

# Автоматическое детектирование и распознавание объектов на изображениях\*

*Дудоров Н. А., Томинин Я. Д., Томинин В. Д., Ерлыгин Л. О.,  
Демидова Ю. О.*

МФТИ(ГУ)

Работа посвящена обнаружению объектов на изображениях, другими словами, разделению изображений на два класса: содержащих искомый объект и не содержащих. Признаковое описание изображения строится с помощью дескрипторов. Наиболее известные из них: HOG, Haar, LBP, SIFT. Целью работы является усовершенствование дескриптора HOG для получения лучшего результата на ROC и DET кривых при заданном методе классификации. Решение задачи востребовано в распознавании людей на изображениях.

## Введение

Цель исследования заключается в получении метода построения признакового описания изображения для получения максимальной точности в задаче классификации. Исследование состоит в модифицировании дескриптора HOG для получения лучшей DET кривой, чем для стандартного HOG в задаче распознавания людей на изображениях. Проблемой существующего дескриптора HOG является то, что вектор, описывающий изображение, строится на основе порядка сканирования изображения, и поэтому не учитывает пространственную структуру изображения. В данной работе решается задача построения трёхмерного тензора на основе градиента гистограмм дескриптора HOG. Предлагается свернуть вектор, получаемый стандартным HOG, в тензор размера  $9 \times \text{число блоков на строку} \times \text{число строк}$ . Анализ результатов нового дескриптора будет производиться на DET кривой с использованием баз данных изображений «CITYSCAPES» и «INRIA Person».

**Постановка задачи** Пусть  $X^l = \{x\}_0^l$  – объекты обучающей выборки (изображения),  $Y = \{y\}_0^l : y \in \{-1, 1\}$  – ответы на обучающей выборке (два класса изображений: содержащие людей, и не содержащие).  $N : x \rightarrow y$  – модель (свёрточная нейронная сеть).

$$MR = \frac{FN}{TP + FN}$$

Задача состоит в получении наилучшей DET кривой:

$$\frac{MR}{FP} \rightarrow \min$$

(где  $FP$  – ложно-положительные ответы модели,  $FN$  – ложно-негативные,  $TP$  – правильные положительные ответы модели) путём замены:

$$x \in G^{h \times 4 \times bpc \times bpr} \rightarrow x' \in G_1^{h \times 2 \times bpr} \times G_2^{2 \times bpc}$$

где  $G, G_1, G_2 = \{0, 1, \dots, 64\}$ ,  $bpr = 7$ ,  $bpc = 15$ ,  $h = 9$