

# «Исследование возможностей обнаружения и классификации изображений ядросодержащих клеток костного мозга с применением нейронных сетей»

**Авторы:** Сергеев Роман Александрович, Сливенко Андрей Константинович, Трошкин Илья Дмитриевич.

**Научные руководители:** кандидат физико-математических наук Шувалова Екатерина Викторовна, кандидат технических наук Поляков Евгений Валерьевич

## 1. Актуальность

В контексте современной медицинской диагностики [1] применение машинного обучения и компьютерного зрения для обнаружения в изображениях ядросодержащих клеток крови и костного мозга обеспечивает выявление на раннем этапе онкологических заболеваний. Данные методы обладают рядом преимуществ: **Высокая точность диагностики:** Нейронные сети, обучаясь на больших объемах данных, позволяют выявлять тонкие и сложные закономерности, часто незаметные человеческому глазу или стандартным методам анализа, что повышает точность обнаружения атипичных клеток костного мозга. **Автоматизация и ускорение процесса:** Использование нейронных сетей позволяет автоматизировать процесс анализа медицинских изображений, сокращая время, необходимое для диагностики. **Снижение риска ошибок:** Искусственный интеллект, а именно глубокое обучение, в задачах диагностики заболеваний способствует минимизации риска человеческих ошибок, связанных с визуальным анализом изображений. **Постоянное обновление и улучшение:** Алгоритмы машинного обучения могут быть легко обновлены и объединены с новыми данными, позволяя искусственному интеллекту постоянно совершенствоваться и адаптироваться к изменениям в структуре и характере ядросодержащих клеток. **Интеграция с другими методами диагностики:** Нейронные сети могут взаимодействовать с другими методами диагностики, улучшая общую эффективность медицинского анализа и обеспечивая комплексный подход к выявлению заболеваний.

## 2. Проблемы

Проблемы обнаружения ядросодержащих клеток костного мозга с применением искусственного интеллекта включают в себя следующие аспекты [2]. **Сложность сбора данных:** Глубокое обучение требует большого количества аннотированных изображений высокого качества, которые могут быть труднодоступны из-за конфиденциальности медицинских данных и необходимости соблюдения этических норм. **Вариабельность образцов:** Изображения клеток костного мозга значительно различаются в зависимости от пациента, методов окрашивания, а также от качества и настроек микроскопического оборудования. **Неоднозначность в классификации:** Процесс классификации изображений клеток костного мозга чрезвычайно сложен, и даже опытные специалисты иногда не могут прийти к единому мнению, что создает трудности в создании однозначных аннотаций для обучения нейронных сетей. **Обработка данных:** Предобработка изображений для выделения клеток и их характеристик представляет собой сложную задачу, требующую разработки специальных алгоритмов сегментации, нормализации и аугментации данных. **Высокие требования к вычислительным ресурсам:** Обучение современных нейронных сетей, особенно сетей глубокого обучения, требует значительных вычислительных мощностей. **Переобучение и обобщающая способность:** Существует риск создания моделей, которые хорошо работают на обучающем наборе данных, но плохо обобщают результаты на новых, неизвестных данных. **Интеграция с клинической практикой:** Сложность представляет интеграция обученной модели в существующие клинические процессы, так как при этом может потребоваться выполнение ряда дополнительных исследований и адаптация под потребности медицинских специалистов.

## 3. Цель

Цель исследования заключается в обучении и последующем исследовании архитектуры нейронной сети для обнаружения изображений атипичных клеток костного мозга для разработки эффективного и точного инструмента для диагностики онкогематологических заболеваний. Нейронная сеть предназначена для автоматического анализа

изображений, таких как снимки костного мозга, с целью раннего выявления атипичных клеток. Основные задачи: **Обеспечение высокой точности** позволит рано выявлять онкологические заболевания. **Разработка быстрой и эффективной системы**, способной быстро и эффективно обрабатывать медицинские изображения с высокой производительностью для оперативного реагирования на признаки роста злокачественных новообразований. **Создание автономной системы**, работающей в реальном времени и с минимальным участием человека, для повышения эффективности работы медицинского персонала и сокращения времени диагностики. **Разработка универсальной модели**, способной анализировать разнообразные данные, включая данные из различных источников, разного разрешения и формата, обеспечивая масштабируемость и универсальность метода. **Интеграция разработанной нейронной сети с существующими медицинскими информационными системами** и введение новых технологий в клиническую практику.

#### 4. Предлагаемый подход

Методика включает в себя следующие шаги. **Формирование большого и разнообразного аннотированного набора данных**, содержащего изображения различных типов и состояний ядроклеток костного мозга. **Разделение на обучающую, проверочную и тестовую выборки** с учетом распределения количества клеток различных классов по набору данных. **Предварительная обработка изображений**, включающая выделение квадратных изображений с центрированием на объекте из файла меток, гауссовскую нормализацию, а также аугментацию изображений (повороты, смещения, отражения). **Создание нейронной сети с архитектурой YOLOv8 [3]**, подходящей для классификации изображений клеток крови. **Настройка гиперпараметров нейросети**, таких как скорость обучения, размер пакета (батча) и количество эпох. **Обучение нейросети на обучающей выборке** для настройки весов и минимизации функции потерь. **Анализ производительности модели на проверочной выборке** с использованием метрик валидации. **Тестирование модели на тестовой выборке** для оценки ее обобщающей способности.

#### 5. Краткое описание результатов работы

В ходе исследования были изучены возможности обнаружения и классификации изображений ядроклеток костного мозга с использованием нейронных сетей. Анализ собранных данных включал в себя 13 типов различных клеток и их характеристики. Разработаны специализированные методики подготовки и обработки медицинских изображений для их последующего использования в расчетах нейросетей. Создан и оптимизирован алгоритм глубокого обучения на основе нейронной сети модели YOLOv8, предназначенный для обнаружения и классификации ядерных клеток. Проведенные расчеты на подготовленном наборе данных с использованием настроенной модели демонстрируют высокую эффективность и точность. Результаты исследования прошли проверку и подтверждение специалистами в области диагностики заболеваний клеток крови. Данная методика и разработанная модель обладают потенциалом масштабирования и могут быть использованы в качестве автоматического аннотатора изображений клеток костного мозга. Исследование проведено в тесном сотрудничестве со специалистами в области диагностики, биоинформатики, машинного обучения и клинической медицины, что позволило комплексно подойти к решению задачи и учесть многогранность медицинской проблематики.

[1] Digital staining in optical microscopy using deep learning - a review /Kreiss L., Jiang S., Li X., Xu S., Zhou K., Lee K., Muhlberg A., Kim K.,Chaware A., Ando M., Barisoni L., Seung Ah L., Zheng G., Lafata K.,Friedrich O., and Horstmeyer R. // PhotoniX .—2023.—10.—Vol. 4

[2] Cheng, Zhizhao and Yuan-Ya Li. "Improved YOLOv7 Algorithm for Detecting Bone Marrow Cells." *Sensors (Basel, Switzerland)* 23 (2023): n. pag.

[3] <https://docs.ultralytics.com/modes/train/> Документация к библиотеке ultralytics YOLOv8.