



ИСП РАН

Андрей Бурсов
bursov@ispras.ru

Создание платформы для анализа биомедицинских данных

С. А. Лебедев



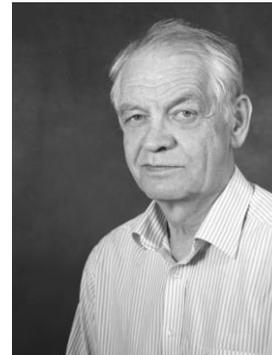
В.А. Мельников



БЭСМ-6 в музее науки Лондона



В.П. Иванников



Л.Н. Королёв



«Если мы глубже разберемся в этом эпохальном советском суперкомпьютере, это позволит пересмотреть заявления времён холодной войны об отставании русской технологии, а также подтвердить или развеять мифы о технологическом совершенстве наших союзников».

Doron Swade, senior curator of computing and information technology



2019: 25 лет ИСП РАН



2023: 75 лет ИТ в России

Цифровая трансформация преобразует не только применяемые технологии, но и культуру, и бизнес-процессы компании. Это фундаментальное переосмысление клиентского опыта, бизнес-моделей и операций. Это поиск новых путей создания ценности, генерации выручки и повышения эффективности.

Основополагающими технологиями цифровой трансформации являются искусственный интеллект, интернет вещей, робототехника, облачные решения.



Результаты цифровой трансформации



Снижение издержек



Новые модели деятельности



Повышение эффективности

Преимущества

- ✓ Есть много модальностей в цифре (сигналы, изображения, текст)
- ✓ Большое количество открытых датасетов
- ✓ Есть готовые открытые инструменты для аннотирования данных
- ✓ Повышается интерес со стороны медицинского сообщества
- ✓ Есть специализированные программы финансирования

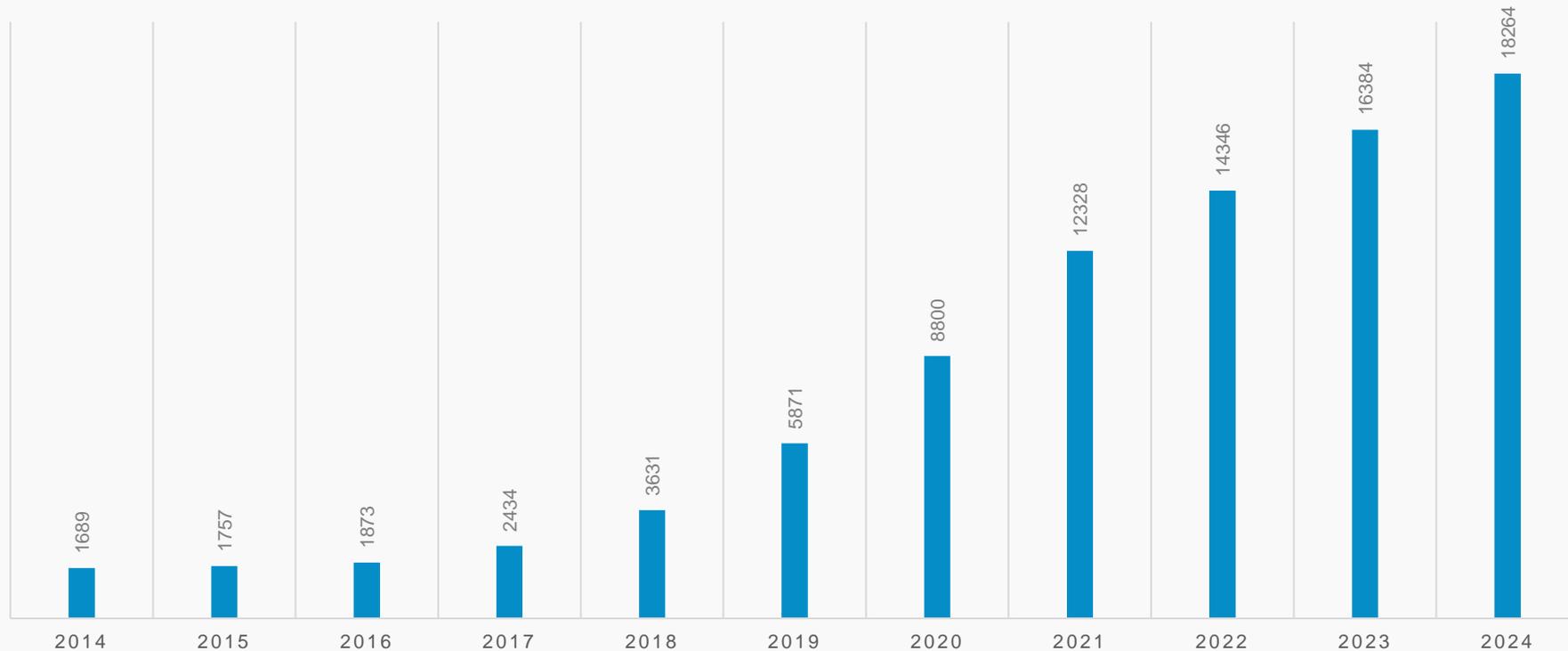


Недостатки

- ✓ Не все организации оцифровывают данные
- ✓ Низкая степень доверия к открытым датасетам (повторяемость)
- ✓ Организационные сложности с получением данных из медучреждений
- ✓ Высокий порог входа в анализ данных в медицине
- ✓ «Порочный круг» анализа данных в медицине

Тренд на применение искусственного интеллекта в медицине и здравоохранении

ARTIFICIAL INTELLIGENCE MEDICAL (PUBMED)



Вызовы

- ✓ Большие объемы медицинских данных
- ✓ Потребность в вычислительных ресурсах для проведения исследований
- ✓ Обеспечение доверия к цифровой медицине
- ✓ Независимость от зарубежных производителей
- ✓ Сложное междисциплинарное взаимодействие



Решения

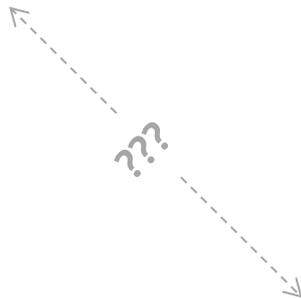
- ✓ Облачные технологии, технологии ИИ и машинного обучения
- ✓ Ресурсы по запросу
- ✓ Защита информации, доверенный ИИ
- ✓ Открытое и отечественное ПО
- ✓ Совместная работа на одной платформе



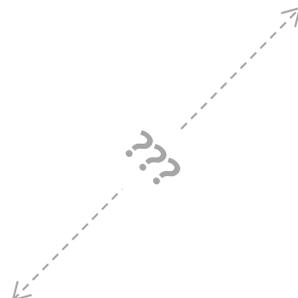
Собранные данные



Исследователи
И ИТ-специалисты



Медицинские работники



Научный центр мирового уровня «Цифровой биодизайн и персонализированное здравоохранение»

Технологическая основа – облачная цифровая экосистема (Платформа) для решения задач биомедицинского домена



2020

Определение требований, техническое проектирование, макет



2023

Начало опытной эксплуатации



2025

Прототип Платформы, готовый к промышленной эксплуатации

Основные задачи Платформы НЦМУ



Датасеты

Сбор и надежное хранение больших медицинских данных



Исследования

Объединение исследователей в междисциплинарные команды, обеспечение совместной работы



Сервисы

Внедрение результатов в реальный сектор



Аннотирование

Управление размеченными данными, синхронизация экспертов



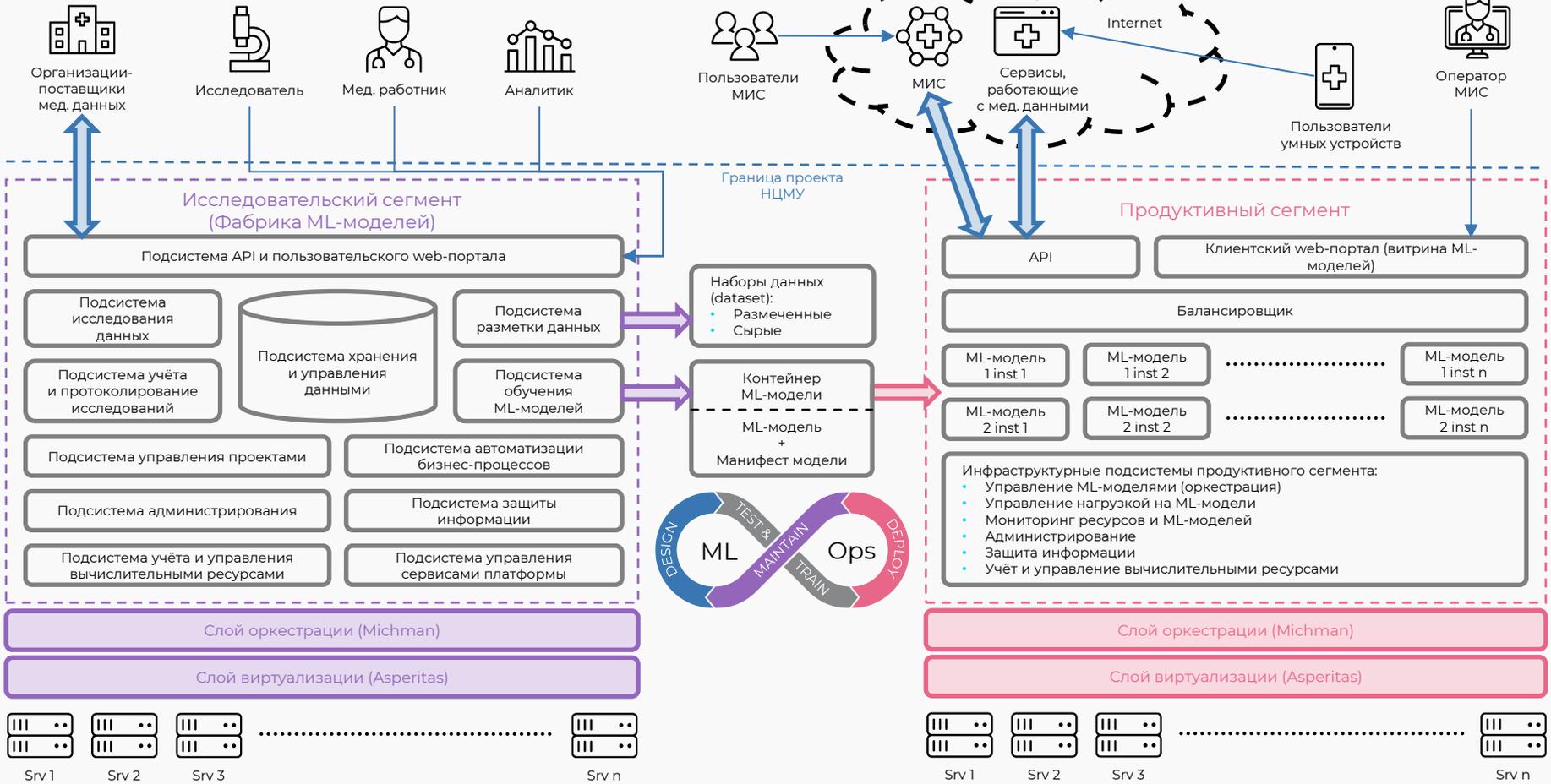
Модели

Быстрая проверка гипотез, проведение экспериментов, получение работоспособных моделей ИИ

Структурная схема Платформы НЦМУ



ИСП РАН





Прикладные инструменты

ИСП РАН Marklab



Streamlit

Технология анализа данных

Сбор открытых данных
Анализ текстов
Анализ графов

Анализ изображений
Анализ временных рядов
Машинное обучение

ИСП РАН

Talisman



python™



PyTorch



Scrapy



NVIDIA

TRITON INFERENCE SERVER

Облачная инфраструктура и системы обработки больших данных

ИСП РАН Asperitas

ИСП РАН Michman

ИСП РАН Fanlight



2023

Опытная эксплуатация Платформы на уровне Первого МГМУ им. И.М.Сеченова, с использованием инструментов Платформы разработаны нейросетевые модели:

- анализа 12-канальных ЭКГ по семи патологиям;
- прогнозирования индивидуального течения онкологических заболеваний.

2024

- Аннотирование 12-канальных ЭКГ врачами.
- Опытная эксплуатация с привлечением всех участников Консорциума (ИБМХ, ИКТИ РАН, НовГУ, Сеченовский университет). Планируется решение следующих задач:
 - анализ 1-канальных ЭКГ;
 - определение вероятности наличия риска злокачественных новообразований;
 - детектирование гипертонической ретинопатии по цифровым изображениям глазного дна;
 - обнаружения лимфоваскулярной инвазии рака легкого на гистологических скан-изображениях рака легкого;
 - предсказание уровня экспрессии белков на основании экспрессии транскриптов.
- Опытная эксплуатация с привлечением внешних участников.

Платформой, по согласованию с ИСП РАН, могут, в проектном формате, воспользоваться исследователи для решения задач в области машинного обучения в медицине и биологии.

Что получает участник?

1. Доступ к сервисам Платформы
2. Вычислительные ресурсы для проведения экспериментов
3. Готовую модель ИИ
4. Опыт совместной работы на единой Платформе

Этапы:



Article | [Open access](#) | [Published: 01 October 2020](#)

Automatic classification of healthy and disease conditions from images or digital standard 12-lead electrocardiograms

[Vadim Gliner](#), [Noam Keidar](#), [Vladimir Makarov](#), [Arutyun I. Avetisyan](#), [Assaf Schuster](#) & [Yael Yaniv](#) 

[Scientific Reports](#) **10**, Article number: 16331 (2020) | [Cite this article](#)

6376 Accesses | **17** Citations | **1** Altmetric | [Metrics](#)

[Home](#) > [Vol 33, No 4 \(2021\)](#) > [ANANEV](#)

Assessment of the impact of non-architectural changes in the predictive model on the quality of ECG classification

 [Vladislav Valerievich ANANEV](#),  [Sergej Nikolaevich SKORIK](#),  [Vsevolod Vladislavovich SHAKLEIN](#),  [Aram Arutyunovich AVETISYAN](#),  [Yurij Emilevich TEREGULOV](#),  [Denis Yuryevich TURDAKOV](#),  [Vadim GLINER](#),  [Assaf SCHUSTER](#),  [Evgeny Andreevich KARPULEVICH](#)

[https://doi.org/10.15514/ISPRAS-2021-33\(4\)-7](https://doi.org/10.15514/ISPRAS-2021-33(4)-7)

Deep Neural Networks Generalization and Fine-Tuning for 12-lead ECG Classification

[Aram Avetisyan](#), [Shahane Tigranyan](#), [Ariana Asatryan](#), [Olga Mashkova](#), [Sergey Skorik](#), [Vladislav Ananev](#), [Yury Markin](#)

Numerous studies are aimed at diagnosing heart diseases based on 12-lead electrocardiographic (ECG) records using deep learning methods. These studies usually use specific datasets that differ in size and parameters, such as patient metadata, number of doctors annotating ECGs, types of devices for ECG recording, data preprocessing techniques, etc. It is well-known that high-quality deep neural networks trained on one ECG dataset do not necessarily perform well on another dataset or clinical settings. In this paper, we propose a methodology to improve the quality of heart disease prediction regardless of the dataset by training neural networks on a variety of datasets with further fine-tuning for the specific dataset. To show its applicability, we train different neural networks on a large private dataset TIS containing various ECG records from multiple hospitals and on a relatively small public dataset PTB-XL. We demonstrate that training the networks on a large dataset and fine-tuning it on a small dataset from another source outperforms the networks trained only on one small dataset. We also show how the ability of a deep neural networks to generalize allows to improve classification quality of more diseases.

Conferences > 2020 Ivannikov Ispras Open Co... 

Non-architectural improvements for ECG classification using deep neural network

Publisher: **IEEE**

[Cite This](#)

[PDF](#)

[Pavel Andreev](#); [Vladislav Ananev](#); [Aram Avetisyan](#); [Vladimir Makarov](#); [Vadim Gliner](#); [Assaf Schuster](#); [Evgeny Karpulevich](#) **All Authors**

76
Full
Text Views



Biomedical Signal Processing and Control

Volume 93, July 2024, 106160



Deep neural networks generalization and fine-tuning for 12-lead ECG classification

[Aram Avetisyan](#)   , [Shahane Tigranyan](#)  , [Ariana Asatryan](#)  , [Olga Mashkova](#)  , [Sergey Skorik](#)  , [Vladislav Ananev](#)  , [Yury Markin](#)  

[Show more](#) 

- Модели обучены на 1,5+ миллионах цифровых ЭКГ из разных регионов (Республика Татарстан, Москва, Великий Новгород) и данных Скорой помощи Москвы.
- В штате ИСП РАН 7 врачей функциональной диагностики.
- Создана собственная система аннотирования ЭКГ.
- Создан сервис нейросетевой классификации ЭКГ ИСП РАН.
- Сервис проходит подготовку к регистрации в качестве медицинского изделия с элементами ИИ.

Список детектируемых синдромов:

1. АВ-блокада 1-ой степени
2. Желудочковая экстрасистолия
3. Наджелудочковая экстрасистолия
4. Неполная блокада правой ножки пучка Гиса
5. Отклонение электрической оси влево
6. Отклонение электрической оси вправо
7. Полная блокада левой ножки пучка Гиса
8. Полная блокада правой ножки пучка Гиса
9. Синусовая брадикардия
10. Синусовая тахикардия
11. Фибрилляция предсердий

Увеличение количества синдромов для предсказания нейронными сетями не является сложной задачей со стороны нейронных сетей, но нужны качественно аннотированные данные

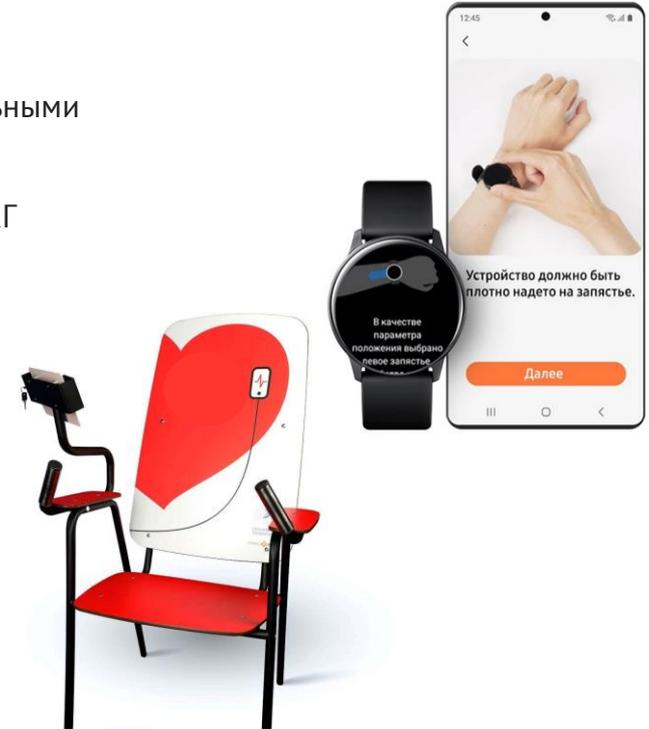


№	Синдром	Чувствительность	Специфичность
1	АВ-блокада 1-ой степени	0.923	0.890
2	Желудочковая экстрасистолия	0.930	0.936
3	Наджелудочковая экстрасистолия	0.930	0.871
4	Неполная блокада правой ножки пучка Гиса	0.890	0.870
5	Отклонение электрической оси влево	0.940	0.870
6	Отклонение электрической оси вправо	0.946	0.925
7	Полная блокада левой ножки пучка Гиса	0.970	0.900
8	Полная блокада правой ножки пучка Гиса	0.980	0.960
9	Синусовая брадикардия	0.970	0.890
10	Синусовая тахикардия	0.939	0.898
11	Фибрилляция предсердий	0.969	0.966

- Результаты работы модели на одноканальных ЭКГ сравнимы с 12-канальными
- Для получения одноканальной ЭКГ не нужен кардиограф
- Всё больше моделей смарт-часов поддерживает сбор одноканальных ЭКГ

Синдром	Чувствительность	Специфичность
Желудочковая экстрасистолия	0.98	0.93
Блокада левой ножки пучка Гиса	0.98	0.91
Синусовая брадикардия	0.95	0.85
Синусовая тахикардия	0.98	0.9
Фибрилляция предсердий	0.97	0.90

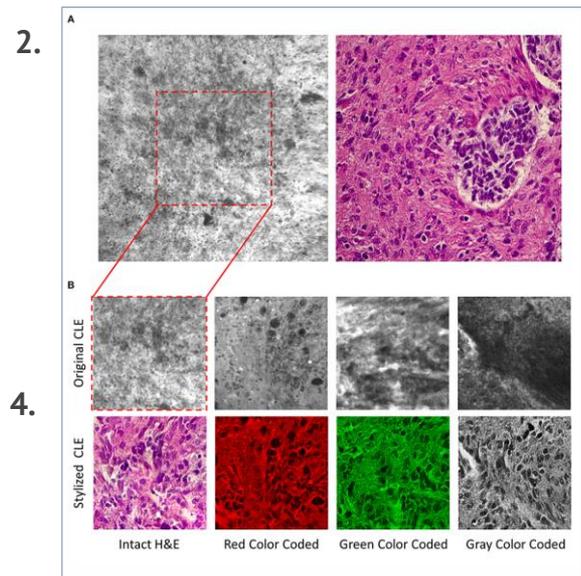
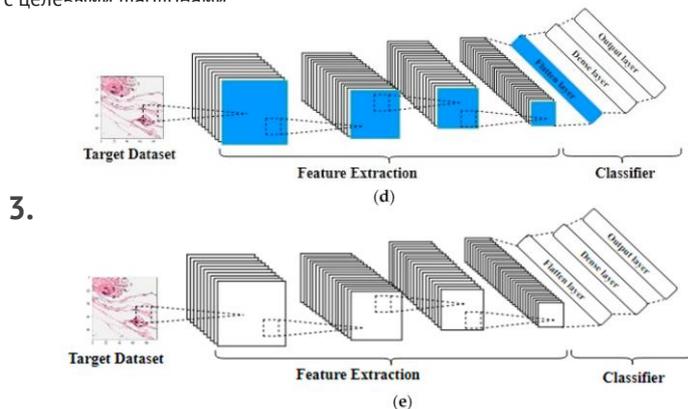
* Данные тестирования на открытом датасете PTB_XL

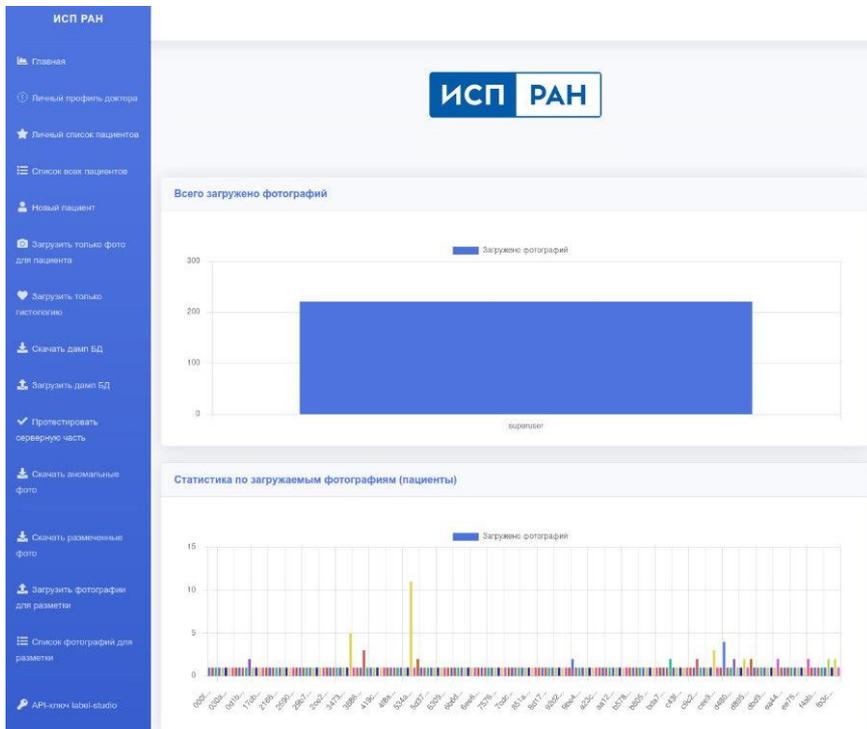


<https://ieeecompsac.computer.org/2024/dicar/>

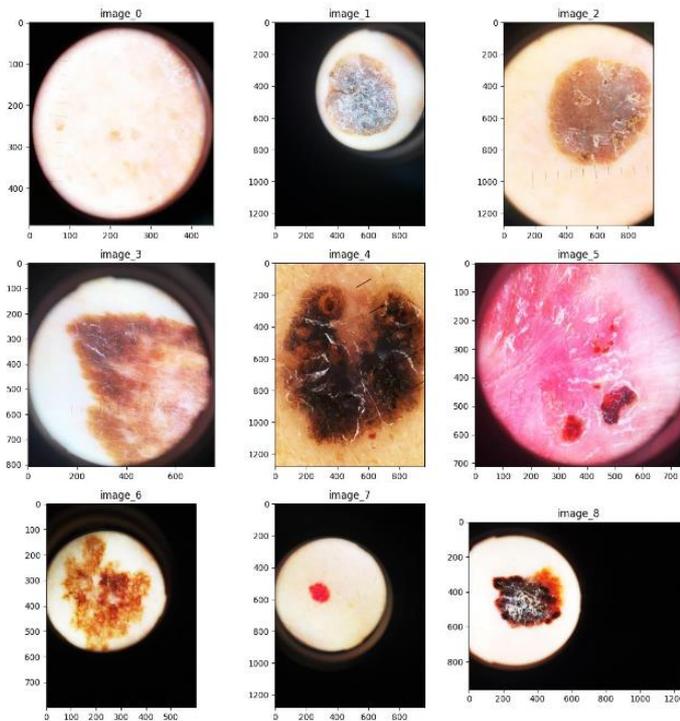
Задача распознавания меланомы на основе анализа меланоцитарных изображений включает следующие подзадачи:

1. **Предобработка изображения** — основная цель заключается в улучшении качества изображения путем уменьшения искажений и артефактов изображения, поиска и улучшения качества признаков изображений.
2. **Обнаружение объекта** — локализация объекта, осуществляемая посредством сегментации изображений и определения положения интересующего объекта.
3. **Извлечение признаков и обучение** — использование статистических методов или методов глубокого обучения для выявления наиболее значимых шаблонов изображения, функций, которые могут быть уникальными для определенного класса изображений и которые впоследствии помогут классификатору различать разные классы.
4. **Классификация изображения** — обнаруженные объекты классифицируются по predetermined классам с использованием подходящего метода классификации, который сравнивает шаблоны изображений с целевыми шаблонами

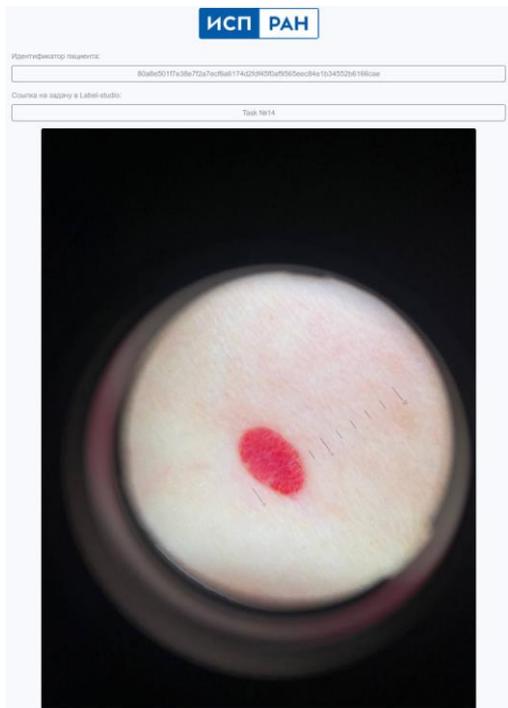




Прототип системы для сбора и разметки дерматоскопических изображений



Примеры из собственной базы дерматоскопических снимков кожи пациентов из РФ



Необходим гистологический анализ:

Нет

Диагноз:

Гемангиома

Вероятность меланомы: [Повторить попытку](#)

Вероятность невуса:

Вероятность актинического кератоза:

Вероятность базально клеточной карциномы:

Вероятность кератозного поражения:

Вероятность дерматофибромы:

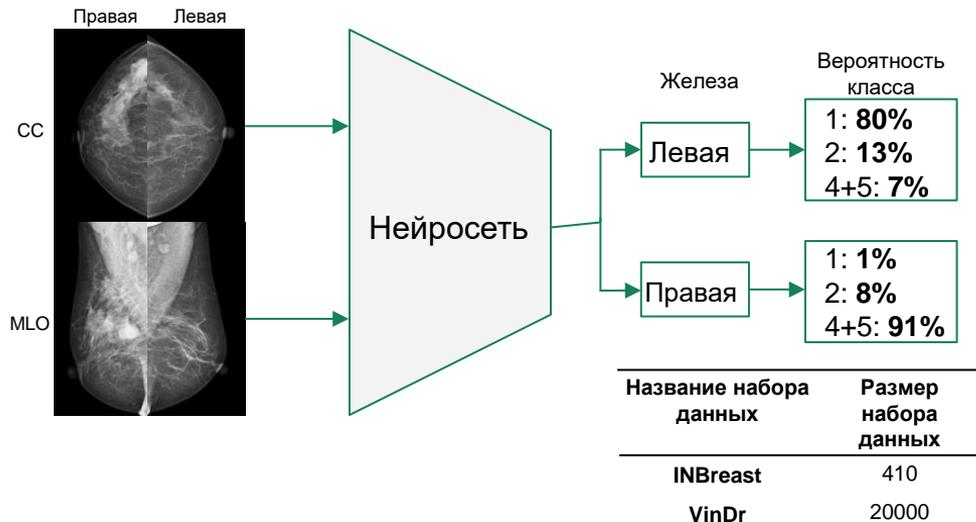
Вероятность рака кожи:

Вероятность сосудистых поражений:

Вероятность неизвестного диагноза:

Время создания:

Интерфейс взаимодействия с системой автоматического анализа дерматоскопических изображений и результат анализа загруженного изображения



Классы:

- 1 - Норма
- 2 - Доброкачественная находка
- 4+5 - Злокачественная находка

Метрики качества обученной модели на датасетах VinDr и INBreast (макро усреднение):

F1 score: **0.6818**
 Precision: **0.7936**
 Recall: **0.629**
 Specificity: **0.8308**

Публикации:

- Senotrusova, S., Ibragimov, A., Litvinov, A., Beliaeva, A., Ushakov, E., & Markin, Y. (2024). Classification of the presence of malignant lesions on mammogram using deep learning. *Digital Diagnostics*, 5(1S), 137-139.
- Ibragimov, A., Senotrusova, S., Litvinov, A., Ushakov, E., Karpulevich, E., & Markin, Y. (2024, July). MamT 4: Multi-View Attention Networks for Mammography Cancer Classification. In *2024 IEEE 48th Annual Computers, Software, and Applications Conference (COMPSAC)* (pp. 1965-1970). IEEE.

РИДы:

- Свидетельства о регистрации ПО 2024613811 (15.02.2024) «Программа детектирования злокачественных образований в маммографических снимках на основе нейросетевой модели»
- Свидетельства о регистрации ПО 2024614104 (20.02.2024) «Программа бинарной классификации маммографических снимков на основе нейросетевой модели»

Открытая конференция ИСП РАН (5-6 декабря 2024 г.)

www.isprasopen.ru

С 2016 г. в Москве при поддержке IEEE, IEEE Computer Society, Samsung, Huawei и др. Более 1000 участников и 100 докладчиков.

«Иванниковские чтения» (17-18 мая 2024, Великий Новгород)

www.ivannikov-ws.org

С 2018 г. в разных городах при поддержке IEEE, IEEE Computer Society и др. Более 200 участников и 40 докладчиков.

«Анализ данных в медицине» (17-18 мая 2024, Великий Новгород)

digitalmed.ru

С 2021 г. в разных городах. В числе организаторов: Сеченовский университет, МНОЦ МГУ, НИИ АГиР им. Д.О. Отта и др.

SYRCoSE (Весенняя конференция молодых ученых по программной инженерии) (29-31 мая, Ставрополь)

С 2007 г. в разных городах.

Более 30 докладчиков. Язык конференции – английский.

OS DAY (20-21 июня 2024, Москва)

osday.ru

С 2014 г. Организаторы: DZ Systems, РусБИТех, Лаборатория Касперского и др. Более 30 докладчиков.

Круглый стол «Системное программирование как ключевое направление противодействия киберугрозам» на форуме «Армия»

С 2018 г. Более 100 участников и 20 докладчиков.



Спасибо за внимание!

Приглашаем к сотрудничеству
и опытной эксплуатации!

ИСП РАН

Андрей Бурсов
bursov@ispras.ru