Обзор статьи <https://arxiv.org/ftp/arxiv/papers/2005/2005.01770.pdf>

Статья предлагает вариант решения проблемы пробок на дорогах.

Одной из причин заторов на дорогах авторы называют то, что традиционные системы управления траффиком, т.е. светофоры, не адаптируются к плотности движения. Как правило, проблему пытаются решить путем изменения времени смены фаз у светофоров в зависимости от времени суток, но это решение часто неэффективно.

Соответственно, авторы предлагают ввести интеллектуальную систему контроля трафика, которая будет оценивать в режиме реального времени плотность траффика на перекрестке или полосе движения с помощью методов обработки изображений и машинного обучения. Для решения задачи используются: Гистограмма направленных градиентов (Histogram of Oriented Gradients - HOG), Локальные бинарные шаблоны (Local Binary Patterns - LBP) и Метод опорных векторов (Support Vector Machine - SVM)

В предлагаемом методе для обучения модели SVM используются данные с камер видеонаблюдения одной полосы движения или с перекрестка. В качестве обучающих изображений с видео сохраняется 200 кадров.

Рис 1. Реальное изображение с камеры видеонаблюдения



Из каждого такого изображения извлекается «область интереса» (ROI). ROI выбирается таким образом, чтобы на изображении присутствовал только трафик, а все ненужные части отбрасывались. Отбрасывание ненужных частей помогает сократить расчеты и повышает точность оценки.

Рис. 2 Пример извлечения ROI из кадра. Ненужные части кадра отсечены - тротуар, люди, здания.



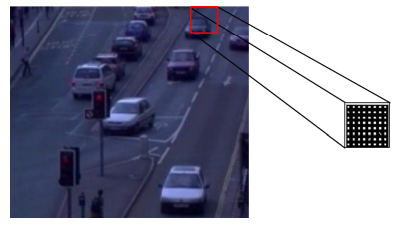
Затем полученное усеченное изображение разбивается на ячейки - для выбранного изображения это 72 ячейки, т.е. 8 х 9. Каждая ячейка размером 44х44 пикселей. Количество ячеек и размер сетки зависят от размера усеченного изображения ROI, который, может быть разным для разных камер, полос движения или перекрестков.

Рис. 3 Пример разбиения изображения по сетке



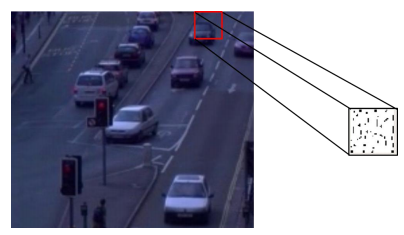
Для каждой ячейки в сетке вычисляется дескриптор HOG. Ячейка делится на более мелкие ячейки 3X3 пикселя и для каждой такой ячейки выполняется расчет гистограммы направлений градиентов. Комбинация этих гистограмм и является дескриптором.

Рис. 4 Пример вычисления дескриптора HOG для отдельной ячейки



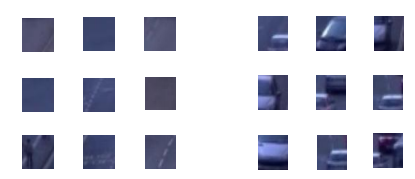
Далее, для каждой ячейки выполняется вычисление локального бинарного шаблона (LBP). Для вычисления LBP исходное изображение преобразуется в черно-белое. Для вычисления LBP используются два параметра – радиус и количество точек вокруг внешнего радиуса. Затем на основе полученного локального бинарного шаблона создается и нормализуется гистограмма.

Рис. 5 Пример вычисления локального бинарного шаблона для отдельной ячейки



Полученные данные передаются в модель SVM для обучения. Мы предполагаем, что ячейка содержит трафик если он имеет 30% или более доли транспортного средства.

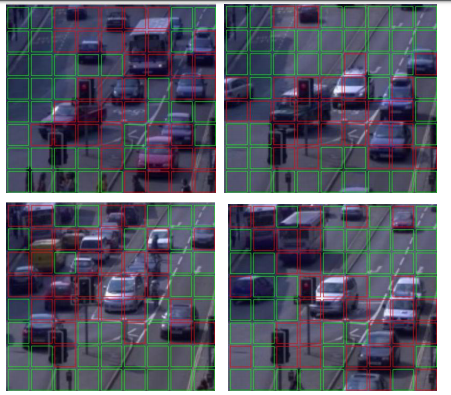
Рис. 6 Слева ячейки без траффика, справа – с траффиком.



Для оценки реальной плотности трафика используется видео с камер наблюдения в реальном времени. Кадр обрезается, отбрасывается лишнее, изображение разбивается на части. Затем обученная модель для каждой ячейки принимает решение – что содержит ячейка, пустую дорогу или машины. После этого считается количество ячеек, имеющих трафик.

На рисунке изображение с камеры разделено на прямоугольники представляющих ячейки. Каждая ячейка окрашена в соответствии с результатом классификации ее моделью SVM. Ячейка красная, если на ней движение, зеленая - на ней пустая дорога.

Рис. 7 Пример обработки ячеек в реальном изображении



Для реального обучения было использовано 200 изображений с камер видеонаблюдения. Изображения разделили на 20000 ячеек. Ячейки были разделены на два класса - трафик, и не трафик. Данные для обучения разделили на тестовый и тренировочный в пропорции 80 и 20.

Accuracy на тестовом наборе - 0,9488248

Precision на тестовом наборе - 0.9400322

Recall на тестовом наборе - 0.9402985

F1 на тестовом наборе - 0.9398987

Рис. 8 Архитектура решения

