

Система компьютерного зрения для детектирования опасных ситуаций на производстве

Фомин Никита Алексеевич, М26 группа

Научный руководитель:

к.ф.-м.н., доцент Солдатенко Илья Сергеевич

Тверской государственный университет

Факультет прикладной математики и кибернетики

Кафедра информатики и информационных технологий

Тверь — 2024

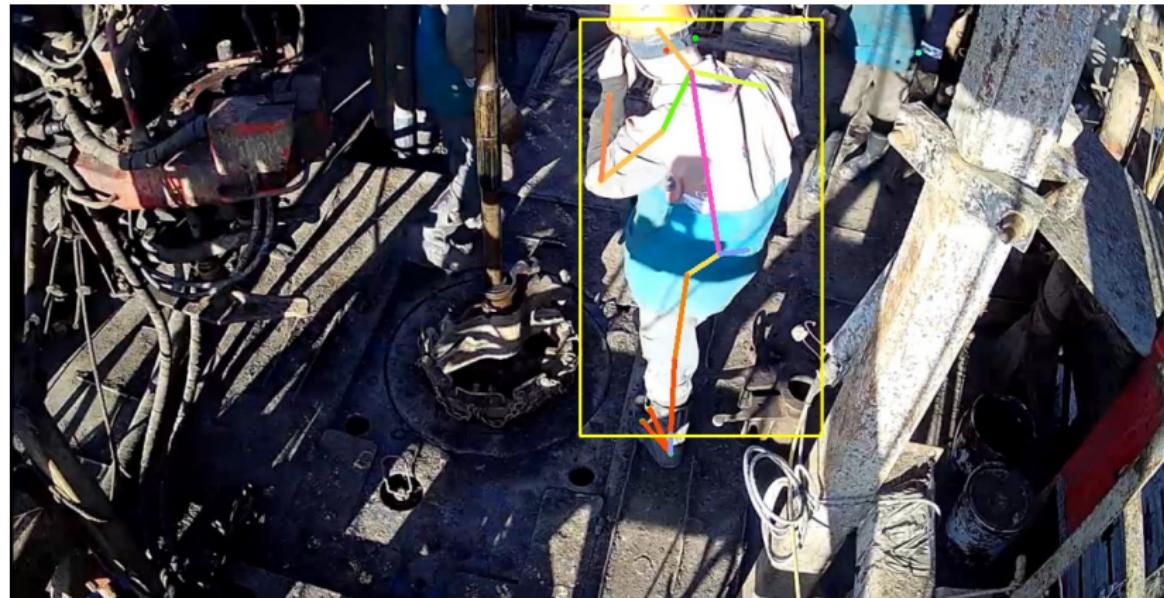
Постановка задачи



Постановка задачи



Постановка задачи



Постановка задачи

Начальные предположения:

- Опасная зона для конкретной камеры задана заранее;
- Детектируем только случай нахождения в опасной зоне без СИЗ.

Список рассматриваемых средств индивидуальной защиты:

- Каска,
- Перчатки,
- Ботинки.

Архитектура решения



Цели и задачи

Цель работы – проектирование системы компьютерного зрения для детектирования опасных ситуаций на производстве.

Задачи работы:

- ① Проектирование архитектуры системы компьютерного зрения,
- ② Разработка метода распознавания и скелетизации,
- ③ Реализация алгоритма детектирования нахождения в опасной зоне,
- ④ Описание принципа принятия решений по заданному кадру.

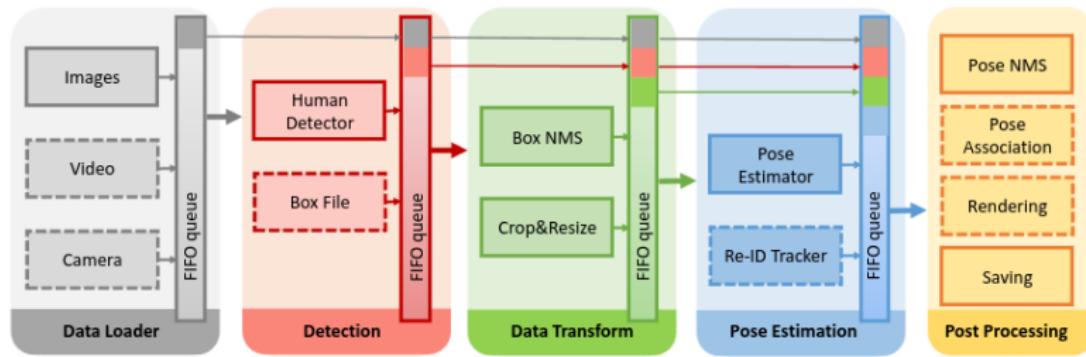
Распознавание людей и их позиций



Скелетизация



Модель AlphaPose



Выходной формат данных скелетной модели

По заданному кадру алгоритм строит множество S скелетных моделей.

Каждый элемент S имеет вид:

$$(ID, K, BB),$$

где:

- ID – уникальный идентификатор человека в кадре,
- K – список ключевых точек скелетной модели,
- BB – ограничивающая рамка для изображения человека.

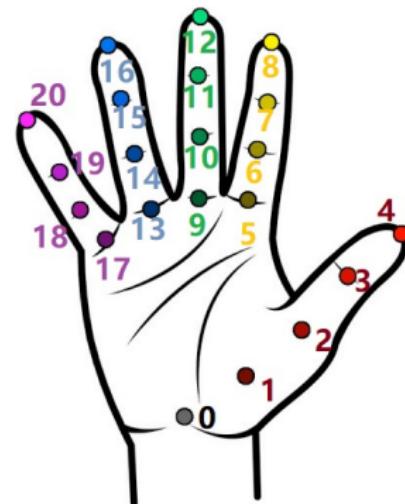
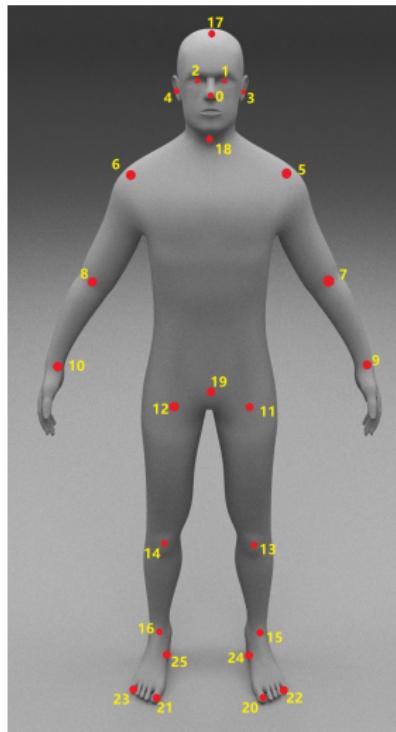
Формат представления ключевых точек

$$K = [(x_i, y_i, c_i)]_{i=0}^{N-1},$$

где

- (x_i, y_i) – координаты ключевой точки на изображении,
- c_i – мера уверенности алгоритма, что ключевая точка предсказана корректно,
- N – общее число ключевых точек (26 тело, 21 левая рука, 21 правая рука).

Нумерация ключевых точек



Рассчет координат и меры уверенности для ключевой точки

Пусть $z_{(x,y)}$ — ненормализованные логиты, выход backbone-части регрессора.

Определим карту теплоты $C = (c_{(x,y)})$ следующим образом:

$$c_{(x,y)} = \sigma(z_{(x,y)})$$

Мера уверенности для данной точки выражается как:

$$conf = \max(C)$$

Нормализованная карта теплоты $P = (p_{(x,y)})$ выражается следующим образом:

$$p_{(x,y)} = \frac{c_{(x,y)}}{\sum C}$$

Простейшее определение опасной зоны

В простейшем случае, опасную зону можно определить следующим образом.

Зафиксируем выпуклый многоугольник A на изображении.

$$x_{min} = \min\{x : \exists y(x, y) \in A\}$$

$$x_{max} = \max\{x : \exists y(x, y) \in A\}$$

Тогда опасной зоной DZ будем называть выпуклую оболочку множества

$$A \cup \{(x_{min}, 0), (x_{max}, 0)\}$$

Определение принадлежности опасной зоне

Задан минимальный порог срабатывания m для принадлежности ключевой точки зоне.

По определению $I_i = 1$, если ключевая точка человека (x_i, y_i, c_i) находится в опасной зоне.

$$I_i = 1 \iff (c_i \geq m) \wedge ((x_i, y_i) \in DZ)$$

Задана минимальная доля t ключевых точек, принадлежащих опасной зоне, при которой считаем человека находящимся в ней.

Считаем, что человек находится в опасной зоне, если:

$$\frac{\sum I_i}{N} \geq t$$

Усложнение определения опасной зоны

Будем считать, что область пространства, обозреваемого камерой, локально представляет из себя \mathbb{R}^3 .

Зафиксируем правильный многоугольник A в плоскости $z = 0$. Тогда опасной зоной будем называть множество:

$$DZ = \{(x, y, z) : (x, y, 0) \in A\}$$

Детектирование нахождения в опасной зоне



Оценка расстояния до объектов

Рассмотренные подходы к оценке расстояния до объектов:

- Калибровка камеры;
- Анализ геометрии помещения;
- Аппроксимация карты глубины.

Карта глубины

Определение

Карта глубины — это изображение, где для каждого пикселя вместо цвета хранится расстояние от него до камеры.

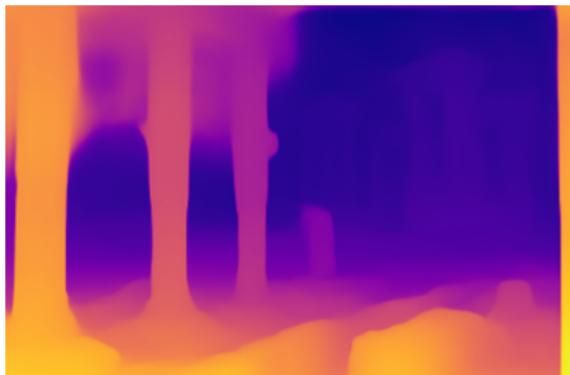
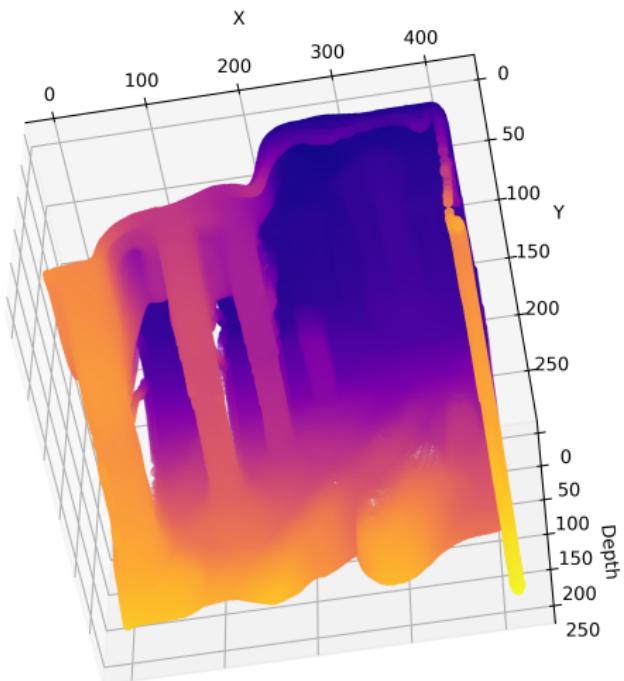
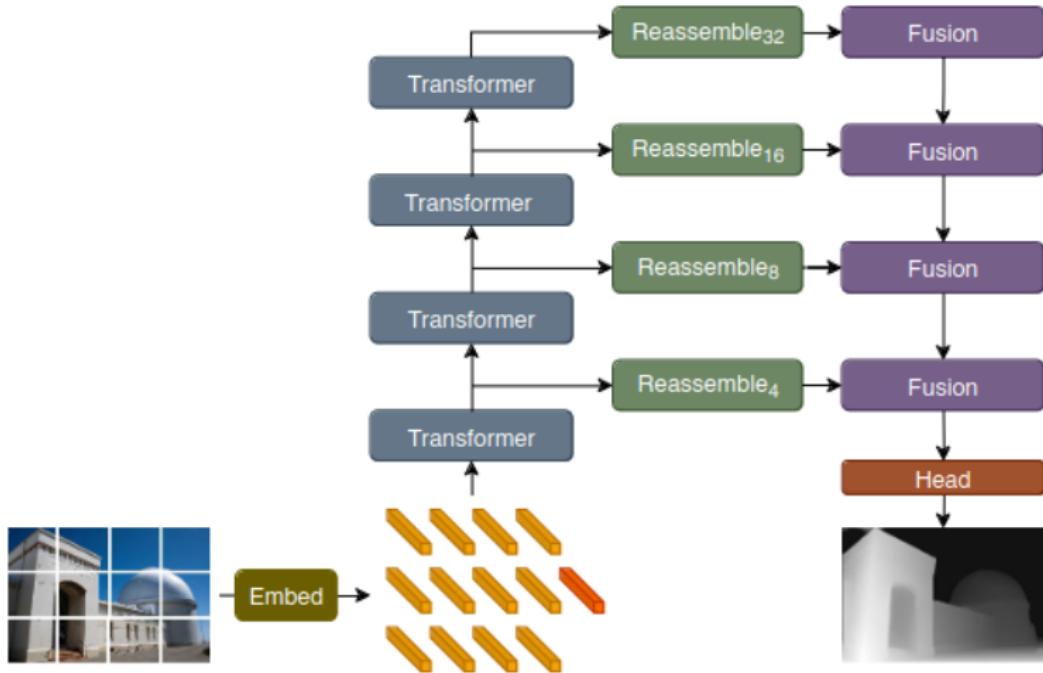


Рис. 1: Пример карты глубины, построенной по изображению

Трёхмерная реконструкция по карте глубины



Модель MiDaS



Получение карты абсолютной глубины

Модель Midas строит по изображению карту относительной глубины M_r . Преобразуем её в карту абсолютной глубины M_a .

Зная точное расстояние d_0 до одной из точек (x_0, y_0) на изображении можно масштабировать карту относительной глубины следующим образом:

$$M_a = c \cdot M_r,$$

где $c = \frac{d}{M_r[x_0, y_0]}$

Уточнение карты абсолютной глубины

Пусть мы знаем расстояния $d = (d_0, \dots, d_{k-1})$ до точек $A = ((x_0, y_0), \dots, (x_{k-1}, y_{k-1}))$.

Определим функцию ошибки приближения следующим образом:

$$L_{(d,A)}(c) = \sum_{i=0}^{k-1} (c \cdot M_r[x_i, y_i] - d_i)^2$$

Методом наименьших квадратов найдём такое c , при котором достигается минимум $L(c)$.

Примеры работы алгоритма проверки принадлежности опасной зоне



Рис. 2: Рабочий в СИЗ не в опасной зоне

Примеры работы алгоритма проверки принадлежности опасной зоне



Рис. 3: Рабочий в СИЗ за опасной зоной

Примеры работы алгоритма проверки принадлежности опасной зоне



Рис. 4: Рабочий без СИЗ в опасной зоне

Принятие решения

Пусть

$[P_i]_{i=0}^{M-1}$ – результат работы модуля распознавания СИЗов,

$[D_i]_{i=0}^{M-1}$ – результат работы модуля детектирования

нахождения в опасной зоне такие, что:

$P_i = 1 \iff$ i-й человек находится в кадре без СИЗ

$D_i = 1 \iff$ i-й человек находится в кадре в опасной зоне

В данном случае M – число обнаруженных людей в кадре.
Тогда будем считать ситуацию на текущем кадре опасной,
если истинно следующее значение:

$$F = \bigvee_{i=0}^{M-1} (P_i \wedge D_i)$$

Матрица возможных ответов

Для расширения спектра возможных ответов требуется введение индикатора присутствия человека без СИЗ в кадре:

$$P = \bigvee_{i=0}^{M-1} P_i$$

D_i	P_i	0	1
0	—	$P \wedge \neg F$	
1	—		F

Результаты

- Спроектирована архитектура системы компьютерного зрения для детектирования опасных ситуаций на производстве;
- Разработан метод построения скелетных представлений людей в кадре с использованием модели AlphaPose;
- Реализован алгоритм детектирования принадлежности скелетного представления опасной зоне с использованием модели MiDaS;
- Описана принцип принятия решений по заданному кадру.

Возможные направления дальнейшего улучшения

- Разработка методов определения иных опасных ситуаций (вроде падения вблизи опасной зоны);
- Исследование возможности применения других методов оценки расстояния до объектов в кадре (например, на основе трёхмерной реконструкции помещения);
- Реализация сглаживания скелетных представлений на основе межкадрового трекинга для увеличения стабильности работы системы.

Основная литература

- ① Fang, H. AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time / H. Fang, J. Li, H. Tang // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2023. — Vol. 45, № 6. — P. 7157–7173.
- ② Ranftl, R. Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-shot Cross-dataset Transfer / R. Ranftl, K. Lasinger, D. Hafner // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2022. — Vol. 44, № 3. — P. 1623–1637.
- ③ Hartley, R. Multiple View Geometry in Computer Vision / R. Hartley, A. Zisserman. — Cambridge University Press, 2004. — 670 p.