

«Большие данные». Лабораторный практикум¹

ВАЖНО. Результаты работы требуется документировать, представляя в формате PDF (лаконично, в свободной форме): группа, ФИО, источник данных, номер задания, фрагменты программного кода с пояснениями и комментариями, диаграммы, результаты и т.д. *Рекомендуется* использовать систему L^AT_EX: [Overleaf](#) или [TeX Live](#). Прием работ: elenakhaa@yandex.ru.

Задание 1 [до 10 баллов, до 19 октября]². Сегментировать по яркости пиксели изображения, представленного файлом `plane.png` (или `plane_294x182.html`).

1. Преобразовать цвет каждого пикселя по правилу:

$$R'_i = G'_i = B'_i = Y_i,$$

где яркость пикселя

$$Y_i = 0,299R_i + 0,587G_i + 0,114B_i.$$

2. По значениям яркости пикселей *программно* построить гистограмму и по ее характеристикам выделить области качественной однородности пикселей (их границами будут локальные минимумы гистограммы).
3. Заменить цвета пикселей, соответствующих первой по частоте области однородности [корпус самолета], на красный. Результат (изображение) представить в виде PNG-файла (или HTML-таблицы).
4. В виде CSV-файла сформировать размеченный набор данных, описывающий классы (категории) пикселей изображения. За признаки (поля, атрибуты) следует принять R , G , B и $label$, где $label$ – порядковый номер области сгущения значений яркости пикселей.

¹Экзаменационная оценка определяется суммой баллов, набранных за семестр, включая экзаменационное задание (до 20 баллов): «удовлетворительно» – от 55 до 70 баллов; «хорошо» – от 71 до 85 баллов; «отлично» – от 86 до 100 баллов.

²Максимальная оценка за работу с признаками недобросовестного заимствования кода – 0 баллов (соглашение распространяется и на последующие задания).