

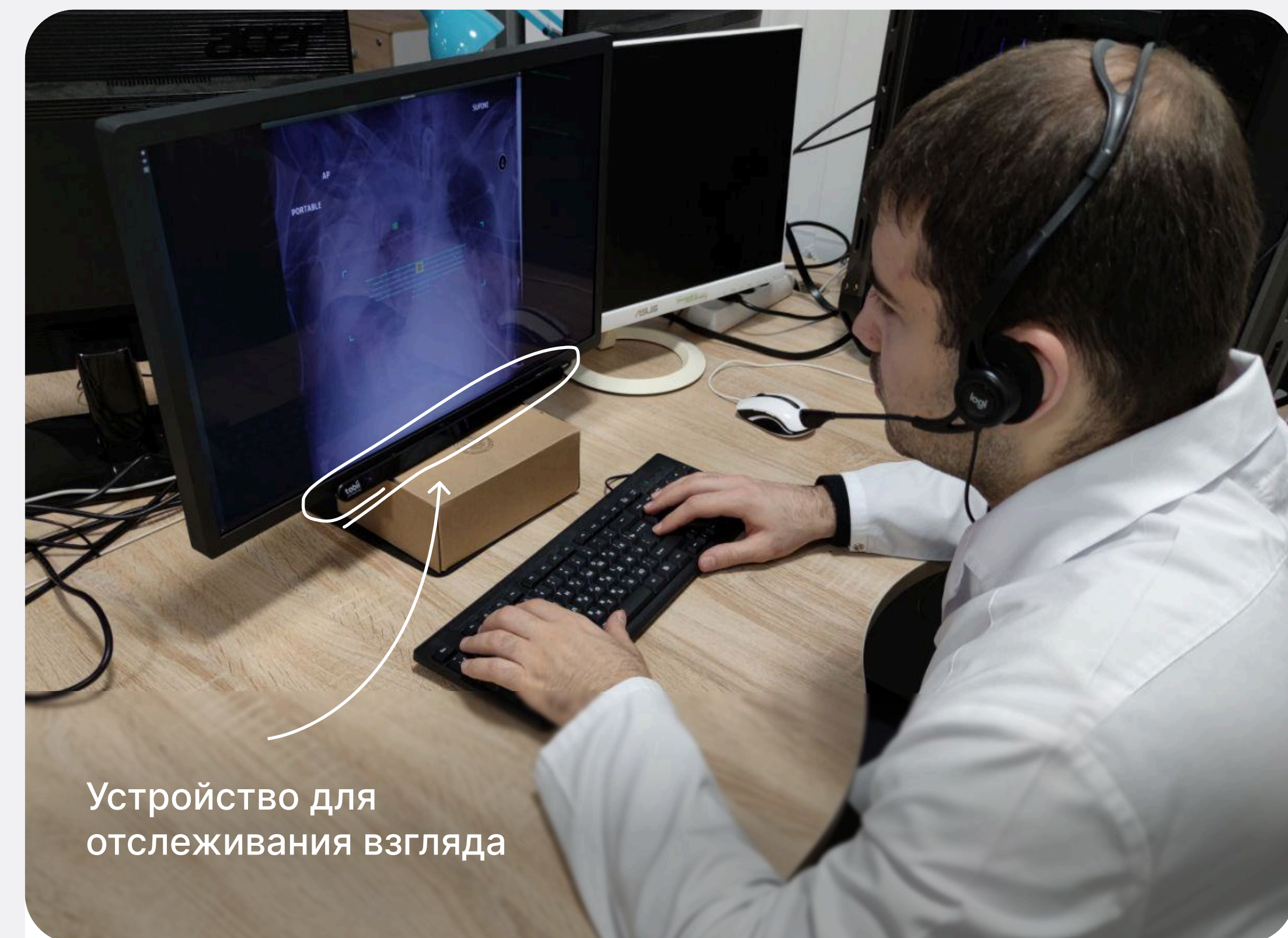
# ОТ ВНИМАНИЯ ВРАЧА К ДИАГНОСТИКЕ: ИНТЕГРАЦИЯ СЕМАНТИЧЕСКОЙ ИНФОРМАЦИИ ДЛЯ ПРОГНОЗИРОВАНИЯ ВЗГЛЯДА ВРАЧА

Першин Илья (докладчик)  
Львов Дмитрий

Лаборатория ИИ в медицине  
Университет Иннополис

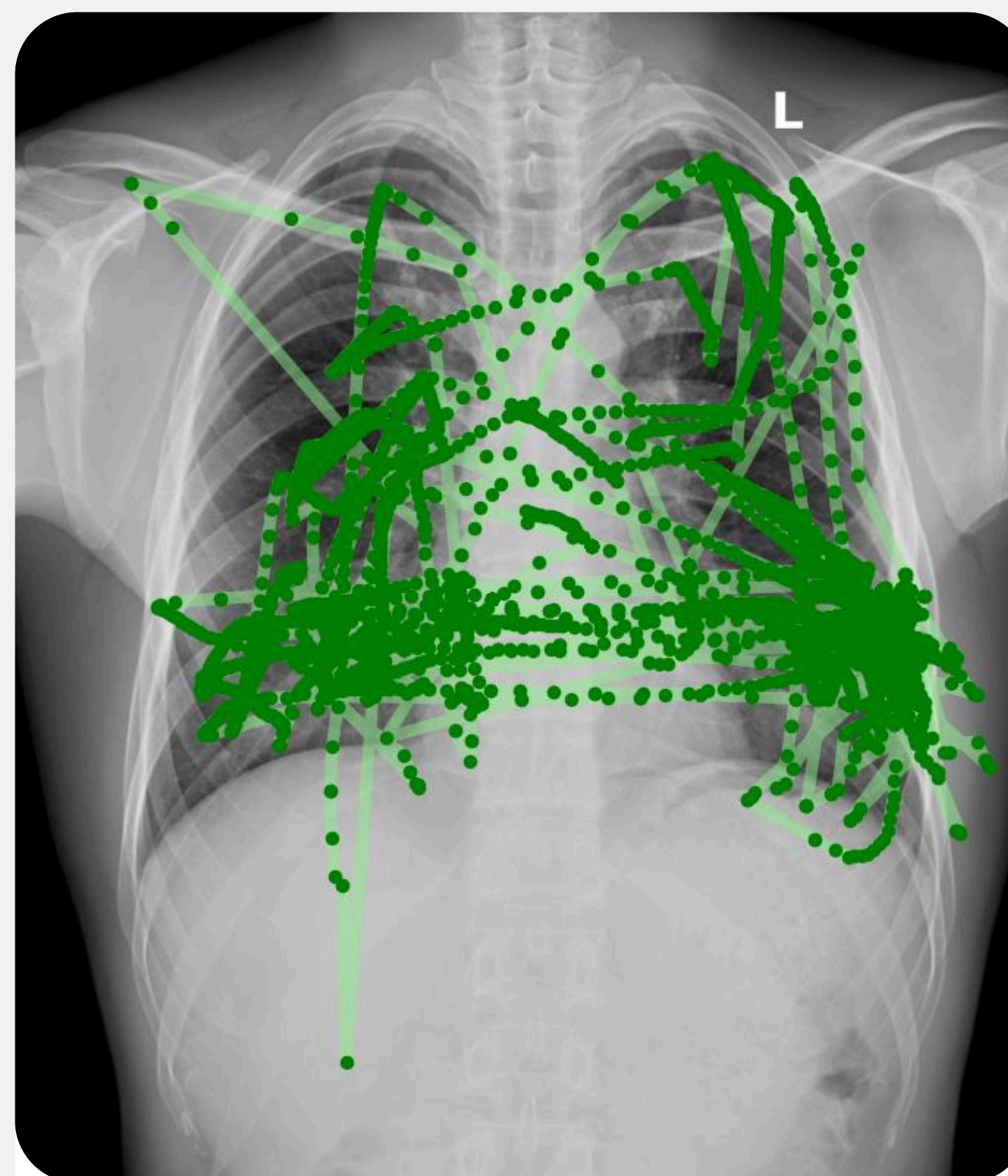
Математика в медицине  
Сеченовский университет, 2025



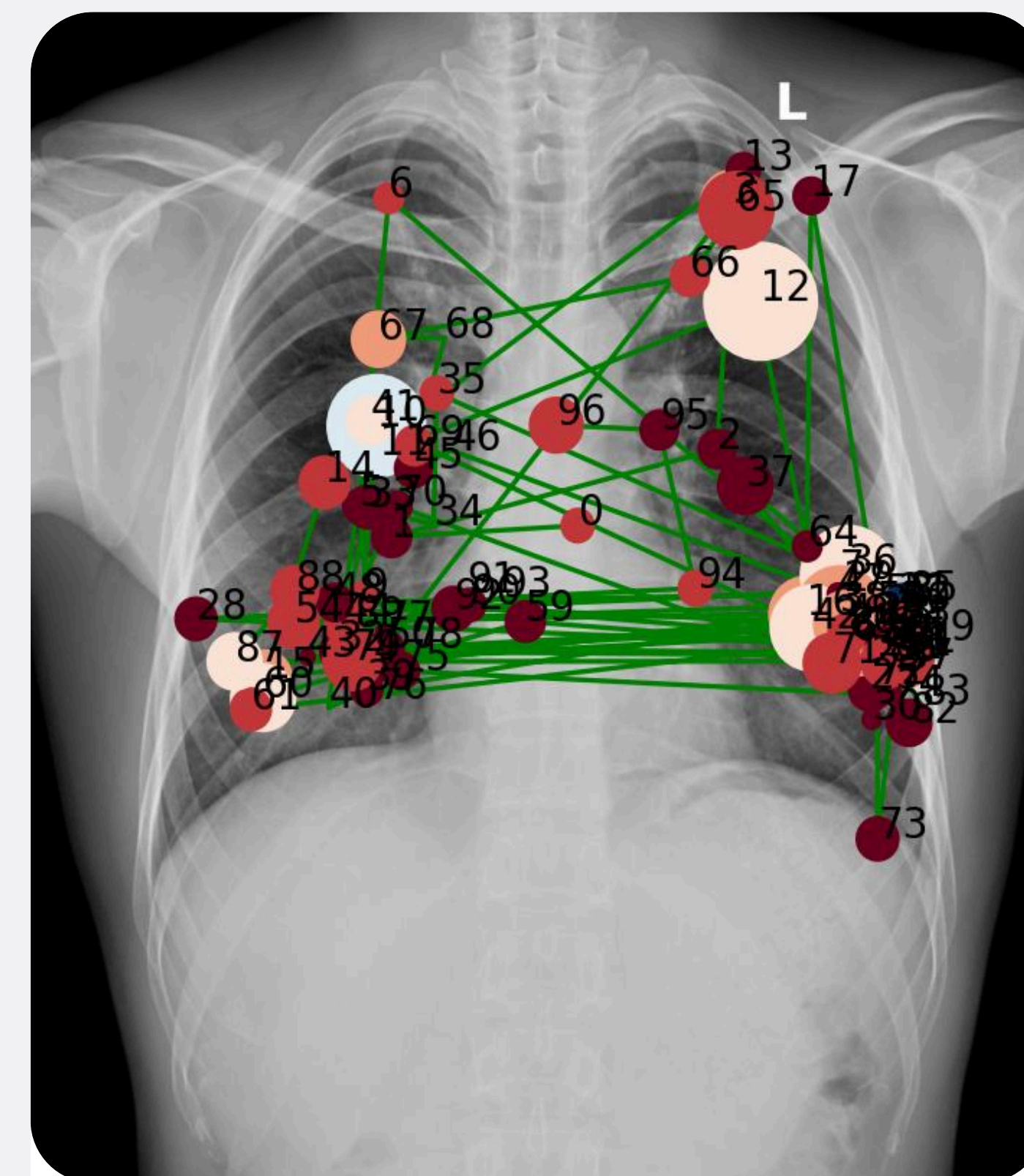


Устройство для  
отслеживания взгляда

Анализ рентгена

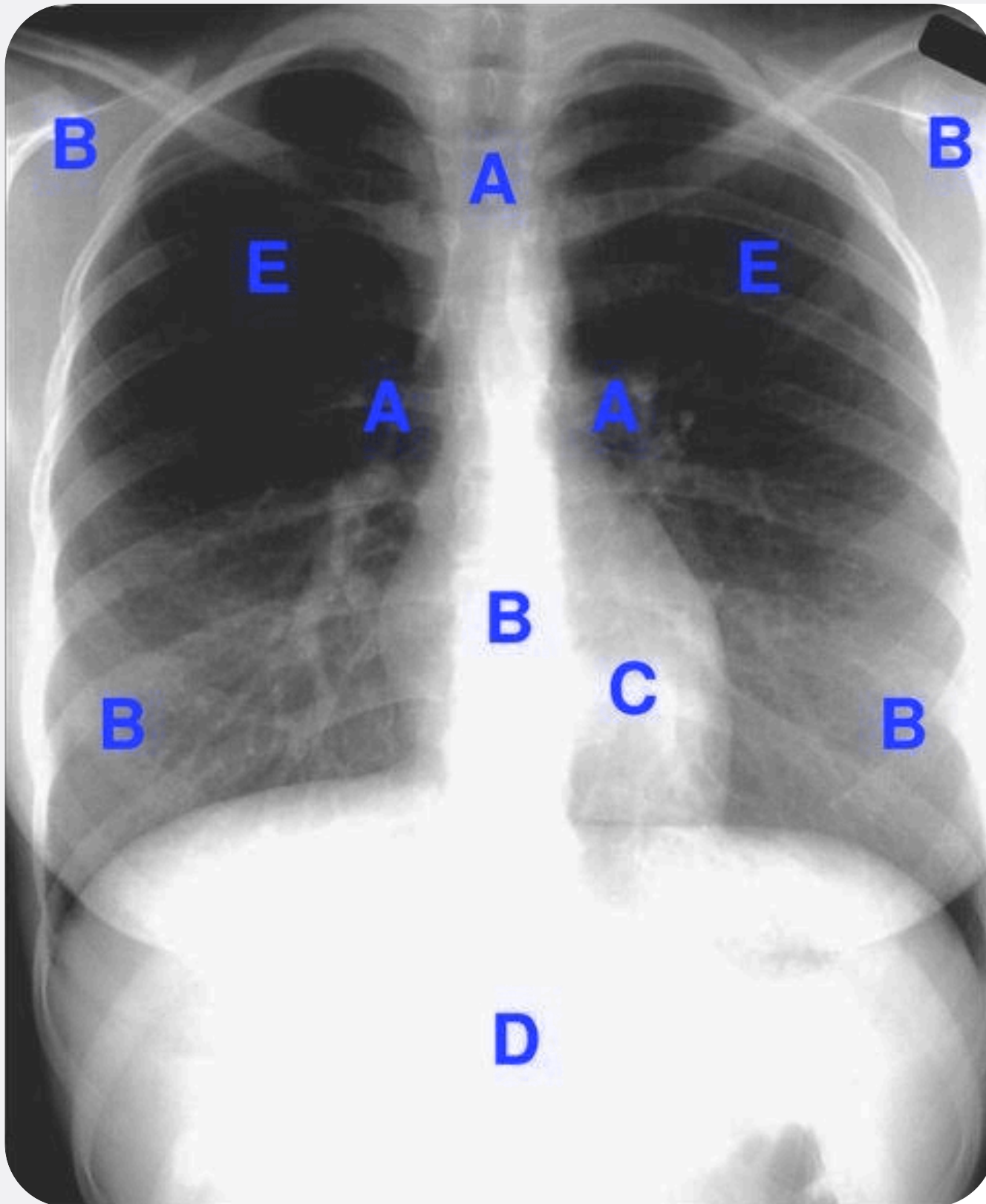


Исходные точки  
взгляда



Точки фиксации  
и саккад



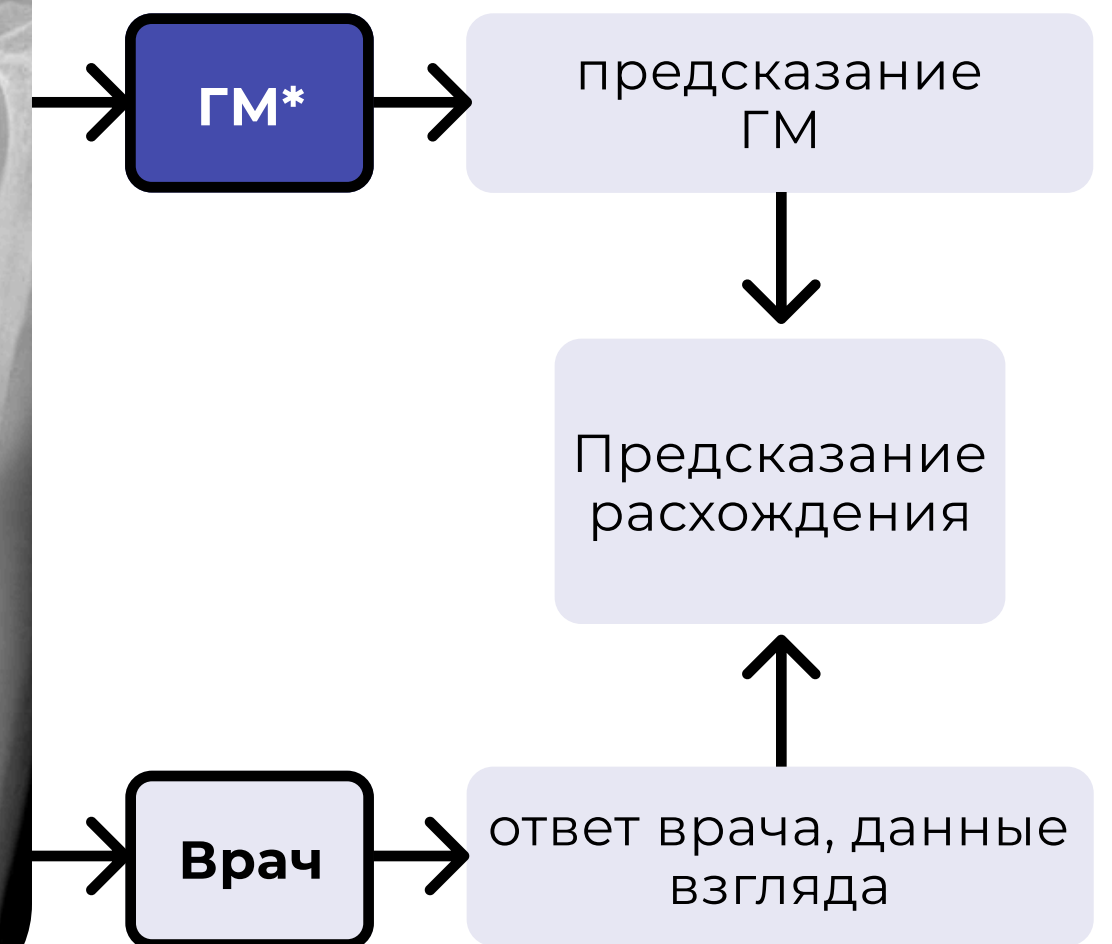


## Протокол ABCDE

- A** (airway) — воздушные пути
- B** (bones) — кости
- C** (cardiac) — сердце
- D** (diaphragm) — диафрагма
- E** (everything else = lungs) — все остальное = лёгкие



ГМ\* - глубокая модель



## Наборы данным со взглядом на рентгеновских снимках грудной клетки:

**Eye-Gaze** [2021] - 1083 снимка

**REFLACX** [2021] - 2616 снимков

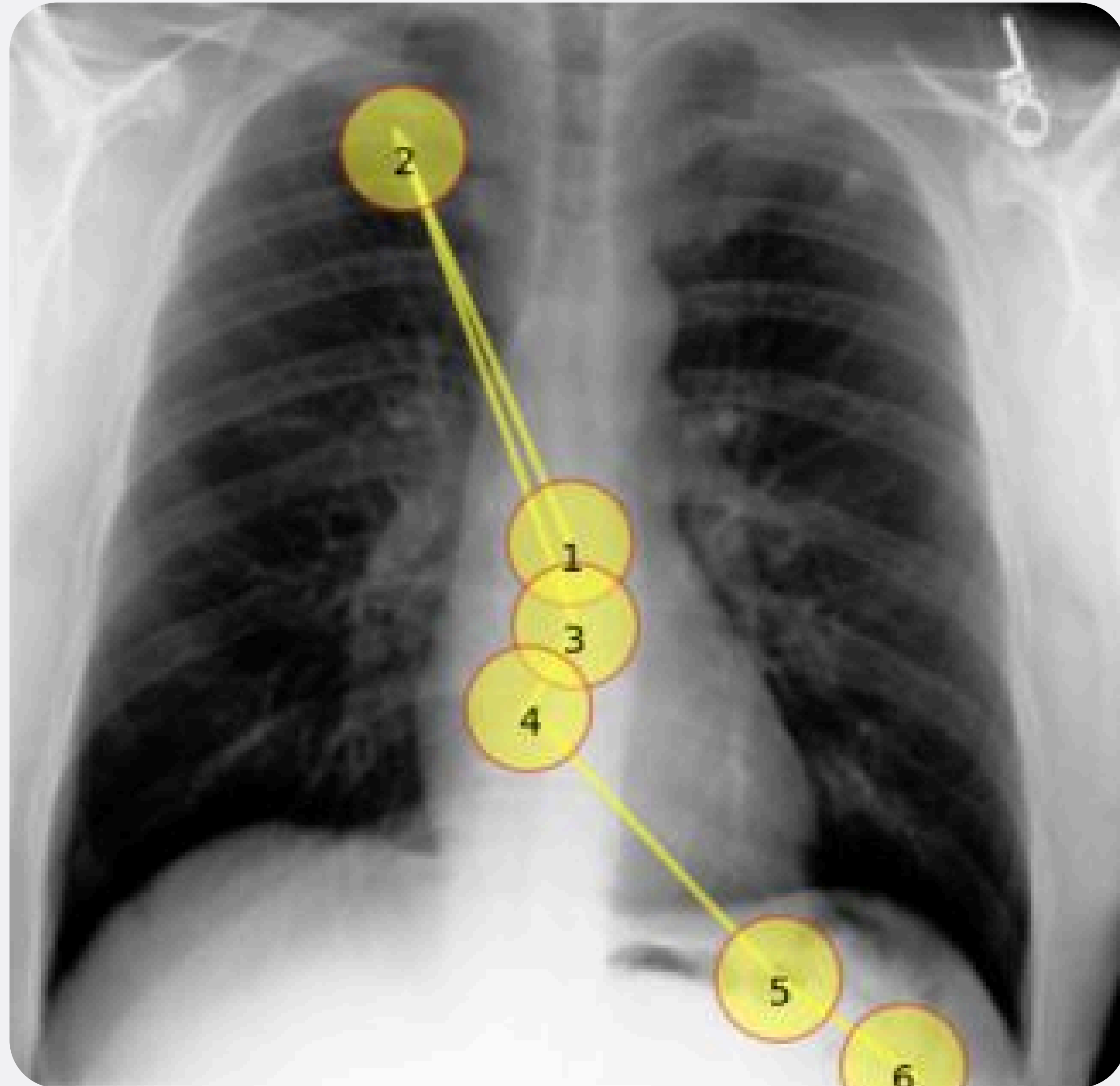
**FG-CXR** [2024] - 2951 снимков

**Итого:** 6 650 снимков

## ЗАДАЧА 1.

3

ПРЕДСКАЗАНИЕ ТРАЕКТОРИИ  
ВЗГЛЯДА ЧЕЛОВЕКА НА  
ИЗОБРАЖЕНИИ



# ЗАДАЧА 1. НАБОРЫ ДАННЫХ И МОДЕЛИ ЦЕЛЕНАПРАВЛЕННОГО ВЗГЛЯДА

4

## Наборы данных

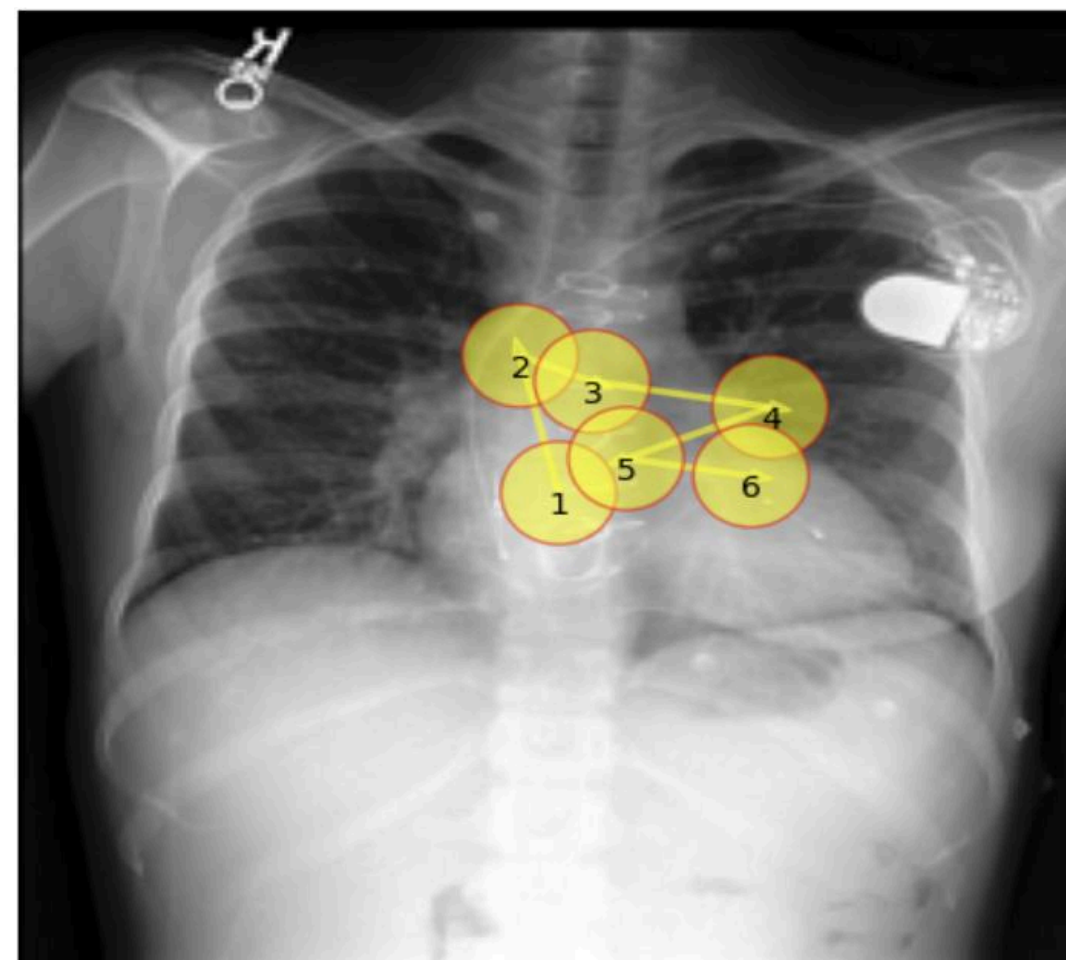
Объект из **COCO-Search18**.  
Цель — автомобиль.



**COCO-Search18** содержит 6202 изображений с разбиением:

- обучение - 70%, 4341
- валидация - 10%, 620
- тестирование - 20%, 1240

Объект из **GazeSearch**.  
Цель - кардиомегалия.



**GazeSearch** содержит 2081 изображения с разбиением:

- обучение - 70%, 1457
- валидация - 10%, 208
- тестирование - 20%, 416

## SOTA модели

**GazeFormer** [2023] - COCO-Search18

**HAT** [2023] - COCO-Search18

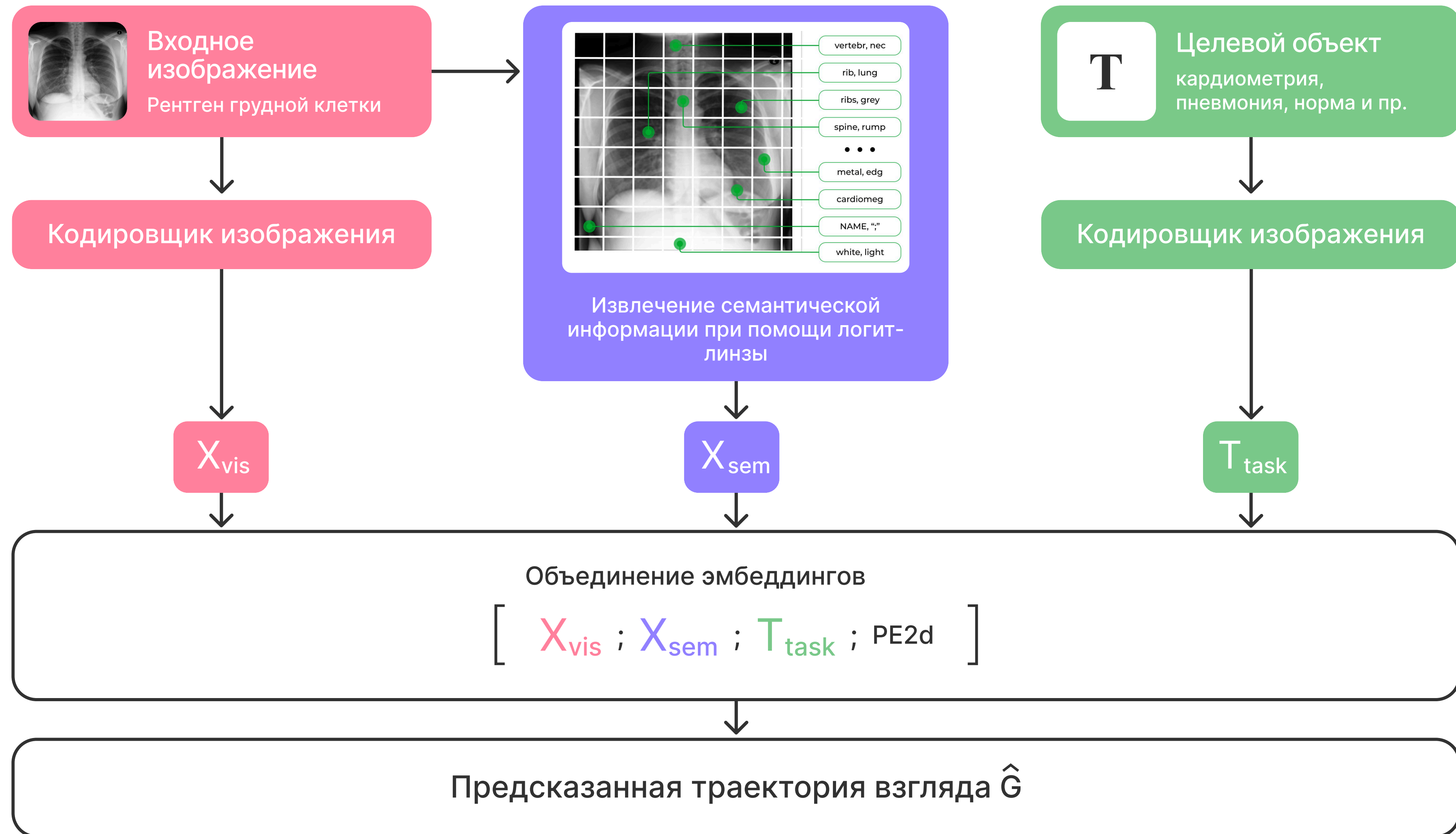
**GazeSearch** [2024] - GazeSearch (рентгены)

В [1] адаптировали метод логит-линзы для VLM.

[1] C. Neo, L. Ong, P. Torr, M. Geva, D. Krueger, and F. Barez, 'Towards interpreting visual information processing in vision-language models', ICLR 2025.

# ЗАДАЧА 1. АРХИТЕКТУРА LOGITGAZE

5





Предположения

- 1. Использование результатов логит-линзы в качестве доп. входа для **GazeFormer** улучшит качество на **COCO18-search**
- 2. Применение логит-линзы для **LLaVa-Med** позволит добиться лучших результатов на **GazeSearch**, чем с LLaVa
- 3. Использование медицинского кодировщика изображений позволит повысить качество результатов на **GazeSearch**

Результаты

Значения метрик схожести траекторий взгляда (чем выше, тем лучше для ScanMatch/STDE, чем ниже, тем лучше для SED) на **GazeSearch**.

Метод	ScanMatch $\uparrow$		SED $\downarrow$	STDE $\uparrow$
	w/o Dur.	w/ Dur.		
GazeFormer	$0.293 \pm 0.021$	$0.201 \pm 0.015$	$5.11 \pm 0.08$	$0.799 \pm 0.004$
HAT	$0.309 \pm 0.020$	–	$5.07 \pm 0.07$	$0.800 \pm 0.004$
GazeSearch	$0.332 \pm 0.019$	$0.223 \pm 0.014$	$4.88 \pm 0.06$	$0.809 \pm 0.004$
LogitGaze	$0.328 \pm 0.018$	$0.225 \pm 0.015$	$5.07 \pm 0.07$	$0.810 \pm 0.004$
LogitGaze-Med (Res)	$0.416 \pm 0.017$	$0.325 \pm 0.012$	$4.68 \pm 0.05$	$0.852 \pm 0.003$
<b>LogitGaze-Med (CheX)</b>	<b><math>0.419 \pm 0.016</math></b>	<b><math>0.330 \pm 0.010</math></b>	<b><math>4.68 \pm 0.05</math></b>	<b><math>0.855 \pm 0.003</math></b>

Тестирование **LogitGaze** на наборе данных **COCO-Search18**.

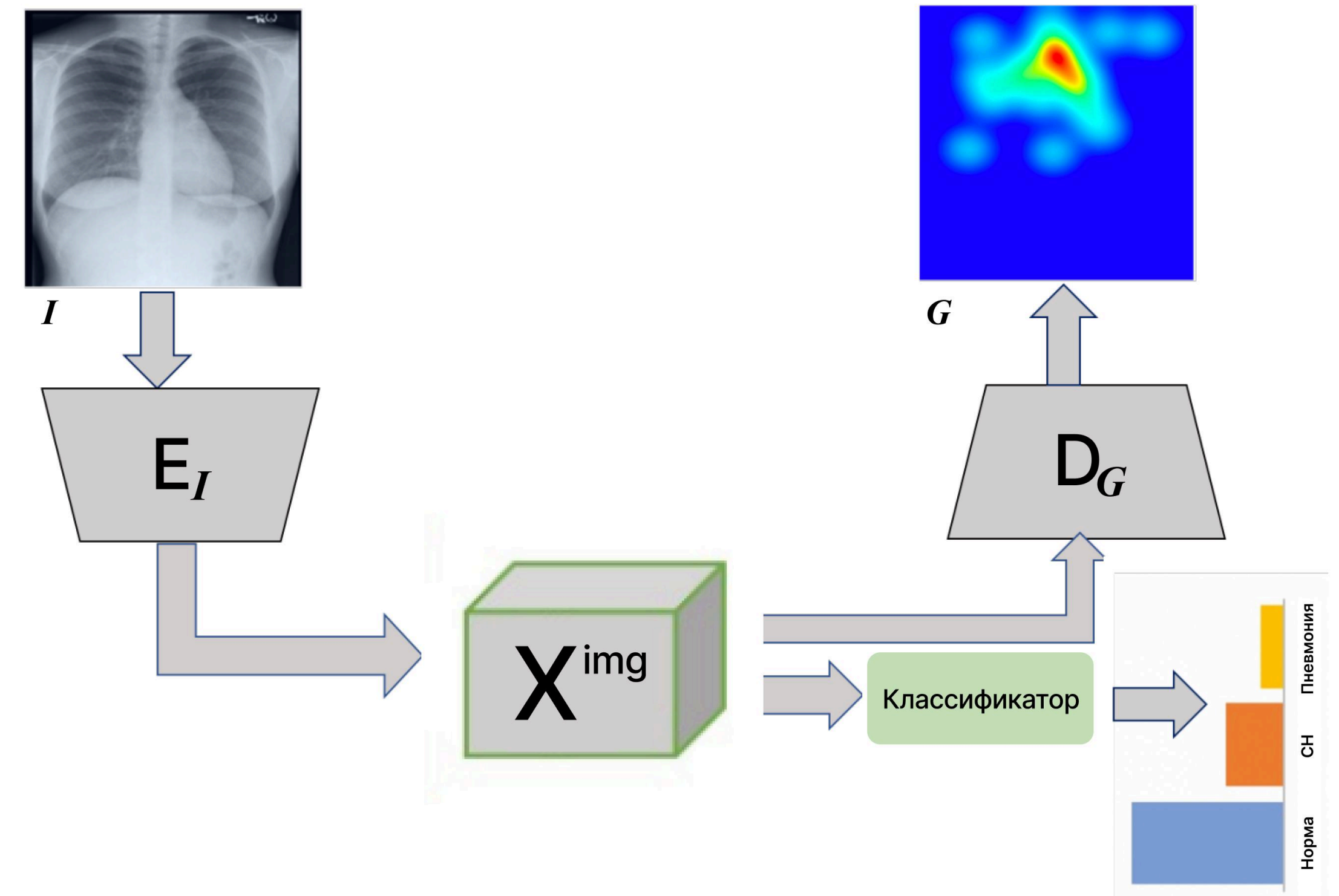
Метод	ScanMatch	ScanMatch
	w/o Dur	w/ Dur
LogitGaze	<b><math>0.527 \pm 0.012</math></b>	<b><math>0.454 \pm 0.013</math></b>
GazeFormer [48]	$0.492 \pm 0.014$	$0.441 \pm 0.015$
LogitGaze-Med	$0.353 \pm 0.018$	$0.289 \pm 0.020$

Вывод: SOTA результат на медицинском и общем домене.

## ЗАДАЧА 2.

7

### ВНЕДРЕНИЕ ВЗГЛЯДА РЕНТГЕНОЛОГА В ГЛУБОКУЮ МОДЕЛЬ КЛАССИФИКАЦИИ ПАТОЛОГИЙ ЛЕГКИХ





Предположения

- В [1] представили
- набор данных взгляда (1083 снимка);
  - метод для улучшения классификации при помощи взгляда рентгенолога
- ✓

1. Замена натуральных данных взгляда на синтетические в методе [1] позволит еще сильнее улучшить результаты
- ✓

2. Использование медицинской модели **LogitGaze-Med** для генерации взгляда лучше, чем **LogitGaze**

Метод проверки

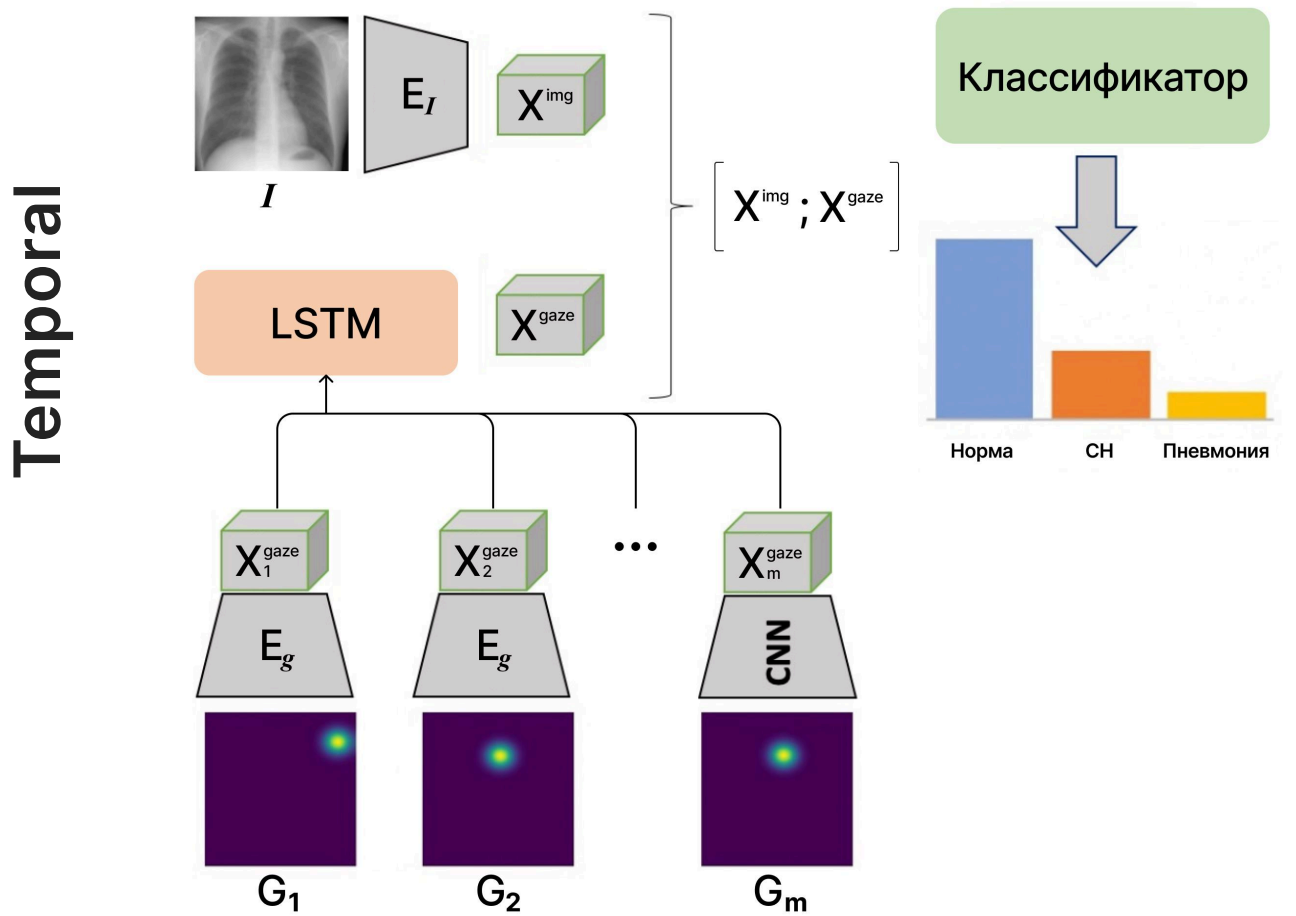
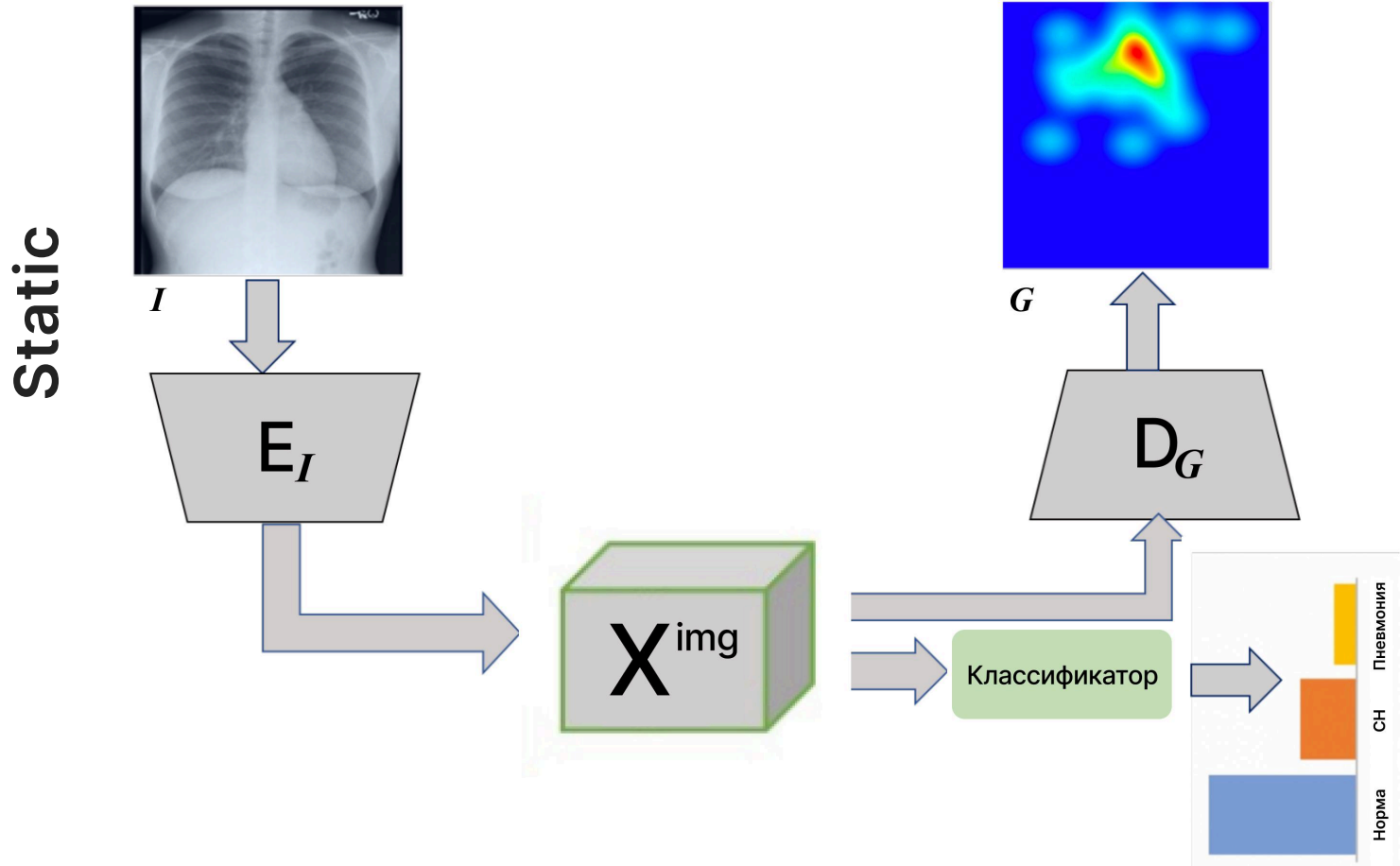
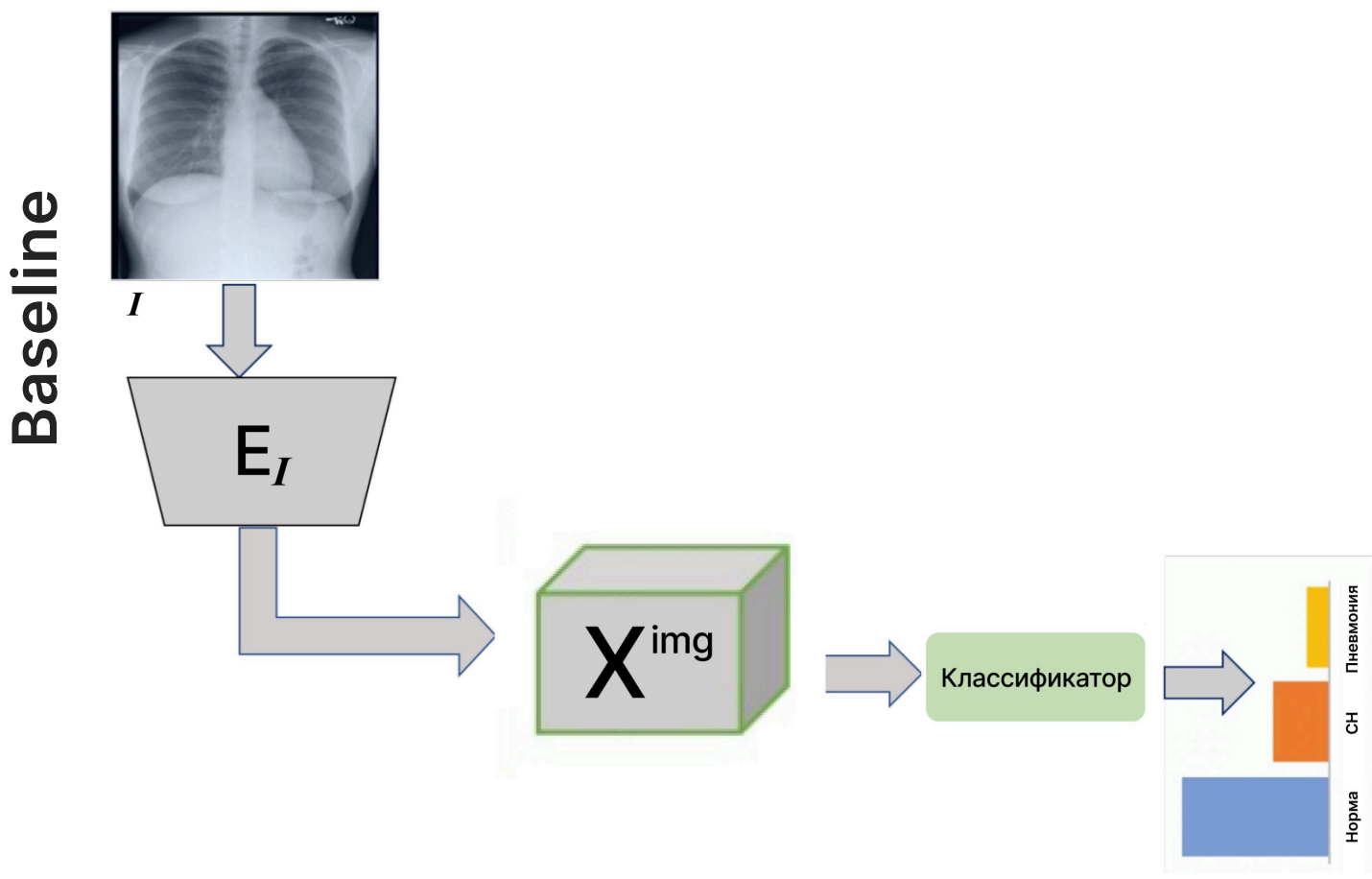
1. Создать синтетический набор данных взгляда (**LogitGaze-Med, LogitGaze, GazeFormer**) на основе **MIMIC-CXR** (227 тыс. рентгенов);
2. Использовать синтетический набор как тренировочный в эксперименте [1];
3. Провести тестирование на тестовом наборе данных из [1].

Результаты

Среднее значение AUROC по патологиям: норма, застойная сердечная недостаточность и пневмония).

Взгляд	Baseline	Temporal	Static
Eye-Gaze [11]	$0.77 \pm 0.02$	$0.82 \pm 0.03$	$0.87 \pm 0.02$
GazeFormer [9]	$0.78 \pm 0.02$	$0.84 \pm 0.02$	$0.89 \pm 0.01$
LogitGaze	$0.80 \pm 0.01$	$0.87 \pm 0.02$	$0.90 \pm 0.01$
<b>LogitGaze-Med</b>	<b><math>0.82 \pm 0.01</math></b>	<b><math>0.90 \pm 0.02</math></b>	<b><math>0.91 \pm 0.01</math></b>

[1] A. Karargyris et al., “Creation and validation of a chest X-ray dataset with eye-tracking and report dictation for AI development,” Sci Data, vol. 8, no. 1, Mar. 2021, doi: [10.1038/s41597-021-00863-5](https://doi.org/10.1038/s41597-021-00863-5).



Результаты приняты на main track NeurIPS 2025 (A\*).

# СПАСИБО ЗА ВНИМАНИЕ!

Илья Першин  
Университет Иннополис  
[i.pershin@innopolis.ru](mailto:i.pershin@innopolis.ru)

