

## ПРИМЕНЕНИЕ СВЕРТОЧНЫХ СЕТЕЙ ДЛЯ ЗАДАНИЯ УСЛОВНОГО РАСПРЕДЕЛЕНИЯ БОЛЬЦМАНОВСКИХ МАШИН МАКСИМАЛЬНОГО ЗАЗОРА

*Гитман Игорь Александрович,  
Кириллов Александр Николаевич*

*Студент, аспирант*

*Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия*

*E-mail: igor.a.gitman@gmail.com, akirillov@cs.msu.ru*

Задача сегментации изображений — это одна из основных задач компьютерного зрения. В данной работе мы предлагаем способ улучшения одного из методов для решения этой задачи: Больцмановских Машин Максимального Зазора (Max-Margin Boltzmann Machines, ММВМс [1]). ММВМс — это вид Больцмановских Машин, с помощью которых можно задавать распределение бинарных масок (сегментации) и скрытых переменных при условии исходной картинки:

$$E(y, h, x) = -y^T W h - h^T (V^1 x^1 + c) - y^T (V^0 x^0 + b) \quad (1)$$

Здесь  $y, h, x = \{x^0, x^1\}$  — это бинарная маска, скрытые переменные и признаки, выделенные из картинки. Остальные переменные — это настраиваемые параметры модели.

Авторы статьи разработали эффективную процедуру обучения с помощью минимизации функции максимального зазора (max-margin function). В качестве признаков  $x^0$  и  $x^1$  для исходных картинок используются гистограммы цветов, гистограммы ориентированных градиентов и прочие дескрипторы, хорошо зарекомендовавшие себя в разных задачах компьютерного зрения. Подобная процедура извлечения признаков весьма трудозатратна. Для каждой конкретной задачи требуется заново настраивать большое количество параметров и выбирать оптимальную комбинацию используемых признаков, чтобы получить наилучшее качество. В данной работе мы предлагаем метод, с помощью которого можно автоматизировать этот процесс и включить его в саму модель. Для этого мы заменили те части исходного функционала (1), в которых использовались «вручную» выделенные признаки на результаты работы сверточных сетей [2] — глубоких нейронных сетей специальной структуры, позволяющих выделять релевантные для конкретной задачи признаки из данных. Исследования последних лет показывают, что использование сверточных сетей дает значительное улучшение результатов во многих

задачах машинного обучения (например, для классификации изображений [3])

Проведенные нами эксперименты подтвердили, что подобную автоматизацию можно сделать с минимальными потерями качества.

В дальнейшем мы планируем обучить более сложную модель, объединяющую исходный алгоритм минимизации зазора и метод обратного распространения ошибки, использующийся для обучения сверточных сетей, что предположительно должно еще сильнее улучшить качество сегментации.

### Литература

1. Jimei Yang, Simon Safar, Ming-Hsuan Yang. Max-Margin Boltzmann Machines for Object Segmentation // Computer Vision and Pattern Recognition, 2014. CVPR 2014
2. Yann LeCun, Léon Bottou, Yoshua Bengio, Patrick Haffner. Gradient-Based Learning Applied to Document Recognition // Proceedings of the IEEE, 1998
3. Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks // Advances in neural information processing systems, 2012