

Искусственный интеллект как инструмент для решения практических задач в нейробиологических исследованиях

Д. С. Бережной¹, Т. К. Бергалиев²

Приведен обзор основных технических задач в рамках современной нейробиологии, которые могут быть успешно решены с применением технологий машинного обучения и искусственного интеллекта. Рассмотрены основные сложности при анализе временных рядов и сегментации изображений, с которыми сталкиваются исследователи, и способы их решения. Дан обзор современных технических решений с использованием сверточных нейросетей для решения практических задач в нейробиологических исследованиях.

Ключевые слова: нейросети, машинное обучение, обработка данных, анализ изображений, анализ временных рядов

1. Введение

Развитие технологий ИИ в последние годы во многом направляется областями их непосредственного применения. Так, интеллектуальный анализ больших данных является одним из основных направлений приложения искусственного интеллекта. Потребность в подобном анализе появилась в частности в связи с развитием цифрового бизнеса, увеличивающимися темпами накопления данных в цифровом виде и необходимостью их систематизации и интерпретации. Также благодаря актуальности стоявших прикладных задач активное развитие получили ряд направлений ИИ, в которых сейчас достигнуты особые успехи: обработка естественного языка, изображений, задачи кластеризации и классификации. Значительного развития достигли технологии нейросетевого анализа и классификации изображений, анализа и предсказания временных рядов.

При этом активное развитие ИИ-технологий наблюдается в прикладных областях — в области рекламы, развлечений, информационной без-

¹ *Бережной Даниил Сергеевич* — н.с. каф. высшей нервной деятельности Биологического ф-та МГУ; ст.н.с. ФГБНУ Научный центр неврологии, e-mail: berezhnoy.daniil@gmail.com.

Berezhnoy Daniil Sergeevich — researcher at Department of Higher Nervous Activity, Faculty of Biology, Lomonosov Moscow State University; senior researcher at Research Center of Neurology

² *Бергалиев Тимур Кайратович* — н.с. НИУ МФТИ; генеральный директор ООО BiTronics Lab, e-mail: timurbergaliyev@gmail.com.

Bergaliyev Timur Kairatovich — researcher at MIPT, head of BiTronics Lab Ltd.

опасности, финансов, автомобилестроения, потребительской техники. В то же время, складывается парадоксальная ситуация, когда в фундаментальных научных исследованиях, требующих достаточно сложных методов анализа данных, системы искусственного интеллекта до последнего времени не получали такого широкого распространения. В биологических науках это может быть связано с нехваткой математического базиса, недостаточной технической подготовкой исследователей, и, как следствие, относительно высоким “порогом входа” для освоения подобных технологий. И лишь в последние 5 лет эта ситуация постепенно начинает меняться в лучшую сторону — появляется все больше научных работ на стыке биологии и ИИ, прототипов и готовых решений для использования ИИ в биологических исследованиях [1]. В данной работе мы постараемся дать обзор подобных решений в нейробиологии и выделить направления их развития.

В рамках современных нейробиологических исследований возникает целый ряд технических задач, которые могут быть успешно решены с привлечением нейросетевых алгоритмов, получивших особое развитие в последние годы. В числе таких задач можно выделить анализ и классификацию временных рядов электрофизиологических данных (биосигналов) и выделение значимых объектов в зашумленных изображениях: как гистологических срезов мозга, так и видеоизображениях испытуемых в эксперименте.

2. Анализ биосигналов мозга

Исследования электрической активности головного мозга классически являются одним из основных направлений нейробиологии. При этом анализируемый сигнал в форме изменения потенциала во времени может представлять собой как активность отдельных клеток или их небольших популяций (LFP — local field potential), так и суммарную электрическую активность коры головного мозга (EEG — electroencephalogram). Если учитывать, что в современных исследованиях количество регистрируемых каналов часто превышает 100, а частота записи более 1 кГц, уже одна запись часового эксперимента ставит вопрос о поиске закономерности в больших данных. Задачей исследователя, в самом общем виде, является выделение определенных значимых событий или паттернов активности в этом сигнале, что может быть формализовано как анализ временных рядов (time-series analysis). Обычными методами здесь является исследование в частотной (например, спектральный анализ) или временной области (например, автокорреляция). Первыми из методов интеллектуального анализа данных, успешно применяемыми в области электрофизиологии, стали методы анализа компонент PCA и ICA, на-

пример, для выделения разрядов отдельных нервных клеток в суммарной записи с электрода [2]. Применяемые уже около 30 лет, благодаря появлению прикладных библиотек вроде scikit-learn для языка python и свободных программных решений эти методы в настоящее время доступны еще более широкому кругу исследователей. В то же время только в последние годы стали активно применяться методы полностью автоматической классификации фрагментов записи ЭЭГ на основе машинного обучения и нейросетевых алгоритмов [3]. Для классификации используются либо сами фрагменты записи — короткие “эпохи” по несколько секунд — либо производные их кепстрального анализа. В качестве методов для классификации в современных исследованиях используются самые различные варианты, которые приведены в подробном мета-анализе [4], но наибольшую эффективность показали сверточные нейросети. Нейросетевой подход с успехом был применен к задачам выделения сна и бодрствования, аффективных и когнитивных состояний, эпилепсии и других видов патологии. В то же время, большинство исследований задействуют открытые библиотеки python, поэтому создание удобного графического программного интерфейса для создания обучающей выборки биосигналов является необходимой задачей для дальнейшего распространения подобного подхода.

3. Подсчет клеток в гистологических исследованиях

Задача обработки и анализа изображений в нейробиологии и, шире, в биологии, является одной из рутинных. В случае биомедицинских исследований анализ морфологического состава биологической ткани, измерения размера или количества определенных объектов по статичному изображению с микроскопа является достаточно простой, но регулярно повторяемой процедурой. Тем не менее, эта чисто техническая операция отнимает значительное количество времени при ручном выполнении, а в опросе 2009 года 71% исследователей отметили, что осуществляют подсчет клеток в гистологических препаратах вручную [5]. Отчасти это объясняется отсутствием удобных программных решений для автоматизированного выполнения подобной задачи, а отчасти — сложностью её осуществления с помощью стандартных методов анализа изображений. Основным инструментом анализа изображений в биологических исследованиях является программа ImageJ, которая является свободно доступной и для которой существует несколько плагинов подсчета клеток [6]. Большинство из них основано на стандартной процедуре выделения порога по одному из каналов (thresholding), что не всегда подходит для зашум-

ленных изображений или пересекающихся объектов сложной формы, к которым относятся и нервные клетки. Одним из перспективных подходов к решению этой проблемы является использование методов распознавания объектов с использованием нейросетей. В последние 5 лет появилось несколько готовых приложений для подсчета количества клеток [7, 8], основанных на предобученных сверточных нейросетях, среди которых наибольшую эффективность показывает архитектура U-Net. Новые разработки по этой теме включают также автоматическую сегментацию изображений с микроскопа и увеличение разрешения с помощью генеративных нейросетей [9].

4. Видео-трекинг и анализ поведения

Задача выделения объекта на изображении в нейробиологических исследованиях имеет и еще одно применение — анализ поведения животного по экспериментальным видеозаписям. Самым простым вариантом этой задачи является видеотрекинг — выделение траектории перемещения человека, животного или выбранной части тела в последовательности изображений. Учитывая сложность экспериментальной среды, возможных перемещений объекта и саму проблему отделения объекта от фона, вопрос применения систем компьютерного зрения здесь является крайне актуальным. В 2018 году для решения этой задачи была разработана система DeepLabCut, позволяющая с высокой точностью осуществить трекинг нескольких объектов по видео из эксперимента без предварительной обработки [10]. Успешность и скорость работы системы основаны на высокой эффективности переноса обучения (transfer learning) предобученных на открытой базе ImageNet нейросетей при их тренировке на новые объекты. Основным же плюсом системы, созданной с использованием открытых библиотек для языка python, является наличие графического интерфейса, показывающего пользователю все необходимые шаги от тренировки и оценки качества обучения нейросети до выделения объектов на тестовых видео. Таким образом, создана интуитивно-понятная система в виде приложения, которая может быть свободно использована любым исследователем без предварительного опыта с нейросетевым анализом изображений. Современные работы с применением данной системы показывают, что она уже вышла за рамки своей исходной прикладной задачи и позволяет решать уже чисто исследовательские вопросы автоматической сегментации и поиска паттернов поведения, незаметных при других способах анализа [11].

Этот пример показывает, что создание подобных интегрированных систем, решающих практические задачи и доступных для исследователей-биологов, не имеющих специальных знаний для примене-

ния ИИ, является необходимым шагом для дальнейшего развития междисциплинарного направления, связывающего исследования искусственного и естественного интеллекта.

Список литературы

- [1] Hassoun S., Jefferson F., Shi X., et al, “Artificial Intelligence for Biology”, *Integrative and Comparative Biology*, 2021.
- [2] Chapin J.K., Nicolelis M. A.L., “Principal component analysis of neuronal ensemble activity reveals multidimensional somatosensory representations”, *Journal of Neuroscience Methods*, **94**:1 (1999), 121–140.
- [3] Gemein L., Schirrmester R., Chrabaszcz P., et al, “Machine-learning-based diagnostics of EEG pathology”, *NeuroImage*, **220** (2020), 117021.
- [4] Roy Y., Banville H., Albuquerque I., et al, “Deep learning-based electroencephalography analysis: a systematic review”, *Journal of Neural Engineering*, **16**:5 (2019), 051001.
- [5] Ongena K., Das C., Smith J., et al, “Determining Cell Number During Cell Culture using the Scepter Cell Counter”, *Journal of Visualized Experiments*, 2010, № 45.
- [6] Grishagin I., “Automatic cell counting with ImageJ”, *Analytical Biochemistry*, **473** (2015), 63–65.
- [7] Antonets, D., Russkikh, D., Sanchez, N., et al, “CellCountCV—A Web-Application for Accurate Cell Counting and Automated Batch Processing of Microscopic Images Using Fully Convolutional Neural Networks”, *Sensors*, **20**:13 (2020), 3653.
- [8] Falk T., Mai D., Bensch R., et al, “U-Net: deep learning for cell counting, detection, and morphometry”, *Nature Methods*, **16**:1 (2019), 67–70.
- [9] Ishii S., Lee S., Urakubo H., et al, “Generative and discriminative model-based approaches to microscopic image restoration and segmentation”, *Microscopy*, **69**:2 (2020), 79–91.
- [10] Mathis A., Mamidanna P., Cury K., et al, “DeepLabCut: markerless pose estimation of user-defined body parts with deep learning”, *Nature Neuroscience*, **21**:9 (2018), 1281–1289.
- [11] Shin J., Song M., Paik S., Jung M., “Spatial organization of functional clusters representing reward and movement information in the striatal direct and indirect pathways”, *Proceedings of the National Academy of Sciences*, **117**:43 (2020), 27004–27015.

Artificial intelligence as a research tool for solving technical tasks in neurobiology studies

Berezhnoy D.S., Bergaliev T.K.

We give an overview of the main technical problems in the modern neurobiology studies, which can be successfully solved using machine learning and artificial intelligence technologies. The main difficulties encountered by researchers in the analysis of biosignals time series and image segmentation are considered. An overview of modern technical

solutions using convolutional neural networks for solving practical problems in neurobiological research is given.

Keywords: neural networks, machine learning, data processing, image processing, time-series analysis

References

- [1] Hassoun S., Jefferson F., Shi X., et al, “Artificial Intelligence for Biology”, *Integrative and Comparative Biology*, 2021.
- [2] Chapin J.K., Nicolelis M.A.L., “Principal component analysis of neuronal ensemble activity reveals multidimensional somatosensory representations”, *Journal of Neuroscience Methods*, **94**:1 (1999), 121–140.
- [3] Gemein L., Schirrneister R., Chrabaszcz P., et al, “Machine-learning-based diagnostics of EEG pathology”, *NeuroImage*, **220** (2020), 117021.
- [4] Roy Y., Banville H., Albuquerque I., et al, “Deep learning-based electroencephalography analysis: a systematic review”, *Journal of Neural Engineering*, **16**:5 (2019), 051001.
- [5] Ongena K., Das C., Smith J., et al, “Determining Cell Number During Cell Culture using the Scepter Cell Counter”, *Journal of Visualized Experiments*, 2010, № 45.
- [6] Grishagin I., “Automatic cell counting with ImageJ”, *Analytical Biochemistry*, **473** (2015), 63–65.
- [7] Antonets, D., Russkikh, D., Sanchez, N., et al, “CellCountCV—A Web-Application for Accurate Cell Counting and Automated Batch Processing of Microscopic Images Using Fully Convolutional Neural Networks”, *Sensors*, **20**:13 (2020), 3653.
- [8] Falk T., Mai D., Bensch R., et al, “U-Net: deep learning for cell counting, detection, and morphometry”, *Nature Methods*, **16**:1 (2019), 67–70.
- [9] Ishii S., Lee S., Urakubo H., et al, “Generative and discriminative model-based approaches to microscopic image restoration and segmentation”, *Microscopy*, **69**:2 (2020), 79–91.
- [10] Mathis A., Mamidanna P., Cury K., et al, “DeepLabCut: markerless pose estimation of user-defined body parts with deep learning”, *Nature Neuroscience*, **21**:9 (2018), 1281–1289.
- [11] Shin J., Song M., Paik S., Jung M., “Spatial organization of functional clusters representing reward and movement information in the striatal direct and indirect pathways”, *Proceedings of the National Academy of Sciences*, **117**:43 (2020), 27004–27015.