

УДК 004.932.2

АНАЛИЗ БИОЛОГИЧЕСКИХ ОБЪЕКТОВ МЕТОДОМ КОМПЬЮТЕРНОЙ МИКРОСКОПИИ В ВИДИМОМ ДИАПАЗОНЕ ЭЛЕКТРОМАГНИТНОГО ИЗЛУЧЕНИЯ С ПРИМЕНЕНИЕМ НЕЙРОННЫХ СЕТЕЙ

В. Г. Никитаев¹, Н. Н. Тупицын², А. Н. Проничев¹, В. В. Дмитриева¹,
Е. В. Поляков¹, К. А. Либериc¹, М. С. Григорьева^{1,3}, А. Д. Палладина²

Рассмотрена междисциплинарная задача (физика, биология, медицина, информатика) распознавания клеток костного мозга для диагностики острых лейкозов и минимальной остаточной болезни. Рассмотрено применение нейронных сетей для определения классов клеток в микроскопическом изображении препарата костного мозга. Изображения формируются в проходящем свете. Представлена структура нейронной сети и выборки изображений для проведения экспериментального исследования. По результатам эксперимента уточнена структура нейронной сети и определены точностные характеристики разработанной системы.

Keywords: компьютерная микроскопия, распознавание изображений, нейронные сети.

Введение. Решение задачи автоматизации диагностики острых лейкозов требует применения междисциплинарного (физика, биология, медицина, информатика) подхода, учитывающего физические особенности формирования микроскопических изображений клеток крови в препаратах костного мозга. С точки зрения биологии и медицины необходимо учесть особенности структуры ядер клеток крови при патологических состояниях организма. Методы информатики – используются для обработки больших объемов данных и принятия диагностических решений в слабоформализованных задачах.

¹ НИЯУ МИФИ, 115409 Россия, Москва, Каширское ш., 31.

² Национальный медицинский исследовательский центр онкологии им. Н. Н. Блохина, 115478 Россия, Москва, Каширское шоссе, 23.

³ ФИАН, 119991 Россия, Москва, Ленинский пр-т, 53; e-mail: kaf46@mail.ru.

Среди прорывных направлений современной науки и техники одна из важнейших ролей принадлежит искусственному интеллекту (распознавание образов, базы знаний, экспертные системы) [1–3].

Анализ крови входит в число первостепенных задач диагностики различных заболеваний [4]. Широко используемым методом исследования для его выполнения является микроскопия в видимом диапазоне электромагнитного излучения. Типовая задача, решаемая врачом клинической лабораторной диагностики во время проведения этой процедуры, – классификация клеток крови.

Стандартный подход к решению задачи распознавания клеток крови на уровне компьютерной микроскопии предусматривает три этапа: предобработку, описание и классификацию изображений.

Каждый вид клеток может быть описан врачом отличительным набором признаков. Эффективное применение компьютерной техники требует формализации информативных признаков [5, 6].

Диагностика острых лейкозов базируется на морфологических, цитохимических и молекулярно-генетических методах исследования лейкоэмических клеток периферической крови и костного мозга. При этом микроскопический анализ клеток крови и костного мозга является обязательным.

Перспективным направлением в развитии методов диагностики острых лейкозов и минимальной остаточной болезни (МОБ) является разработка систем цифровой микроскопии с применением технологий искусственных нейронных сетей [7]. Оснащение отдаленных лечебно-профилактических учреждений (ЛПУ) компьютерной микроскопией в сочетании с искусственным интеллектом позволит обеспечить адекватный и своевременный анализ ответа на поддерживающую терапию, получаемую больным в ЛПУ по месту жительства. Важно, что именно на основании регулярных анализов МОБ в костном мозге больного возможно выявление раннего начала рецидива заболевания с принятием надлежащих мер.

В настоящее время для распознавания изображений типов клеток применяются различные подходы [7–9]. При этом одна из нерешенных проблем – автоматический анализ изображений препаратов клеток костного мозга при диагностике острых лейкозов и минимальной остаточной болезни. Представленные в публикациях данные (порою впечатляющие) базируются зачастую на малых объемах и разных составах выборок клеток, что затрудняет сравнение рассматриваемых методов и не дает исчерпывающей информации об их эффективности и корректности полученных результатов. Наше

исследование основано на уникальном архиве препаратов костного мозга Лаборатории иммунологии гемопоэза Национального медицинского исследовательского центра онкологии им. Н. Н. Блохина. Представленный материал отражает вариации клеток костного мозга по изначально выбранным видам и опыт врачей клинической лабораторной диагностики.

Целью работы является предварительное исследование нейросетевого классификатора при распознавании клеток костного мозга в диагностике острых лейкозов и минимальной остаточной болезни.

Материалы и методы. Создание системы распознавания клеток костного мозга на основе нейросетей представляет сложную масштабную задачу, требующую поэтапных решений. В данной связи в настоящей работе рассмотрен первый этап исследований по применимости нейросетей для распознавания клеток костного мозга.

Поставленная в работе цель связана с решением следующих задач:

- формирование выборки изображений;
- определение структуры нейронной сети;
- физическая реализация нейронной сети;
- экспериментальное исследование нейронной сети.

Формирование выборки изображений. Исходные цифровые микроскопические изображения препаратов костного мозга получены с помощью системы компьютерной микроскопии, включающей роботизированный световой микроскоп Olympus BX43, цифровую камеру Imregx IPX-4M15T-GCFB и компьютер. Электронное изображение получается в результате воздействия на светочувствительную матрицу камеры потока электромагнитного излучения видимого диапазона длин волн, формируемого осветителем микроскопа и прошедшего через препарат костного мозга на предметном стекле микроскопа. В ячейках светочувствительной матрицы камеры накапливается заряд, пропорциональный количеству попавших в ячейки фотонов. Распределение зарядов по ячейкам соответствует пространственному распределению доли фотонов, прошедших через окрашенные структурные элементы клеток костного мозга, находящихся на предметном стекле микроскопа. При последовательном электронном сканировании ячеек матрицы камеры пропорционально накопленному в них заряду на входе аналого-цифрового преобразователя (АЦП), входящего в устройство камеры, формируется электрическое напряжение, в соответствии с которым на выходе АЦП генерируется цифровой код. Полученный код записывается в память компьютера. Таким образом, полу-

ченная в компьютере совокупность цифровых кодов для всех ячеек матрицы камеры и представляет цифровое изображение, отражающее распределение окрашенных структур в клетках костного мозга.

Для формирования выборки изображений использовались цитологические препараты костного мозга 27 больных с подтвержденным диагнозом “острый лейкоз” (ОЛ). Препараты изготовлены в Лаборатории иммунологии гемопоэза Национального медицинского исследовательского центра онкологии им. Н. Н. Блохина. С помощью компьютерной программы на изображениях для каждой клетки выделялось ядро, окружающий фон закрашивался черным цветом. Получившееся изображение ядра сохранялось в файле с размером 192×192 пикселя с описанием цвета в RGB модели и восьмибитным кодированием цветовых компонент (всего 24 бит на пиксель).

Одной из проблем при обучении нейронных сетей для распознавания клеток костного мозга является малый объем верифицированных данных – изображений клеток костного мозга, классифицированных опытными врачами – экспертами. В представляемой работе в эталонную выборку было включено 8195 изображений клеток костного мозга следующих типов: миелоцит (126), сегментоядерный нейтрофил (277), нормобласт (288), палочкоядерный нейтрофил (92), бласт (6625), лимфоцит (507), моноцит (147), метамиелоцит (91), эозинофил (42).

Эталонная выборка была разделена на две части – тестовую и обучающую. Обычно при обучении искусственных нейронных сетей объем тестовой выборки меньше обучающей, а соотношение их размеров варьируется в широких пределах от 1 к 4 до 1 к 25 [10–12]. При этом, чем больше объем обучающей выборки, тем точнее может быть настроена нейронная сеть, а чем больше объем тестовой выборки, тем достовернее результат оценки точности классификации с применением нейронной сети. Ввиду относительно небольшого размера выборки исходных изображений клеток костного мозга нами принято соотношение тестовой выборки к обучающей как 1 к 4. Ввиду неравномерного представительства в исходной выборке разных типов клеток была выполнена процедура балансировки выборки, основанная на создании дополнительных изображений путем поворота исходного изображения на несколько разных углов. Таким образом, для проведения эксперимента по обучению нейронной сети и оценке точности классификации клеток костного мозга общее количество изображений клеток каждого типа доведено до 6625.

Определение структуры нейронной сети. Вначале структура нейронной сети представлена двумя слоями. Входной слой состоял из 110592 нейронов. Данное количество обусловлено исходным размером изображения 192×192 пикселей и тремя цветовыми компонентами. Выходной слой состоял из 9 нейронов, число которых соответствует количеству классифицируемых типов клеток. Тип клетки определяется путем выбора максимального значения на выходах нейросети, диапазон значений каждого из выходов от 0 до 1. Окончательная структура нейронной сети выбиралась в ходе эксперимента. Для изменения первоначальной структуры добавлялись попарно сверточный слой и слой максимальной подвыборки – субдискретизации. Для перехода от слоев свертки к выходному слою использовался слой сглаживания и дополнительный полносвязный слой. В эксперименте варьировалось число пар сверточный слой + слой максимальной подвыборки.

Физическая реализация нейронной сети. Программная реализация задачи осуществлялась на языке Python. Версии программного обеспечения и использованных библиотек: Python 3.7, PyCharm 2019.3.1, OpenCV 4.2.0, TensorFlow 2.0 (модули: Image, Io, Data, Keras).

Экспериментальное исследование нейронной сети. Цель эксперимента – уточнение структуры нейронной сети и определение точности классификации типов клеток костного мозга с применением разработанной нейронной сети.

Вычислительный эксперимент проводился на персональном компьютере с процессором Intel i5-9600KF и ОЗУ 32ГБ.

В результате эксперимента определена следующая структура нейронной сети:

- входной слой (110592 нейронов);
- первый внутренний слой – сверточный;
- второй слой – субдискретизационный;
- третий слой – сверточный;
- четвертый слой – субдискретизационный;
- пятый слой – сверточный;
- шестой слой – сглаживающий, реализует переход от сверточной сети к вектору;
- седьмой полносвязный слой (64 нейрона);
- выходной слой (9 нейронов).

При операции “свертка” использовался размер ядра свертки 3×3 , субдискретизация выполнялась посредством определения максимального значения в окне размером 2×2 , в качестве функции активации нейрона использовалась ReLU.

Настройка нейросети выполнялась с применением обучающей выборки (по 5300 изображений клеток каждого типа). Обучение проводилось в течение 10 эпох. На рис. 1 представлена зависимость средней точности распознавания девяти типов клеток костного мозга на обучающей и тестовой выборках в зависимости от номера эпохи обучения. После шестой эпохи наблюдается уменьшение точности на тестовой выборке при продолжающемся росте точности на обучающей выборке. Это свидетельствует о переобучении нейросети. Поэтому принято, что на шестой эпохе обучения нейросеть настроена оптимально.

Созданная нейронная сеть была экспериментально проверена на тестовом наборе данных (по 1325 изображений клеток каждого типа). Результаты эксперимента для разных типов клеток представлены в табл. 1. Среднее значение точности распознавания клеток костного мозга по исследованным девяти типам клеток составило 85%.

При заболевании острым лейкозом субстрат злокачественной опухоли представлен бластными клетками, поэтому важным является обнаружение бластных клеток в исследуемом препарате костного мозга. На тестовой выборке в 1325 бластных клеток разработанная нейронная сеть продемонстрировала точность распознавания в 91%.

Полученные результаты могут быть использованы при проектировании автоматизированных систем диагностики острых лейкозов и минимальной остаточной болезни.

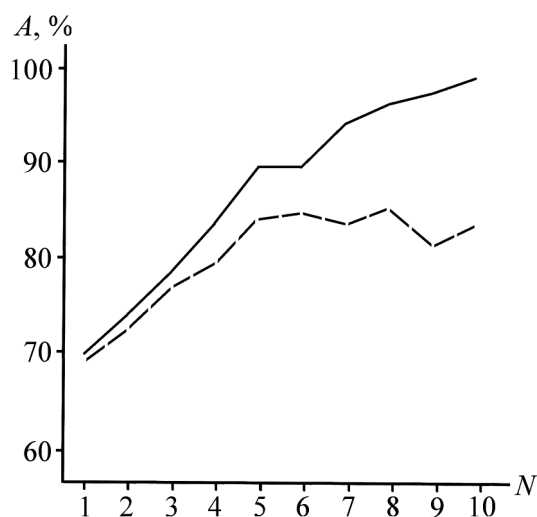


Рис. 1: График зависимости средней точности распознавания клеток костного мозга в % от количества эпох N (сплошная линия для обучающей выборки, для тестовой — прерывистая).

Т а б л и ц а 1

Точность распознавания клеток костного мозга

Тип клетки	Точность распознавания для обучающей выборки в %	Точность распознавания для тестовой выборки в %
Миелоцит	86	81
Сегментоядерный нейтрофил	98	97
Нормобласт	74	87
Палочкоядерный нейтрофил	91	87
Бласт	98	91
Лимфоцит	71	68
Моноцит	99	70
Метамиелоцит	92	88
Эозинофил	97	97

Заключение. Настоящая работа посвящена междисциплинарной (физика, биология, медицина, информатика) задаче распознавания клеток костного мозга для диагностики острых лейкозов и минимальной остаточной болезни.

Рассмотрено применение нейронных сетей для определения классов клеток в микроскопическом изображении препарата костного мозга. Изображения формировались в видимом диапазоне электромагнитного излучения. Проведено предварительное исследование применимости нейронных сетей. При проведении классификации изображений клеток костного мозга использовалась многослойная нейронная сеть.

Проанализированы исходные данные и создан алгоритм классификации, основанный на оптимальной структуре сети.

В эксперименте рассмотрено девять типов клеток. В среднем точность распознавания исследованных клеток костного мозга с применением нейронной сети составила 85%. При этом для распознавания бластных клеток, составляющих субстрат злокачественной опухоли при острых лейкозах, получена 91% точность.

Представленные результаты исследования полезны при создании автоматизированных систем диагностики онкогематологических заболеваний.

Работа выполнена при поддержке РФФИ по проекту № 18-29-09115.

Л И Т Е Р А Т У Р А

- [1] V. G. Nikitaev, Measurement Techniques **58**, 214 (2015). DOI:10.1007/s11018-015-0688-0.
- [2] V. G. Nikitaev, Measurement Techniques **58**, 467 (2015). DOI:10.1007/s11018-015-0736-9.
- [3] V. G. Nikitaev, Measurement Techniques **58**, 719 (2015). DOI:10.1007/s11018-015-0781-4.
- [4] T. Matsuda and M. Niino, Japanese journal of clinical oncology **46**(3), 290 (2016). DOI:10.1093/jjco/hyw031.
- [5] M. M. Amin, S. Kermani, and A. Talebi, Journal of medical signals and sensors **5**(1), 49 (2015). DOI:10.4103/2228-7477.150428.
- [6] A. N. Pronichev, E. V. Polyakov, N. N. Tupitsyn, et al., J. Phys.: Conf. Ser. **784**, 012003 (2017). DOI:10.1088/1742-6596/784/1/012003.
- [7] Mu-Chun Su, Chun-Yen Cheng, and Pa-Chun Wang, The Scientific World Journal **2014**, 9 (2014). DOI:10.1155/2014/796371.
- [8] H. Liu, H. Cao, and E. Song, Journal of Medical Systems **43**, 82 (2019). DOI:10.1007/s10916-019-1185-9.
- [9] P. Gabel, M. Crysandt, R. Herwartz, et al., Lecture Notes in Computer Science **11039**, Springer, Cham, 2018. DOI:10.1007/978-3-030-00949-6_10.
- [10] C. J. Shallue, J. Lee, J. Antognini, et al., Journal of Machine Learning Research **20**, 1 (2019).
- [11] M. S. Seyfioglu, S. Z. Gurbuz, IEEE Geoscience and remote sensing letters **14**(12), 2462 (2017). DOI:10.1109/LGRS.2017.2771405.
- [12] A. G. Shumixin and A. S. Boyarshinova, Computer Research and Modeling **7**(2), 243 (2015). DOI:10.20537/2076-7633-2015-7-2-243-251.

Поступила в редакцию 30 июля 2020 г.

После доработки 5 сентября 2021 г.

Принята к публикации 6 сентября 2021 г.