

УДК 004.932

ОСОБЕННОСТИ ИСПОЛЬЗОВАНИЯ СРЕДСТВ МАШИННОГО ОБУЧЕНИЯ ПРИ АНАЛИЗЕ МЕДИЦИНСКИХ ИЗОБРАЖЕНИЙ

© Трубников А.А.

*Самарский национальный исследовательский университет
имени академика С.П. Королева, г. Самара, Российская Федерация*

e-mail: antei3mer@gmail.com

Большое количество задач диагностики в современной медицине можно свести к задаче анализа изображений. Рентгенография, компьютерная томография, микроскопия являются базовыми медицинскими исследованиями и подчас единственными доступными. Экспоненциальный рост количества изображений как данных в медицине дал возможность применять методы машинного обучения, которые в некоторых случаях достигают уровня точности человека [1].

Существует большое количество проектов по реализации средств медицинской диагностики с использованием средств машинного обучения, однако пока что широкого распространения во врачебной практике они не получили. Так как программное решение влияет на постановку диагноза, то к нему предъявляют особые требования. Специфика данных для обучения также накладывает дополнительные ограничения на модель.

В качестве примера рассмотрим конкретную задачу. Острый лимфобластный лейкоз – раковое заболевание крови, при котором системы кроветворения начинает бесконтрольное производство незрелых лимфоидных клеток (лимфобластов) (рис. 1). Данное заболевание является самым распространенным злокачественным заболеванием в детском и юношеском возрасте [2].

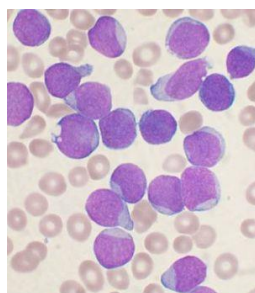


Рис. 1. Аспират костного мозга больного человека

Первичными исследованиями, проводимыми при диагностике этого заболевания, являются доступные анализ крови и мазок периферической крови. Доступность микроскопической фотографии сделала возможным использовать методы машинного обучения в задаче бинарной классификации (рис. 2).



Рис. 2. Фотографии белых клеток: (а) лимфобласт, (б) здоровая клетка [3]

Как видно из рисунков 1 и 2, клетки очень схожи внешне, что затрудняет постановку диагноза врачом, однако, используя методы машинного обучения, мы можем повысить качество (точность, скорость) диагностики.

Как и в других похожих исследованиях, для создания программного решения поставленной задачи используются язык программирования Python 3 и библиотека машинного обучения tensorflow [4]. Данная библиотека предоставляет удобный интерфейс работы с данными и построением моделей. Признаки извлекаются при помощи сверточной нейронной сети, в качестве классификатора используется полносвязная многослойная нейронная сеть.

Библиографический список

1. Torsten W.K., Ingrid S., Til A. Challenges of medical image processing // Computer Science – Research and Development. 2011. Vol. 26. DOI: 10.1007/s00450-010-0146-9.
2. Acute Lymphoblastic Leukemia. URL: <https://www.cancer.gov/types/leukemia/patient/child-all-treatment-pdq> (дата обращения: 20.02.2021).
3. C-NMC dataset.
4. Tatdow P., Siripen W., Kittiya K., Aniruth P. Convolutional Neural Networks for Recognition of Lymphoblast Cell Images // Computational Intelligence and Neuroscience. 2019. Vol. 2019. DOI: 10.1155/2019/7519603.