

Шеретов М.А.

**СОВРЕМЕННЫЕ МЕТОДЫ АНАЛИЗА И КЛАССИФИКАЦИИ
КЛЕТОК КРОВИ: НЕОБХОДИМОСТЬ ИНТЕРПРЕТИРУЕМЫХ
АЛГОРИТМОВ**

В данной статье рассмотрены различные модели для анализа, подсчёта и классификации клеток крови. Традиционные методы, такие как метод опорных векторов и преобразование Хафа, показали свою эффективность, но требуют дорогостоящего оборудования и сложных процедур преданализа, что затрудняет их применение в современных клиниках. Современные подходы с использованием свёрточных нейронных сетей предлагают автоматизацию и снижение требований к преданализу за счёт автоматического выделения признаков. Однако основной недостаток свёрточных нейронных сетей заключается в сложности интерпретации результатов. Поднимается вопрос о необходимости разработки инструмента для интерпретации результатов работы свёрточных нейронных сетей, выявления причинно-следственной связи.

Ключевые слова: классификация клеток крови, метод опорных векторов, преобразование Хафа, свёрточные нейронные сети, преданализ.

SHERETOV M.A.

**MODERN METHODS OF BLOOD CELL ANALYSIS AND
CLASSIFICATION: THE NEED FOR INTERPRETABLE ALGORITHMS**

This article reviews various models for analyzing, counting and classifying blood cells. Traditional methods, such as support vector methods and Hough transforms, have been shown to be effective but require expensive equipment and complex pre-analytic procedures, making them difficult to implement in modern clinics. Current approaches using convolutional neural networks offer automation and reduced pre-analytics requirements through automatic feature extraction. However, the main drawback of convolutional neural networks is the difficulty in interpreting the

results. The question of the need to develop a tool for interpreting the results of convolutional neural networks and identifying causality is raised.

Keywords: Blood cell classification, support vector method, Hough transform, convolutional neural networks, pre-analytics.

Существует множество моделей для анализа, подсчёта и классификации клеток крови. К примеру Lou и др. использовали метод опорных векторов на данных, полученных после спектрального анализа, для подсчёта эритроцитов [1]. Но этот подход требует дорогостоящего оборудования, не доступного большинству современных клиник. Cruz и др. создали систему, использующую пороговое значение насыщенности оттенка и связную маркировку для независимого подсчёта эритроцитов, лейкоцитов и тромбоцитов [2]. Acharya and Kumar построили систему подсчёта эритроцитов используя разделение вокруг медоида и преобразование Хафа [3]. Singhal и Singh классифицировали злокачественные и доброкачественные лейкоциты, используя 256 локальных бинарных шаблонов [4]. Затем они увеличили количество используемых ими шаблонов до 4096, чтобы улучшить результаты классификации [5]. И наоборот, Bhattacharjee и Saini использовали только 8 шаблонов для работы метода k-ближайших соседей для классификации лейкоцитов [6]. Ученые из университета Джорджа Мейсона предложили использовать метод опорных векторов, используя легко интерпретируемые признаки, извлечённые при помощи алгоритмов компьютерного зрения[7].

Но каждый из представленных учеными методов очень требователен к набору данных, на котором он производится, увеличивая требования к ресурсам для процедуры преданализики — сбору и окраске исходного материала, что делает их трудно воспроизводимыми. При нарушении чётких правил пробоподготовки, результат сильно снижается.

Популярным методом анализа и классификации изображений, способным понизить требования к этапу преданализики, является использование сверточных нейронных сетей [8]. Эта модель вычислений очень популярна в анализе медицинских изображений [9,10]. Свёрточные нейронные

сети способны самостоятельно определить важные, для решения конкретной задачи, признаки [8]. Kutlu и другие, а также Long' использовали различные свёрточные нейронные сети для классификации разных типов лейкоцитов [11,12]. Существует множество различных архитектур для CNN, включая VGGNet [13], Resnet [14] и YOLO [15]. Трансферное обучение особенно полезно, когда исследователь располагает набором данных с небольшим количеством наблюдений или экземпляров. При таком подходе заданная архитектура сначала тренируется на очень большом наборе данных. Затем предварительно обученная свёрточная нейронная сеть тренируется на задаче, которую исследует аналитик [8].

Один из главных недостатков использования свёрточных нейронных сетей это сложность интерпретации модели используемых ею признаков, вне зависимости от архитектуры. Некоторые исследователи снижают количество признаков, используемых для классификации, и добивались успеха. Так, например Sahlol и др. использовали трансферное обучение модели VGGNet в сочетании со статистически улучшенным алгоритмом Salp Swarm Algorithm (SESSA) для извлечения более полезных признаков для классификации здоровых и злокачественных лейкоцитов [10]. В результате было отобрано более 1000 признаков из 25088 возможных, которые затем были использованы в модели классификации на базе метода опорных векторов [10]. Однако этот подход всё ещё не даёт достаточного представления о важных, для определения доброкачественных и злокачественных лейкоцитов, признаках. Кроме того, эта система в основном занимается классификацией лейкоцитов и не способна отличить лейкоциты от эритроцитов или тромбоцитов.

В заключение можно сказать, что анализ, подсчёт и классификация клеток крови остаются сложными задачами, требующими высокоточных и эффективных методов. Традиционные подходы, такие как метод опорных векторов, пороговое значение насыщенности оттенка и преобразование Хафа, несмотря на их эффективность, часто требуют дорогостоящего оборудования и сложных процедур преданализики. Современные методы, такие как свёрточные

нейронные сети, предлагают более гибкие и автоматизированные решения, способные снизить требования к этапу преданализа за счёт автоматического выделения признаков.

Однако основным недостатком использования свёрточных нейронных сетей является сложность интерпретации результатов их работы. Несмотря на достижения в области трансферного обучения и использования различных архитектур свёрточных нейронных сетей, таких как VGGNet, Resnet и YOLO, остаются проблемы с выделением и интерпретацией ключевых признаков, особенно при классификации различных типов клеток крови.

Необходимо разработать алгоритм, который бы позволял интерпретировать результаты работы свёрточных нейронных сетей, обеспечивая более чёткое понимание важных признаков, используемых для классификации клеток. Такой алгоритм способствовал бы дальнейшим исследованиям на тему классификации клеток на пробах жидкостей человека.

В целом, дальнейшие исследования и разработки необходимы для создания более точных, доступных и интерпретируемых методов анализа клеток жидкостей человека.

Библиографический список

1. J. Lou, M. Zhou, Q. Li, C. Yuan and H. Liu, "An automatic red blood cell counting method based on spectral images," 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Datong, China, 2016, pp. 1391-1396, doi: 10.1109/CISP-BMEI.2016.7852934.

2. D. Cruz *et al.*, "Determination of blood components (WBCs, RBCs, and Platelets) count in microscopic images using image processing and analysis," *2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, Manila, Philippines, 2017, pp. 1-7, doi: 10.1109/HNICEM.2017.8269515.

3. Acharya, V., Kumar, P. Identification and red blood cell automated counting from blood smear images using computer-aided system. *Med Biol Eng Comput* 56, 483–489 (2018). <https://doi.org/10.1007/s11517-017-1708-9>

4. V. Singhal and P. Singh, "Local Binary Pattern for automatic detection of Acute Lymphoblastic Leukemia," 2014 Twentieth National Conference on Communications (NCC), Kanpur, India, 2014, pp. 1-5, doi: 10.1109/NCC.2014.6811261.

5. Singhal, V., Singh, P. (2016). Texture Features for the Detection of Acute Lymphoblastic Leukemia. In: Satapathy, S., Joshi, A., Modi, N., Pathak, N. (eds) Proceedings of International Conference on ICT for Sustainable Development. Advances in Intelligent Systems and Computing, vol 409. Springer, Singapore. https://doi.org/10.1007/978-981-10-0135-2_52.

6. R. Bhattacharjee and L. M. Saini, "Robust technique for the detection of Acute Lymphoblastic Leukemia," 2015 IEEE Power, Communication and Information Technology Conference (PCITC), Bhubaneswar, India, 2015, pp. 657-662, doi: 10.1109/PCITC.2015.7438079.

7. William Franz Lamberti, Blood cell classification using interpretable shape features: A Comparative study of SVM models and CNN-Based approaches, Computer Methods and Programs in Biomedicine Update, Volume 1, 2021, 100023, ISSN 2666-9900, <https://doi.org/10.1016/j.cmpbup.2021.100023>.

8. Jiuxiang Gu, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Gang Wang, Jianfei Cai, Tsuhan Chen, Recent advances in convolutional neural networks, Pattern Recognition, Volume 77, 2018, Pages 354-377, ISSN 0031-3203, <https://doi.org/10.1016/j.patcog.2017.10.013>. (<https://www.sciencedirect.com/science/article/pii/S0031320317304120>)

9. Esteva, A., Chou, K., Yeung, S. *et al.* Deep learning-enabled medical computer vision. *npj Digit. Med.* 4, 5 (2021). <https://doi.org/10.1038/s41746-020-00376-2>.

10. Sahlol, A.T., Kollmannsberger, P. & Ewees, A.A. Efficient Classification of White Blood Cell Leukemia with Improved Swarm Optimization of Deep Features. *Sci Rep* 10, 2536 (2020). <https://doi.org/10.1038/s41598-020-59215-9>.
11. Hüseyin Kutlu, Engin Avci, Fatih Özyurt, White blood cells detection and classification based on regional convolutional neural networks, *Medical Hypotheses*, Volume 135, 2020, 109472, ISSN 0306-9877, <https://doi.org/10.1016/j.mehy.2019.109472>.
12. Fei Long, Jing-Jie Peng, Weitao Song, Xiaobo Xia, Jun Sang, BloodCaps: A capsule network based model for the multiclassification of human peripheral blood cells, *Computer Methods and Programs in Biomedicine*, Volume 202, 2021, 105972, ISSN 0169-2607, <https://doi.org/10.1016/j.cmpb.2021.105972>.
13. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv:1409.1556 [cs] (2015). ArXiv: 1409.1556, <http://arxiv.org/abs/1409.1556>.
14. K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.
15. M.M. Alam, M.T. Islam, Machine learning approach of automatic identification and counting of blood cells, *Healthc Technol Lett* 6 (4) (2019) 103–108, <https://doi.org/10.1049/htl.2018.5098>. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6718065/>

Сведения об авторах

Шеретов Марк Алексеевич – магистрант кафедры ИТАС, Пермского национального исследовательского политехнического университета, группы АСУ8-23-1м, г. Пермь, e-mail: mark.sheretov@gmail.com