

УДК 004.93

ПРИМЕНЕНИЕ КЛАССИЧЕСКИХ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

Аликрицкий М.С. студент гр. ИСТм-42, 2 курс
Поволжский государственный университет телекоммуникаций и информати-
ки
г. Самара

Рассматриваются классические методы машинного обучения, применяемые для анализа и обработки изображений, в частности методы опорных векторов (SVM) и k-ближайших соседей (k-NN). Актуальность их исследования в эпоху глубокого обучения обусловлена необходимостью создания интерпретируемых, ресурс эффективных и быстрых решений для прикладных задач, где объем размеченных данных невелик, а вычислительные ресурсы ограничены. К таким задачам можно отнести, например, автоматизированный визуальный контроль качества деталей в реальном времени на производственной линии, классификацию стандартизированных документов или базовое распознавание лиц в системах контроля доступа с малой пользовательской базой [1, 2]. В условиях, когда сбор и разметка больших данных затруднены или экономически невыгодны, а требования к прозрачности принятия решения высоки, классические алгоритмы демонстрируют значительный практический потенциал [4].

Классические методы машинного обучения, такие как SVM и k-NN, обладают рядом преимуществ в условиях ограниченных вычислительных ресурсов: они не требуют мощных графических ускорителей, легко интерпретируются и позволяют работать с небольшими размеченными выборками. В отличие от глубоких нейросетей, для их обучения достаточно нескольких сотен или тысяч примеров на класс, что существенно снижает затраты на сбор и разметку данных [5].

Метод опорных векторов (SVM) основан на построении гиперплоскости, разделяющей классы изображений в многомерном пространстве признаков с максимальным зазором [1]. Математически задача обучения SVM сводится к решению задачи квадратичной оптимизации: необходимо минимизировать норму вектора весов w при условии корректной классификации всех обучающих примеров, что записывается как:

$$y_i(w * x_i + b) \geq 1$$

где - x_i - вектор признаков объекта, y_i - его метка класса в интервале $[-1, 1]$, w - вектор весов, которые определяют ориентацию гиперплоскости, а b - смещение.

Для случаев, когда данные линейно неразделимы, используется механизм мягкого зазора, позволяющий работать с нелинейными разделяющими поверхностями без явного преобразования признаков в пространство высокой размерности [1]. Такой подход хорошо работает при задаче бинарной классификации и позволяет достигать высокой точности на структурированных данных, полученных, например, с помощью дескрипторов LBP или HOG.

Основной особенностью SVM можно выделить то, что он, использует функции, с помощью которых строятся нелинейные поверхности, которые разделяют без явного перехода в большую размерность. Если же рассматривать практически задачи, то, можно заметить, что, на практике для распознавания изображения часто применяют базисные функции (RBF), они дают показывают отличные результаты при условии, что были правильно подобраны гиперпараметры C и γ [2].

Далее рассмотрим метод k -ближайших соседей (k -NN), его можно описать несколькими словами, а именно, простой и интуитивно понятный. Суть этого алгоритма, в том, что решение о принадлежности изображения к определенному классу изображений принимается на основании большинства среди k соседей по признаку заданных примеров [5]. Метод обучается очень медленно, так как не строит собственную явную модель во время обучения, в результате все вычисления проводятся до самого момента классификации нового найденного объекта.

Для классификации каждого нового объекта алгоритм подсчитывает расстояние от него, до каждого нового объекта в обучающей выборке, вследствие чего алгоритм определяет степень их близости. Класс же объекта определяется путем так называемого голосования, среди всех классов этих соседей, что выражается формулой:

$$\sum_{i=1}^n \frac{1}{d^2(X, Y_i)}$$

где $d^2(X, Y_i)$ - квадрат расстояния от известного объекта Y_i до нового X , n - количество известных записей класса, для которого рассчитываются голоса. В качестве меры близости чаще всего используется евклидово расстояние между векторами признаков, которое вычисляется по формуле:

$$d(x, x_i) = \sqrt{\sum_{j=1}^m (x_j - x_{ij})^2}$$

где m — размерность пространства признаков, а x_j и x_{ij} — значения j -го признака для классифицируемого объекта и i -го объекта обучающей выборки

соответственно. Выбор метрики расстояния может варьироваться в зависимости от специфики данных.

Кроме евклидова расстояния, в задачах распознавания изображений иногда используют манхэттенское расстояние или косинусную меру близости. Выбор метрики влияет на качество классификации и может быть оптимизирован с помощью кросс-валидации. Также важным параметром является само значение k , слишком большое может привести к сглаживанию границ между классами.

Алгоритм не требует явного этапа обучения, что упрощает его реализацию и позволяет быстро адаптироваться к изменениям в обучающей выборке, однако время классификации линейно зависит от ее размера, что может стать ограничением при работе с большими объемами данных. k -NN широко применяется для распознавания образов, в том числе изображений, особенно в сочетании с такими дескрипторами, как HOG или LBP, и часто служит базовым методом для сравнения с более сложными алгоритмами, благодаря своей простоте и наглядности [3].

Оба метода имеют свои сильные стороны. При работе с данными с большой размерностью подойдет метод SVM, который показывает хорошую обобщающую способность на не очень больших выборках, однако, для обучения модели метода нужно решать задачи по оптимизации, а также очень ответственно подходить к выбору ядра и гиперпараметров. Метод k -NN, напротив, легко адаптируется к различным новым появляющимся классам, но минус в том, что классификация очень сильно зависит от объема обучающей выборки. SVM хорошо подходит для задач, где имеются стабильные и структурированные данные, а k -NN хорошо себя показывает в задачах где важна интерпретируемость и адаптация [2].

В итоге можно отметить, что сочетание алгоритмов SVM и k -NN с дескрипторами LBP и HOG, дает возможность создавать устойчивые системы, для распознавания картинок, при низких вычислительных затратах. Практическая же ценность данных методов уже подтверждается на практике в прикладных областях, где создание глубоких нейронных сетей не имеет смысла или же затраты при их создании, будут слишком высоки. Классические методы машинного обучения не стоит списывать со счетов, они до сих остаются актуальными, в некоторых случаях даже заменяя сложные подходы.

Список литературы:

1. Созинов П. А., Шелухин О. И. Метод опорных векторов в задаче распознавания образов с использованием гистограмм ориентированных градиентов // Информационные технологии. - 2018. - № 4. - С. 243-249.
2. Вишняков В. С., Дмитриев В. И. Сравнительный анализ методов опорных векторов и k -ближайших соседей для классификации изображений // Вестник компьютерных и информационных технологий. - 2020. - № 6 (192). - С. 22-28. - DOI 10.14489/vkit.2020.06.pp.022-028.

3. Гонсалес Р., Вудс Р. Цифровая обработка изображений : пер. с англ. - М. : Техносфера, 2012. - 1104 с.
4. Журавлёв Ю. И., Рязанов В. В., Сенько О. В. Распознавание. Математические методы. Программная система. Практические применения. - М. : Фазис, 2006. - 176 с.
5. Мерков А. Б. Распознавание образов: Введение в методы статистического обучения. - М. : URSS, 2021. - 256 с.