

ЦИФРОВАЯ ТРАНСФОРМАЦИЯ АНАЛИЗА СНИМКОВ УЗИ ЩИТОВИДНОЙ ЖЕЛЕЗЫ

Ильин В.И. (НИЯУ МИФИ), Ложкин И.А. (НИЯУ МИФИ), Зайцев К.С.
(НИЯУ МИФИ)

Научный руководитель – доктор технических наук, профессор Зайцев К.С.
(НИЯУ МИФИ)

Введение. Цифровая трансформация анализа медицинских снимков представляет собой комплексную задачу, объединяющую в себе множество задач и технологий компьютерного зрения, обработки естественного языка и иных разделов машинного обучения и программирования. Вынесение своевременного диагноза является главным залогом успешного лечения в самых разных областях медицины, включая исследования узлов щитовидной железы, и в условиях острой нехватки высококвалифицированных кадров создание системы автоматического анализа снимков УЗИ щитовидной железы представляется особенно актуальным. Для обеспечения надежной работы такой системы необходимо качественно решить задачи обнаружения, сегментации, отслеживания и классификации узловых образований по системе EU-TIRADS на изображениях и кинопетлях УЗИ.

Основная часть. В рамках исследования был сформирован датасет, включающий 441 кинопетлю УЗИ щитовидной железы с узлами (139 пациентов) и 150 отдельных изображений без узлов (TIRADS1, 37 пациентов), в наборе данных присутствовал дисбаланс классов. Предварительная обработка включала удаление нерелевантных областей, изменение размеров и нормализацию изображений, а также отбор фиксированного числа кадров из кинопетель для обеспечения однородности входных данных при обучении и тестировании моделей. Разработанный пайплайн автоматизированного анализа состоит из нескольких ключевых этапов: предобработка снимков, детекция узлов с помощью модели YOLO12n [1], трекинг узлов в кинопетлях посредством трекера Bot-SORT [2], сегментация областей с узлом с использованием UNet++ [3] с энкодером EfficientNet-B7 [4] и финальная классификация узлов ансамблем моделей DL и ML. Особое внимание было уделено детекции: после дообучения модели YOLO11s на расширенном датасете (4560 кадров, включая данные из набора TN3K [5] из Интернета; от каждой кинопетли бралось до 4 кадров) удалось достичь повышения показателей: precision до 84,3%, recall до 74,1%, mAP50 до 79,7%. После обучения моделей 12-ой версии YOLO получены лучшие результаты на тесте у модели YOLO12n: precision = 88,4%, recall = 72,8%, mAP50 = 82,5%. Стоит отметить, что намеренно была выбрана модель с ориентацией на precision, чтобы уменьшить количество ложноположительных предсказаний на кинопетлях — для корректировки ложноотрицательных предсказаний в таком случае отработает трекер и алгоритм на постобработке. Для сегментации увеличенных областей с узлом модель UNet++ продемонстрировала высокий результат с Dice Coefficient равным 95%. Особенностью исследования стал трехуровневый ансамбль для классификации узлов по системе EU-TIRADS. Для обучения и тестирования моделей от каждой кинопетли брался кадр, на котором узел имеет наибольшую площадь. На первом уровне ансамбля задействованы шесть моделей DL: YOLO11-cls [1] для бинарной классификации узлов (определяющих близость к одному из двух классов по каждой паре классов от TIRADS2 до TIRADS5), показавшие точность от 77,1% до 95,1%. На втором уровне работают две мета-модели XGBoost [6], обучаемые на признаках, полученных от CV-моделей, а также с учетом типа проекции снимка (продольная или поперечная). Эти модели достигли accuracy до 73,9%. На третьем

уровне финальная мета-модель интегрирует предсказания двух предыдущих мета-моделей, обеспечивая итоговую точность классификации на четырех классах TIRADS2–TIRADS5 на уровне 74,6%, при этом precision (macro) составила 77,1%, recall (macro) — 68,6%, f1-score (macro) — 70,8%. Стоит отметить, что на текущем несбалансированном наборе данных небольшого объема приведенные метрики, рассчитанные для одной модели CV, классифицирующей узлы сразу на четыре класса TIRADS2–TIRADS5, не превышали 70%.

Выводы. В текущей работе дано представление цифровой трансформации анализа снимков УЗИ щитовидной железы. По результатам исследования был получен пайплайн, интегрирующий DL- и ML-модели для анализа многокадровых кинопетель и отдельных изображений УЗИ щитовидной железы, который демонстрирует, что ансамблевый подход позволяет повысить точность отдельных методов.

Список использованных источников:

1. Jocher G., Qiu J., Chaurasia A. Ultralytics YOLO [Электронный ресурс]. 2025. URL: <https://github.com/ultralytics/ultralytics> (дата обращения: 19.11.2025).
2. Aharon N., Orfaig R., Bobrovsky B.-Z. BoT-SORT: Robust Associations Multi-Pedestrian Tracking // arXiv preprint arXiv:2206.14651. 2022.
3. Zhou Z. Towards Annotation-Efficient Deep Learning for Computer-Aided Diagnosis: PhD thesis. — Arizona State University, 2021.
4. Tan M., Le Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks // arXiv preprint arXiv:1905.11946. 2020. URL: <https://arxiv.org/abs/1905.11946> (дата обращения: 19.11.2025).
5. Gong H., Chen G., Wang R. et al. Multi-task learning for thyroid nodule segmentation with thyroid region prior // 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). IEEE, 2021. P. 257–261.
6. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). New York, NY, USA: ACM, 2016. P. 785–794. DOI: 10.1145/2939672.2939785.

Автор _____ Ильин В.И.

Научный руководитель _____ Зайцев К.С.