

Министерство образования и науки РФ
Федеральное государственное автономное образовательное
учреждения высшего образования
«Национальный исследовательский технологический университет «МИСИС»
Кафедра инженерной кибернетики

**Отчет по лабораторной №5 на тему:
«Детекция»**

по дисциплине:

«Компьютерное зрение»

Направление подготовки: 01.03.04 Прикладная математика

Семестр 5

Выполнил:
Фейгенберг М.И.

БПМ-23-ПО-3

30.11.2025

Проверил:
Полевой Д.В.

(оценка)

(дата проверки)

Подпись: _____

Подпись: _____

Москва 2025

Цель работы

Разработка и реализация методов детекции объектов на изображении с использованием многомасштабного анализа, оценки локализации объектов и анализа их качества.

Задание

- Локализовать объекты на изображении, учитывая разные масштабы.
- Оценить достоверность локализации объектов с использованием метрики **score**.
- Визуализировать результаты детекции и сравнить их с эталонными результатами.

Инструкция по сборке и запуску

Весь код используемый и описываемый в данной лабораторной работы можно найти по ссылке:

(<https://github.com/Asbecov/misis2025f-23-03-feigenberg-m-i/tree/master>)

Сборка проекта выполняется с использованием CMake:

```
cmake .  
cmake --build ./build
```

Запуск программы:

- Для детекции объектов изображения
`./task05 --detect <input_lst_path> <output_dir>`
- Для оценки качества детекции на сравнении с эталоном
`./task05 --assess <input_img_path> <input_gt_detections_path>`
`<output_dir>`

Теоретическая часть

Детекция объектов

Детекция подразумевает задачи нахождения и локализации объектов на изображении. Целью является не только нахождение объекта, но и правильное его позиционирование в пространстве изображения, определение границ и выделение объектов в контексте окружающих элементов.

Реализация

Функция выполняющая детекцию объектов:

```
std::vector<Detection> detect(const cv::Mat& img, const std::vector<double>& scales = {1.0, 0.75, 0.5});
```

Для выделения объектов интереса воспользуемся уже готовой функцией сегментации на основе алгоритма водораздела, реализованной в предыдущей лабораторной работе.

Как и было заметно в предыдущей лабораторной работе, водораздел с функционалом *distance transform* является эффективным решением для сегментации плотно упакованных объектов, таких как монеты.

Выбор масштаба и мульти尺度ный анализ

Зависимость сегментации от разрешения

Качество сегментации водоразделом зависит от масштаба: при уменьшении изображения (масштаб < 1.0) - монеты сближаются, их границы сливаются, следовательно *distance transform* формирует один пик, отсюда водораздел выдаёт один сегмент на группу.

При увеличении изображения (масштаб > 1.0): текстура фона и шум усиливаются, что может порождать ложные компоненты или фрагментацию монет.

Для преодоления этой зависимости реализован ансамблевый анализ:

Обработка выполняется на семи масштабах: $\{2.0, 1.75, 1.5, 1.25, 1.0, 0.75, 0.5\}$. На каждом масштабе применяется функция сегментирования, для каждой метки строится бинарная маска, маска и *bbox* масштабируются обратно к исходному разрешению.

Все детекции каждого из масштабов объединяются в единый пул с оценкой, формируемой из среднего оценок всех детекций на масштабе.

Оценка достоверности локализации

Каждая сегментированная компонента оценивается по следующему score:

$$\text{score} = \text{std::clamp}(0.6 \cdot \text{boundary_contrast} + 0.4 \cdot \text{compactness}, 0, 1)$$

boundary-contrast - среднее значение градиента Собеля вдоль границы объекта, нормированное на глобальный максимум градиента изображения. Монеты на светлом фоне обладают резким перепадом яркости, что даёт высокий контраст. Нормировка обеспечивает устойчивость к изменению освещения.

compactness - отношение площади маски к площади ограничивающего прямоугольника:

$$\text{compactness} = \min(1.0, \text{area}(\text{mask}) / \text{area}(\text{bbox}))$$

Для идеального круга ~ 0.785 ; для монет — близко к 1. Низкое значение указывает на аномальную форму или шум.

Детекции с $\text{score} < 0.2$ отбрасываются как мало достоверные.

Визуализация детекций

```
cv::Mat visualize_detection(const cv::Mat& img, const std::vector<Detection>& detections, const double alpha = 0.4);
```

Для каждой детекции накладывает полупрозрачную цветную маску ($\text{alpha} = 0.4$) только внутри бинарной маски, рисует bounding box толщиной 2 пикселя и пишет score детекции в центре bbox.

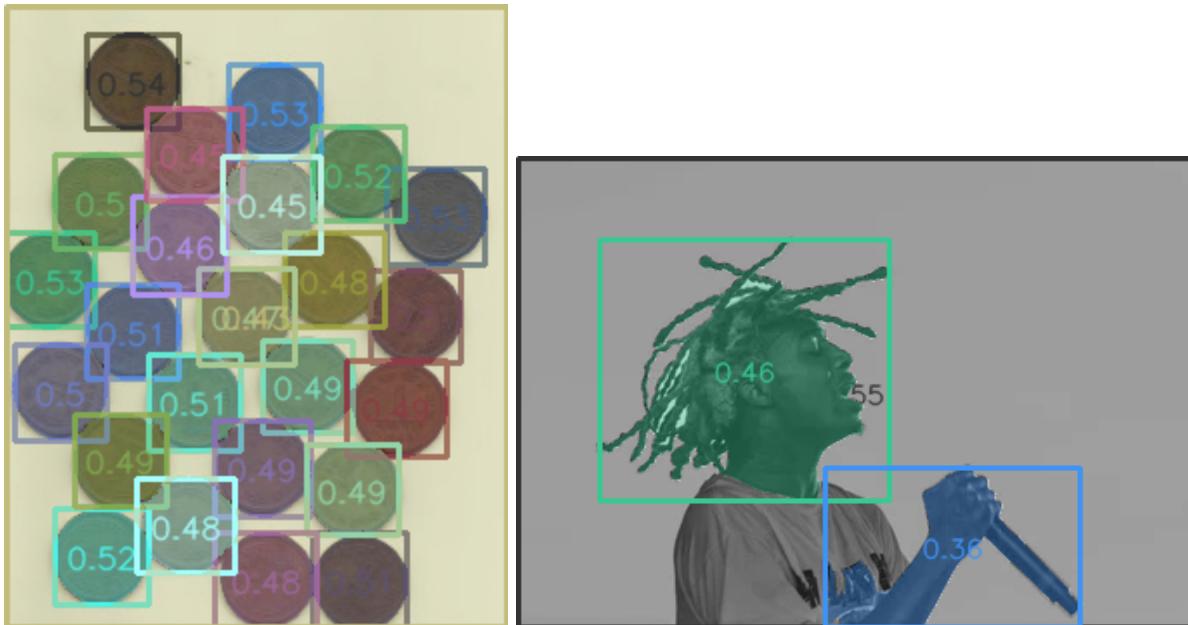


Рис 1-2, визуализация детекций с ее метрикой

Цвета генерируются по индексу, обеспечивая визуальную различимость объектов.

Сравнение с эталоном и визуализация ошибок:

```
cv::Mat visualize_detection_with_gt(const cv::Mat &img, const
std::vector<Detection> &detctions, const std::vector<cv::Rect>& gt_bboxes,
const double alpha = 0.4);
```

Для анализа качества эталонные bbox'ы загружаются из текстового файла в формате x, y, width, height.

На изображение накладываются:

- Зелёные прямоугольники эталона
- Синие детекции системы
- Красные ложные срабатывания

Совпадение определяется по $\text{IoU} \geq 0.5$.

Считается средний IoU полученных детекцией bbox'ов на сравнении с эталонными.

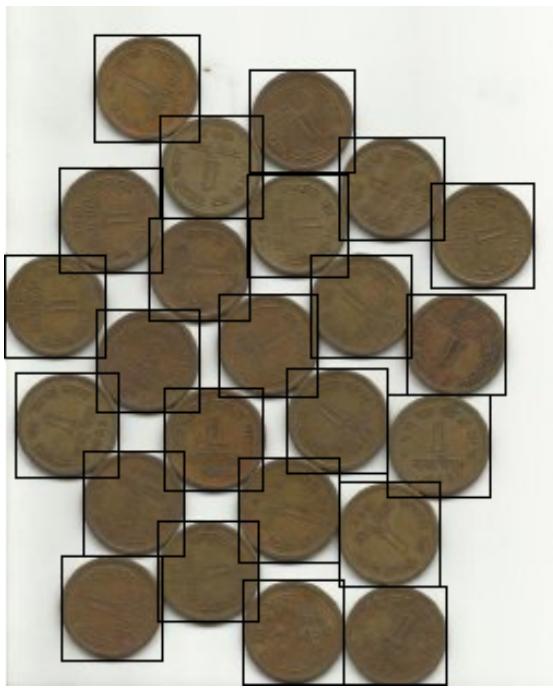


Рис 3, эталонные детекции объектов на изображении

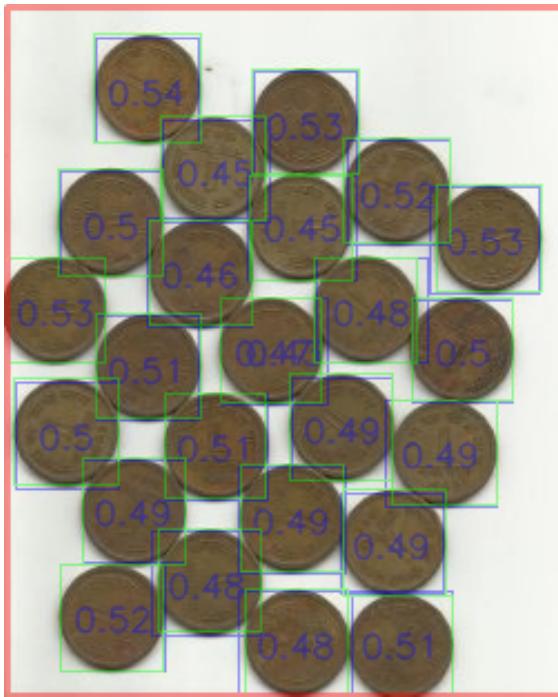


Рис 4, визуализация полученной детекции на сравнении с эталонной

При сравнении с эталонной детекцией описанный алгоритм получил хорошие результаты $\text{IoU} = 0.926$.

Заключение

В ходе лабораторной работы была реализована система детекции объектов на изображениях с использованием многомасштабного анализа. Для оценки качества локализации использовалась метрика **score**, которая была основана на контрасте границы и компактности объекта. Этот подход позволил точно локализовать объекты и исключить ложные детекции.