

Домашнее задание (HSE–CV): «Границы и локальные дескрипторы — Sobel, Canny, HOG»

ФИО студента, группа

17 сентября 2025 г.

Общее описание

Формат сдачи: один Jupyter Notebook для Google Colab (`.ipynb`), содержащий все вычисления и **детальные пояснения в Markdown**.

Срок: 1 неделя. **Рекомендуемое время:** 6–10 часов.

Цель: вручную реализовать ключевые этапы обработки изображений из лекции (свёртка, Sobel, Canny, HOG), провести эксперименты с параметрами и **объяснить** наблюдаемые эффекты.

Требование к воспроизводимости: при генерации случайных величин фиксируйте `seed`. Везде, где возможно, указывайте версии библиотек.

Оглавление ноутбука (рекомендуемый шаблон)

1. **Title & Authors** (ФИО, группа, дата).
2. **Setup** (установка пакетов, импорты, служебные функции).
3. **Данные** (описание и загрузка изображений).
4. **RGB → Gray: BT.601 vs среднее**.
5. **2D-Convolution & Sobel** (I_x, I_y, M, Θ , паддинг, сравнение с OpenCV).
6. **Canny (ручная реализация):** blur → градиенты → NMS → double-threshold → hysteresis.
7. **Сравнение с cv2.Canny** + подбор порогов.
8. **HOG (ручная реализация):** ячейки, бининги, L2-Nys, визуализация.
9. **Эксперименты с гиперпараметрами HOG**.
10. (Опционально) **Мини-кейс: HOG + линейный SVM**.
11. **Ответы на вопросы (Markdown)**.
12. **Выводы**.

Набор задач и критерии

0) Подготовка окружения и данных (обязательный раздел, без баллов)

В разделе *Setup* установите при необходимости: `numpy`, `matplotlib`, `opencv-python-headless`, `scikit-image`, `scikit-learn`. Импортируйте: `numpy as np`, `matplotlib.pyplot as plt`, `cv2`, `skimage.*`.

Данные: используйте минимум 2 встроенных изображения (`skimage.data`) и 2–3 собственных (разные сцены, контраст, шум).

1) RGB → Grayscale: веса против простого среднего (10 баллов)

Задача:

- 1) Реализуйте две функции конвертации RGB в серый:

$$I_{\text{gray}} = 0.299 R + 0.587 G + 0.114 B \quad (\text{BT.601}),$$

и простое среднее: $I_{\text{avg}} = (R + G + B)/3$.

- 2) Сравните на 2–3 изображениях: визуально и количественно (например, гистограммы яркости, среднее/дисперсия).
- 3) Объясните, почему взвешенная формула лучше согласуется с человеческим восприятием и как это влияет на дальнейшие шаги (градиенты).

Что показать: примеры до/после, гистограммы, 1–2 абзаца пояснений.

2) Свёртка и фильтры Собела: I_x, I_y , величина и направление (25 баллов)

Задача:

- 1) Реализуйте **2D-свёртку** с поддержкой паддинга `zero` и `reflect` (не использовать `cv2.filter2D`/готовые конволюции).
- 2) Реализуйте фильтры Sobel 3×3 :

$$G_x = \begin{pmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{pmatrix}, \quad G_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{pmatrix}.$$

Получите карты I_x, I_y для 3–4 изображений в обоих режимах паддинга.

- 3) Постройте величину и направление градиента:

$$M(x, y) = \sqrt{I_x^2 + I_y^2}, \quad \Theta(x, y) = \text{atan2}(I_y, I_x).$$

- 4) Сравните ваш M с `cv2.Sobel` (MSE/PSNR). Объясните расхождения (нормировки, паддинг, типы данных).
- 5) Исследуйте влияние паддинга: покажите кропы у границ при `zero` vs `reflect` и объясните артефакты.

Что показать: визуализации I_x, I_y, M , карту/палитру Θ , табличку MSE/PSNR, краткое резюме про шаблон 1–2–1 и про паддинг.

3) Canny руками: NMS и гистерезис (35 баллов)

Цель: вручную реализовать ядро детектора Кэнни (Gaussian blur допускается из cv2 или skimage).

Задача:

- 1) **Сглаживание:** примените Gaussian-blur с $\sigma \in \{0.8, 1.0, 1.4\}$; поясните компромисс шум vs размытие границ.
- 2) **Градиенты:** используйте свои $I_x, I_y \Rightarrow M, \Theta$.
- 3) **NMS:** реализуйте подавление немаксимумов вдоль направления Θ (квантуйте Θ до $0^\circ, 45^\circ, 90^\circ, 135^\circ$).
- 4) **Double Threshold + Hysteresis:** реализуйте сильные/слабые/фоновое, отслеживание связности (8-соседство).
- 5) **Сравнение:** сопоставьте с cv2.Canny на наборе изображений. Подберите $(T_{\text{low}}, T_{\text{high}})$ (например, $T_{\text{high}} \approx 2-3 \times T_{\text{low}}$), опишите методику (гистограмма M , Otsu, эмпирика).
- 6) **Отчёт:** для каждой σ покажите лучший пример с подписью, почему выбранные пороги оптимальны; отметьте типичные артефакты и их диагностику.

Что показать: поэтапные визуализации (после blur, M , NMS, итог), сравнение с cv2.Canny, пояснения про пороги и влияние σ .

4) HOG-дескриптор: ячейки, бининги, нормировка L2-Nys (25 баллов)

Задача:

- 1) Реализуйте HOG для окна:
 - градиенты I_x, I_y (можно ваши Sobel), M и Θ ;
 - диапазон $0-180^\circ$, 9 бинов (шаг 20°);
 - билинейная интерполяция по двум бинам по углу;
 - ячейка: 8×8 пикс.; блок: 2×2 ячейки; шаг блока: 1 ячейка;
 - **L2-Nys** нормировка:

$$f \leftarrow \frac{v}{\sqrt{\|v\|_2^2 + \varepsilon^2}}, \quad f_i \leftarrow \min(f_i, 0.2), \quad f \leftarrow \frac{f}{\sqrt{\|f\|_2^2 + \varepsilon^2}}.$$

- 2) Визуализируйте HOG (стрелочки/«ёжик») поверх изображения.
- 3) **Мини-исследование гиперпараметров:** ячейка 8×8 vs 16×16 ; 9 vs 12 бинов; L2 vs L2-Nys. Сравните визуально и численно (размерность, время, косинусное сходство).
- 4) (*Опционально*, +5б) Игрушечный датасет: HOG + линейный SVM vs порог по границам (из Canny) как простой базовый признак.

Что показать: визуализации, сравнительные коллажи/таблицы, 2–3 абзаца анализа влияния гиперпараметров на детальность/устойчивость/скорость.

5) Теоретические вопросы (Markdown-раздел, 5 баллов)

Ответьте кратко (2–6 предложений на пункт):

- 1) Разница между «операцией свёртки» и «обучаемым свёрточным слоем».
- 2) Почему atan2 предпочтительнее $\arctan(I_y/I_x)$ для углов градиента?
- 3) Почему шаблон $1-2-1$ в Sobel снижает влияние шума?

- 4) Чем плох глобальный единый порог по M , и как гистерезис это исправляет?
- 5) Ограничения НОГ и способы частичной компенсации масштабной чувствительности.

6) Бонус-задания (+ до 15 баллов)

Выберите не более двух:

- **Multi-scale Canny:** пирамида изображений, объединение границ (логическое ИЛИ) с нескольких масштабов; сравнение с одношкальным результатом.
- **Автоподбор порогов Canny:** Otsu/percentile-based для T_{high} и правило для T_{low} (например, $T_{\text{low}} = 0.4 T_{\text{high}}$); оцените устойчивость на наборе изображений.
- **Signed-HOG (0–360°) vs unsigned (0–180°):** когда принципиальна разница? Покажите на собственном примере.

Оценивание

Компонент	Баллы
RGB \rightarrow Gray (BT.601 vs среднее)	10
2D-свёртка & Sobel (I_x, I_y, M, Θ , сравнение, паддинг)	25
Canny (ручная реализация: NMS, double-threshold, hysteresis)	35
НОГ (реализация, визуализация, гиперпараметры)	25
Теоретические ответы (Markdown)	5
Итого (без бонуса)	100
Бонус (раздел 6)	+15

Требования к визуализации и отчётности

Каждое изображение/график снабжайте подписью (что показано, параметры, ключевой вывод). Для сравнений рекомендуется *side-by-side*/коллажи с подписями параметров под панелями. По возможности фиксируйте *время выполнения* (`%timeit`) для ключевых шагов. Функции снабжайте краткими *docstring* и проверками граничных условий (формы массивов, типы/диапазоны значений).

Частые ошибки и диагностика

- **Типы и шкалы:** используйте `float32/64`, нормируйте интенсивности в $[0,1]$ при необходимости.
- **Паддинг:** проверяйте карты у границ; сравнивайте `zero` и `reflect` на кроках.
- **NMS:** некорректная квантизация угла даёт «рваные» или пропавшие границы; убедитесь в корректности 4 направлений и выборе соседей.
- **Гистерезис:** учитывайте 8-соседство; слабые пиксели повышаются до границ *только* при связности со *strong*.
- **НОГ:** не забывайте интерполяцию по бинам и *перекрывание* блоков; проверяйте размерность вектора признаков.

Что сдавать

Один файл `.ipynb` (совместимый с Colab), содержащий все графики и выводы. Если используете внешние данные, приложите инструкцию по загрузке (например, из Google Drive). В начале ноутбука добавьте **сводку результатов** (5–10 предложений).

Мини-чеклист перед сдачей

- ☐ Две версии Grayscale и их сравнение.
- ☐ Собственная 2D-свёртка с `zero` и `reflect`.
- ☐ Собственные I_x, I_y, M, Θ и сравнение с OpenCV Sobel.
- ☐ Реализованы NMS и гистерезис; показан пайплайн Canny и сравнение с `cv2.Canny`.
- ☐ Реализован HOG (ячейки, бины, блоки, L2-Hys, визуализация) и исследованы гиперпараметры.
- ☐ Есть ответы на 5 теоретических вопросов.
- ☐ Все графики подписаны, параметры указаны, выводы сформулированы.

Формулы из задания (для удобства)

RGB \rightarrow Gray (BT.601): $I_{\text{gray}} = 0.299 R + 0.587 G + 0.114 B$,

Sobel: $G_x = \begin{pmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{pmatrix}, \quad G_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{pmatrix},$

Градиент: $M(x, y) = \sqrt{I_x^2 + I_y^2}, \quad \Theta(x, y) = \text{atan2}(I_y, I_x),$

L2-Hys: $f \leftarrow \frac{v}{\sqrt{\|v\|_2^2 + \varepsilon^2}}, \quad f_i \leftarrow \min(f_i, 0.2), \quad f \leftarrow \frac{f}{\sqrt{\|f\|_2^2 + \varepsilon^2}}.$

Удачи! Важно не только получить корректные результаты, но и **продемонстрировать ход мысли**: что вы пробовали, почему так решили и чем ваши выводы подтверждаются на графиках и числах.