

Использование рецептивных полей в методах машинного обучения для обработки изображений, представленных выборками случайных отсчетов

В.Е. Анциперов

*Институт радиотехники и электроники им. В.А.
Котельникова РАН
Москва, Россия
anticperov@cplire.ru*

В.А. Кершнер

*Институт радиотехники и электроники им. В.А.
Котельникова РАН
Москва, Россия
vladkershner@mail.ru*

Аннотация—В работе предлагается и обосновывается способ обработки изображений, полученных путем формирования выборок случайных фотоотсчетов, с использованием рецептивных полей. Данный способ основан на известных фактах сенсорной нейробиологии, в частности в области функциональных особенностей нейронов зрительной системы человека. Предлагаемый метод позволит эффективно обрабатывать фотоотсчетные изображения средствами машинного обучения.

Ключевые слова— *Выборки случайных отсчетов, автоэнкодеры, рецептивные поля, сверточные нейронные сети, машинное обучение.*

1. ВВЕДЕНИЕ

Понятие рецептивного поля неразрывно связано с описанием функционирования соответствующих нейронов, в случае со зрительной системой такими нейронами являются ганглиозные клетки сетчатки глаза. Современные представления об этом классе зрительных нейронов достаточно расширились, что позволяет нам использовать уже имеющиеся модели зрительной системы на примере животных [1,2]. В свою очередь рецептивные поля связаны с глубокими сверточными сетями, которые широко используются в задачах по классификации изображений.

В ранних работах, посвященных новым способам представления данных для анализа изображений [3], было предложено использование выборок случайных фотоотсчетов для кодирования слабоинтенсивных изображений. На основе этих выборочных представлений с использованием принципов автокодирования был описан простой метод восстановления изображений. Стоит отметить, что автоэнкодеры являются относительно простой архитектурой нейронных сетей, направленной на обучение системы без учителя. Основная задача автокодировщика – преобразовать входные данные в подобные данные на выходе таким образом, чтобы максимально уменьшить объем входных данных и минимизировать любые искажения при соответствующих преобразованиях.

Однако для решения задач глубокого обучения, использование автоэнкодеров становится недостаточным, и требуется использование более эффективных методов машинного обучения, связанных с использованием более сложных алгоритмов, позволяющих не только преобразовывать данные из

одного представления в другое, но и выделять из них скрытую информацию - метаданные. В данной работе обсуждаются способы и методы обработки изображений для распознавания образов, обнаружения особенностей и т.д. с использованием рецептивных полей сенсорных нейронов.

2. СТРУКТУРА МОДЕЛИ РЕЦЕПТИВНЫХ ПОЛЕЙ ДЛЯ ОБРАБОТКИ ФОТООТСЧЕТНЫХ ИЗОБРАЖЕНИЙ

Поскольку подавляющее большинство известных методов обработки изображения были инспирированы зрительной системой животного, и в частности человека, стоит отметить некоторые сходства данных методов с этой системой. Основным объектом обработки информации в любой сенсорной системе человека является световое излучение, отраженное от поверхностей объектов, окружающих и находящихся в поле зрения наблюдателя. Регистрируемая входная информация, осуществляемая через зрительные нейроны сетчатки глаза (колбочки и палочки), проходит через биполярные клетки, передается в ганглиозные клетки и далее, уже в виде нервных импульсов, через зрительный нерв поступает в зрительную кору головного мозга (рис.1). Для наглядности эту систему можно упростить, представив в виде некоторой простой нейронной сети типа автоэнкодера с множеством слоёв. В таком случае входной слой будет состоять порядка из ста миллионов фоторецепторов, выходным слоем является зрительная кора, а набор из скрытых слоёв представлен наборами горизонтальных, биполярных, ганглиозных и других клеток. Стоит отметить, что ганглиозные клетки сетчатки глаза образуют мозаичную структуру из дендритов, покрывающую визуальное пространство сетчатки, и имеют соответствующие рецептивные поля.

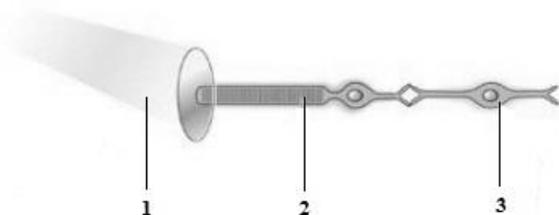


Рис. 1. Рецептивное поле зрительного нейрона: 1 – падающее световое излучение, 2 – фоторецепторная клетка, 3 – биполярная клетка

Понимания структуры сетчатки глаза и знания о нейронных сетях позволили нам разработать модель обработки изображения. Как и в случае со зрительной системой человека, модель обработки можно представить в виде аналогичной нейронной сети. В качестве предварительного этапа (первого слоя) используются выборки случайных фотоотсчетов. Использование выборочных представлений позволяет уменьшить объем исходного изображения, а значит более эффективно обрабатывать данные уже на начальном этапе. Следующим этапом является формирование параметрической модели, особенностью которой является упорядочивание компонент в области изображения. Такая модель имеет решетчатую структуру, что непосредственно связывает её с вышеуказанными фактами визуального восприятия человека.

На рис. 2 представлен пример обработки изображения. В качестве исходного изображения было взято цветное изображение 11860 (Country Pleasure ASB 3) из базы данных Emotive2000 [7]. Изображение было задано изначально в формате JPG с размерами 900 x 884 пикселей и имело глубину цвета 24 бита. Для уменьшения объема вычислений было произведено преобразование в формат GIF с размером 512 x 512 пикселей и в серую палитру с глубиной цвета 8 бит. Было сформировано выборочное представление с размером 500.000 отсчетов, осуществленное методом семплирования с отклонением. Количество использованных отсчетов было выбрано, исходя из двух параметров - конечного объема данных файла выборки, который используется в последующей обработке данных и отвечает за её скорость, и размера соответствующего изображения выборки, определяющего возможность интерпретировать обрабатываемые данные наблюдателю. На последнем этапе компоненты, представляющие собой отдельные отсчеты, были упорядочены в узлах решеточной структуры, размер самих узлов и их шаг были определены таким образом, чтобы конечное изображение оставалось четким, но сама решетчатая структура визуально не воспринималась.

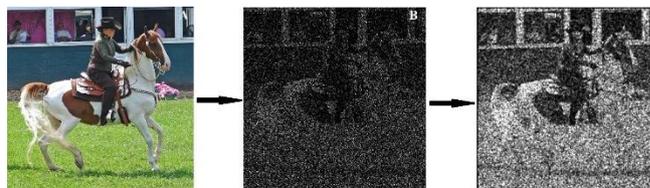


Рис. 2. Формирование параметрической модели: А – исходное изображение, В – выборка случайных фотоотсчетов, С – обработанное изображение

На основе этих данных была сформирована параметрическая модель исходного изображения, которая, как мы считаем, позволит использовать такие инструменты машинного обучения как сверточные нейронные сети для распознавания объектов, выделения особенностей и событий на изображениях, а также проведения оценок и анализа изображений.

БЛАГОДАРНОСТИ

Работа выполнена за счет бюджетного финансирования в рамках государственного задания в ИРЭ им. В.А. Котельникова РАН.

ЛИТЕРАТУРА

- [1] Antsiperov, V.E. New Generative Image Model for Variational Autoencoders Based on Counts Partition / V.E. Antsiperov // International Conference on Information Technology and Nanotechnology (ITNT). – 2021 – P. 1-8.
- [2] Hubel, D.H. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex / D.H. Hubel, T.N. Wiesel // The Journal of physiology. – 1962. – Vol. 160(1). – P. 106-154.
- [3] Hubel, D.H. Receptive fields of single neurones in the cat's striate cortex / D.H. Hubel, T.N. Wiesel // Brain Physiology and Psychology. – University of California Press, 2020. – P. 129-150.
- [4] Wienbar, S. The dynamic receptive fields of retinal ganglion cells / S. Wienbar, G.W. Schwartz // Progress in retinal and eye research. – 2018. – Vol. 67. – P. 102-117.
- [5] Fossum, E.R. Photon-Counting Image Sensors / E.R. Fossum, N. Teranishi. – MDPI, 2017.
- [6] Schwartz, G.W. The spatial structure of a nonlinear receptive field / G.W. Schwartz // Nature neuroscience. – 2012. – Vol. 15(11). – P. 1572-1580.
- [7] Robb, D.A. Crowdsourced Feedback With Imagery Rather Than Text: Would Designers Use It? / D.A. Robb, S. Padilla, B. Kalkreuter, M.J. Chantler // CHI'15: 33rd Annual ACM Conference on Human Factors in Computing Systems Proceedings, 2015.