации | Softmax, LogSoftmax |
| Алгоритм обучения | Adam, Adadelata, Adagrad, GradientDescent |
| Скорость обучения | 0.2, 0.02, 0.002, 0.0002, 0.00002 |
| Коэффициент снижения скорости обучения | 0.85, 0.7, 0.25, 0.05, 0.001 |
| Функция скорости обучения | Exp, NatExp, Inverse, Cosine, LinearCosine |

патологий. Для простоты реализации нашей нейросети (single-label) с целью принципиальной оценки ее эффективности уточнение патологий в архитектуре нейросети не реализовано, ограничиваясь 2 классами — норма или патология. Для обучения нейросети использовалась классифицированная заранее врачом-рентгенологом база медицинских снимков, полученных на цифровых рентгеновских аппаратах. Были выделены два класса изображений. Объем базы 22000 снимков — 11000 снимков здоровых людей, 11000 снимков с различными заболеваниями. Из них для обучения использовалось 20000 снимков, для тестирования качества работы нейросети использовались оставшиеся 2000 снимков. Таким образом, соотношение изображений распознаваемых классов в обучающей и тестирующей базе составило 1 к 1.

АРХИТЕКТУРА РАЗРАБАТЫВАЕМОЙ СВЕРТОЧНОЙ НЕЙРОСЕТИ

За основу для построения глубокой сверточной нейросети была принята конфигурация нейросети Inception ResNet [8]. В которой совмещаются архитектуры построения “широких” нейросетей Inception и “глубоких” нейросетей ResNet. На международных соревнованиях по классификации изображений Imagenet Recognition Challenge [9] первое место последние несколько лет занимают именно глубокие нейронные сверточные сети, построенные по данной архитектуре [10]. Сеть была реализована в программном модуле на языке Python с использованием библиотеки машинного обучения Tensorflow от компании Google.

В целом архитектура сверточной нейронной сети состоит из слоев: свертки (конволюции) — выделение существенных признаков на изображении; макспулинга — уменьшение размерности изображения за счет объединения соседних пикселей; перцептронов — обычная многослойная нейросеть. Первые два типа слоев, чередуясь между собой, формируют вектор признаков на вход многослойного перцептрона. Серьезной проблемой является

выбор подходящего количества слоев в глубину и ширину в рамках базовых блоков нейросети. Также при построении архитектуры нейросети важны вопросы выбора размера батча для обучения, выбор функции нормализации, выбор функции пулинга, выбор алгоритма и параметров обучения, выбор способа модификации базы используемых снимков для увеличения их количества и разнообразия.

Было проведено большое количество итерационных вычислительных экспериментов по обучению и тестированию нейросети для разного набора параметров ее архитектуры. Основные исследованные параметры архитектуры нейросети представлены в табл. 1, в результате выбраны наиболее эффективные. Построенная нейросеть получила название “эталон”. Результаты итоговых экспериментов по тестированию эффективности представлены в табл. 2.

ПРЕДОБРАБОТКА DATA AUGMENTATION

Значительную роль в формировании эффективной нейросетевой модели, как показали исследования, играет правильно подготовленная обучающая база снимков. В виду ограниченности размера доступной медицинской базы данных, статистической редкости, малого разнообразия или размера отдельных патологий возникает необходимость искусственного повышения разнообразия выборки рентгенограмм с целью выявления достаточных для эффективной классификации нейронной сетью патологических признаков. Одно из перспективных направлений для подготовки базы снимков для обучения — Data Augmentation, суть которого в трансформации уже имеющейся выборки изображений для увеличения разнообразия классифицируемых признаков. Основные методы преобразования, используемые в рамках данного подхода:

— вращение, в частности горизонтальное и вертикальное зеркалирование;

Таблица 2. Тестирование базы снимков 2000 штук с соотношением 1/1

| | Эталон | Data Augmentation | 32 × 32 (сжатие) | 4 × 4 E2 | 16 × 16 E3 | Ансамбль |
|------------------|--------|-------------------|------------------|-------------|---------------|----------|
| TP (Sensitivity) | 85.40 | 87.70 | 87.50 | 86.30 | 91.1 | 92.1 |
| FN | 14.60 | 12.30 | 12.50 | 13.70 | 8.9 | 7.9 |
| FP | 17.20 | 19.10 | 27.00 | 20.20 | 24.1 | 18.6 |
| TN (Specificity) | 82.80 | 80.90 | 73.00 | 79.80 | 75.9 | 81.4 |
| PPV | 83.24 | 82.12 | 76.42 | 81.03 | 79.1 | 83.2 |
| NPV | 85.01 | 86.80 | 85.38 | 85.35 | 89.5 | 91.2 |
| AUC | 0.914 | 0.918 | 0.882 | 0.905 | 0.913 | 0.953 |

– сжатие и растяжение (коэффициенты 0.8 и 1.5 соответственно), масштабирование.

Исходная выборка изображений объединяется с различными сочетаниями трансформаций и обученная на итоговой расширенной базе снимков сверточная нейронная сеть тестируется на исходной немодифицированной проверочной выборке рентгеновских изображений легких. При этом увеличение размера обучающей базы обуславливает необходимость использования дополнительных методов предобработки снимков с целью упрощения процесса обучения нейросети и ускорения вычислений [11]. В результате использования представленных подходов база снимков для обучения увеличилась до 80000. На увеличенной базе изображений было произведено обучение “эталонной” нейросети. Результаты итоговых экспериментов по тестированию эффективности представлены в табл. 2.

ВЫЧИСЛЕНИЕ ЭНТРОПИИ ФРАГМЕНТОВ ИЗОБРАЖЕНИЯ

Исходные снимки имеют стандартный медицинский формат DICOM и имеют разрешение 512×512 пикселей. Обучение нейросети на имеющейся базе снимков занимает достаточно много времени, для сокращения времени обучения сети было решено использовать подход вычисления энтропии [12] фрагментов изображения для сжатия – уменьшения разрешения исходных снимков и, как следствие, сокращения времени обучения модели.

Информационная энтропия – мера хаотичности информации, в том числе содержащейся в изображении. Для расчета энтропии существует множество методов, таких как формула Шеннона, ее модификации на основе градиентов и матрицы коинцидентности, а также ансамбли из данных алгоритмов.

В рамках представленного подхода все исходное изображение разделяется на фрагменты – окна. Рассматривались следующие варианты размеров окна: 128×128 , 64×64 , 32×32 и 16×16 (еди-

ниц). При их использовании формируется новое изображение уменьшенного размера – матрица элементов: 4×4 , 8×8 , 16×16 и 32×32 (пикселей) соответственно. Путем такой обработки из исходных снимков формировались новые уменьшенного размера. Каждый вариант размера окна анализировался отдельно для поиска наиболее эффективного.

Расчет элементов матрицы – пикселей нового уменьшенного изображения осуществляется проходом по всем пикселям исходного изображения в рамках текущего окна в ширину и высоту с приращением исходно нулевых элементов частоты распределения цветов для каждого соответствующего фрагмента в матрице энтропии. Затем массивы распределения цветов нормализуются, и по формуле Шеннона вычисляется энтропия в каждом из фрагментов. В результате каждый пиксель уменьшенного изображения – нормализованное по максимуму значение энтропии фрагмента, переведенное в цветовой RGB канал. При этом важно отметить, что числовое значение в интервале от 0 до 255 для каждого канала определялось своим алгоритмом подсчета энтропии: R – классический алгоритм Шеннона (E1), G – модифицированный градиентный алгоритм Шеннона второго порядка (E2), B – модифицированный алгоритм Шеннона – Aura Matrix (E3). На рис. 1 представлена визуализация карты энтропии снимка для случаев вычисления одиночной энтропии и использования описанного RGB ансамбля алгоритмов вычисления энтропии.

Описанный подход применялся для предобработки всей базы снимков 22000 изображений. Полученная в результате база обработанных снимков использовалась для обучения и тестирования нейросети аналогично процедуре, описанной в предыдущем разделе. Благодаря уменьшению разрешения снимков и исключения из архитектуры слоя макспулинга удалось значительно уменьшить время, затрачиваемое на обучение и тестирование нейросети.

В ходе проведенных экспериментов было установлено, что наиболее эффективно использовать

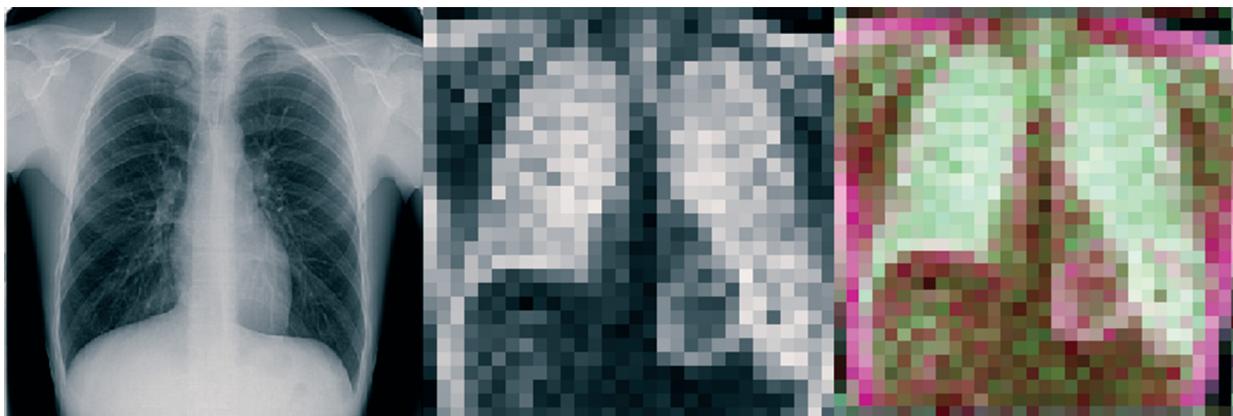


Рис. 1. Пример снимка и его карты одиночной энтропии и RGB ансамбля.

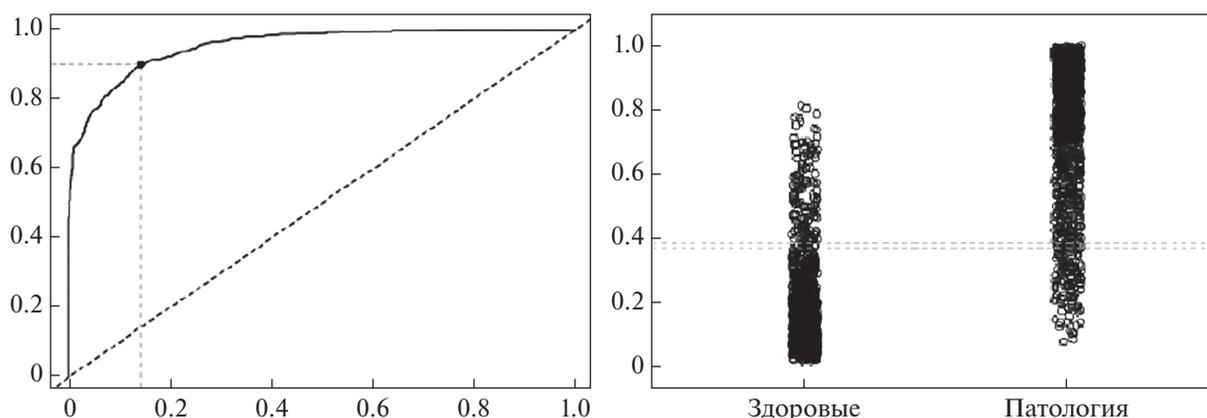


Рис. 2. ROC кривая для ансамбля и распределение вероятностей для классов.

окно размером 16×16 в результате применения которого получается уменьшенное изображение размером 32×32 пикселя. Ускорение по времени вычислений в этом случае составляет в три раза по сравнению с вариантом, когда использовались необработанные снимки размером 512×512 пикселей. Накладные расходы по времени, связанные с подготовкой – сжатием исходных снимков, не учитывались ввиду их незначительности. При этом качество работы, обученной на сжатых снимках нейросети по параметру AUC, оказалось практически равным значению, полученному на базе исходных необработанных снимков. Обученная на таких снимках нейросеть получила название “сжатие”. Результаты итоговых экспериментов по тестированию ее эффективности представлены в табл. 2.

АНСАМБЛЬ ОБУЧЕННЫХ НЕЙРОСЕТЕЙ

Проведенные эксперименты показали, что наибольшую эффективность по параметру AUC показывает “эталонная” нейросеть, обученная на

исходной необработанной базе снимков. Для повышения значения AUC было решено использовать ансамбль из нейросетей, обученных с использованием разных подходов к предобработке базы снимков: “эталон”, “сжатие”, “скользящее окно”. Суть подхода заключается в том, что очередной поступающий при тестировании снимок проходит через все обученные нейросети и берется среднее значение вероятности наличия патологии. Полученное в результате проведенного тестирования значений AUC представлено в табл. 2 под названием “ансамбль”. Оно оказалось максимальным. На рис. 2 показана ROC кривая для ансамбля и распределение вероятностей для рассматриваемых классов изображений: здоров, патология.

ОЦЕНКА РЕЗУЛЬТАТОВ

Для удобства сравнения наилучшие результаты всех проведенных экспериментов по тестированию нейросети для каждого рассмотренного в статье подхода были собраны в одну сводную таб-

лицу. Для оценки качества работы сети кроме основного критерия AUC использовались также значения метрик: TP – истинно положительные; FN – ложно отрицательные; FP – ложно положительные; TN – истинно отрицательные; PPV – positive predictive value, вероятность того что пациент, определенный как “больной”, действительно имеет заболевание; NPV – negative predictive value, вероятность того, что пациент, определенный как “здоровый”, действительно здоров.

По итогам работы на основе архитектуры InceptionResNet построена глубокая сверточная нейросеть для бинарной классификации рентгеновских изображений легких. Подобраны наиболее эффективные параметры конфигурации сети на макроуровне. Полученное достаточно высокое значение критерия AUC = 0.914, и оно превышает результаты, полученные в аналогичных исследовательских проектах. Использование для обучения нейросети увеличенной за счет модификации изображений DataAugmentation базы снимков позволяет незначительно увеличить это значение AUC = 0.918.

Предложенный подход к предобработке базы снимков с целью ускорения процесса обучения нейросети – уменьшение размерности снимков “сжатие” за счет вычисления энтропии фрагментов изображения и дальнейшей нормализации вычисленных значений с целью представления в RGB формат – оказался достаточно эффективен по критерию уменьшения времени обучения нейросети за счет значительного сжатия базы снимков для обучения (ускорение процесса обучения в три раза) полученное значение AUC = 0.882 сравнимо со значением, полученным для случая “эталонной” нейросети, использующей для обучения исходные снимки.

Предложенный подход к предобработке базы снимков с целью выявления скрытой информации о взаимосвязи между фрагментами изображения – вычисление энтропии фрагментов снимка методом “скользящего окна” с шагом в 1 пиксель – показал результаты сравнимые с “эталонной” нейросетью, наилучшее значения AUC = 0.913.

Наилучший результат показало использование ансамбля из рассмотренных нейросетей значение AUC = 0.953.

ЗАКЛЮЧЕНИЕ

Проведенное исследование подтверждает высокий потенциал применения подхода вычисления энтропии для сжатия изображений с сохранением пространственных признаков, необходимых для классификации нейронными сетями. Незначительная разность между наилучшими результатами для энтропии и исходных изображений озна-

чает большой процент покрытия патологических признаков на флюорографических изображениях, что позволяет в значительной мере использовать данный подход при многозначной классификации для ускорения вычислений и упрощения локализации патологий на карте признаков. Была выявлена эффективность алгоритмов расчета пространственной энтропии, а также сочетаний данных алгоритмов в ансамбле. В ходе проведенных экспериментов были определены наиболее эффективные параметры модели “эталонной” нейросети InceptionResNet: количество базовых блоков, размер batch, функции пулинга и нормализации, алгоритм и параметры обучения. Показано, что использование сжатия для базы изображений путем вычисления энтропии фрагментов слабо ухудшает качество обучения нейросети, но, значительно уменьшает вычислительную сложность процесса обучения. Использование для предобработки базы изображений подхода скользящего окна для вычисления энтропии фрагментов не дает повышения качества обучения нейросети. Использование ансамбля нейросетей, обученных на базах изображений, предобработанных различными подходами к вычислению энтропии, дает значимое повышение качества классификации снимков.

СПИСОК ЛИТЕРАТУРЫ

1. Доклад о состоянии здоровья населения и организации здравоохранения. М.: Министерство здравоохранения РФ, 2015. 161 с.
2. Минязев Р.Ш., Румянцев А.А., Дыганов С.А., Баев А.А. // Изв. РАН. Сер. физ. 2018. Т. 82. № 12. С. 1685; Minyazev R.Sh., Rumyantsev A.A., Dyganova S.A., Baev A.A. // Bull. Russ. Acad. Sci. Phys. 2018. V. 82. № 12. P. 1529.
3. Минязев Р.Ш., Румянцев А.А., Баев А.А., Баева Т.Д. // Изв. РАН. Сер. физ. 2019. Т. 83. № 12. С. 1655; Minyazev R.Sh., Rumyantsev A.A., Baev A.A., Baeva T.D. // Bull. Russ. Acad. Sci. Phys. 2019. V. 83. № 12. P. 1494.
4. Румянцев А.А., Минязев Р.Ш., Дыганов С.А. и др. // Вестн. КГТУ. 2018. Т. 21. № 8. С. 124.
5. Babu S. // Int. J. Innov. Res. Sci. Eng. Technol. 2015. V. 2. № 9. P. 596.
6. Klassen V.I., Safin A.A., Maltsev A.V. et al. // J. eHealth Technol. Appl. 2018. V. 16. № 1. P. 28.
7. Paras L., Sundaram B. // Radiology. 2017. V. 284. № 2. P. 574.
8. Szegedy C., Ioffe S., Vanhoucke V. // Proc. AAAI–17. (San Francisco, 2017). P. 4278.
9. He K., Zhang X., Ren S., Sun J. // Proc. CVPR 2016. (Las Vegas, 2016) P. 770.
10. Krizhevsky A., Sutskever I., Hinton G.E. // Proc. NIPS 2012. V. 25. (Lake Tahoe, 2012). P. 1090.
11. Ioffe S., Szegedy C. // Proc. 32nd ICML. V. 37. (Lille, 2015). P. 448.
12. Шеннон К. Работы по теории информации и кибернетике. М.: Изд-во иностранной литературы, 1963. 830 с.