

АЛГОРИТМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ НА ОСНОВЕ НЕЙРОННОЙ СЕТИ

Рассматриваются алгоритмы применения сверточных нейронных сетей для распознавания изображений

ВВЕДЕНИЕ

Данная работа посвящена распознаванию изображений коллекции CIFAR-10, с помощью искусственных нейронных сетей, с использованием гистограммы ориентированных градиентов и полутоновой матрицы смежности в качестве дополнительных источников признаков, а также тестированию и сравнению методов распознавания. Предложен подход к предобучению сверточных нейронных сетей. Указанная тема выбрана в связи с актуальностью задач распознавания изображений и задач компьютерного зрения в целом.

I. СВЕРТОЧНАЯ НЕЙРОННАЯ СЕТЬ

Сверточной нейронной сетью называют нейросеть со следующей архитектурой, удобно визуализируемой втрехмерном пространстве (см.рис.1.):

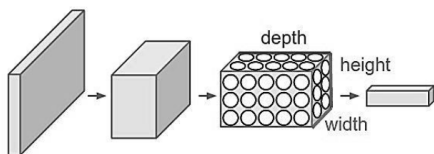


Рис. 1 – Архитектура Сверточной нейронной сети

В качестве входа каждый слой сети принимает трехмерный тензор, на практике чаще всего являющийся изображением.

Выходом сверточной сети является большой набор преобразованных данных, содержащих свойства входного изображения, но не конечный результат задачи аппроксимации. Таким образом, CNN является вспомогательным инструментом. Часто после CNN ставится MLP (многослойный перцептрон), который производит оценку извлеченных свойств.

II. СРАВНИТЕЛЬНЫЙ АНАЛИЗ МОДЕЛЕЙ

Сравнению были подвергнуты следующие модели:

1) MLP с тремя скрытыми слоями по 300, 200 и 50 нейронов соответственно.

Родионов Вадим Сергеевич, магистрант кафедры вычислительных методов и программирования БГУИР, v.rodionov997@gmail.com.

Научный руководитель: Матвеевко Владимир Владимирович, кандидат физ.-мат. наук, доцент, vladimir66@bsuir.by.

2) Сверточная нейронная сеть с тремя слоями свертки и одним полносвязным слоем из 200 нейронов. Параметры слоев свертки: 100 фильтров (3x3), 100 фильтров(2x2), 150 фильтров(2x2).

3) Модифицированная версия сети из п.2, использующая следующий алгоритм построения и обучения: обучается CNN из п.2. стандартным образом, к выходу 3го слоя свертки подключается полносвязный слой, принимающий также HOG и GLCM, полученная модель применяется к исходной тестовой выборке, формируя новую, производится предобучение MLP с тремя скрытыми слоями по 300, 100 и 100 нейронов на новой выборке. MLP подключается к выходу CNN.

4) Версия сети из п.3, предобучающая только первый скрытый слой.

Таблица 1 – Сравнение моделей

№	Ошибка на обучающей выборке	Ошибка на валидационной выборке	Точность на валидационной выборке	Точность на тестовой выборке
1	0.0412	0.0535	0.6635	0.6521
2	0.0313	0.0383	0.7315	0.7234
3	0.235	0.313	0.1034	0.0967
4	0.0023	0.0421	0.7438	0.7456

III. ВЫВОДЫ

Наибольшую эффективность показала сеть №4, наихудшую – сеть №3. Результаты, достигнутые моделями №4 и №2, отличаются незначительно, что может говорить о невысоком содержании дополнительных данных в HOG и GLCM. Предобучение для сети №3 оказалось неэффективным и только ухудшило результат. Возможно, это связано с тем, что данная задача требует более специфического подхода к предобучению.

1. Stanford CS class “CS231n: Convolutional Neural Networks for Visual Recognition” [Electronic resource] / Fei-Fei Li, A. Karpathy. – Mode of access: <http://cs231n.stanford.edu/>.