

ПРОБЛЕМА КОЛЛАПСА НЕЙРОННЫХ СЕТЕЙ ПРИ ИСПОЛЬЗОВАНИИ TRIPLET LOSS

Рогачёв Н. Е.

Белорусский государственный университет, Минск, Беларусь,
e-mail: mikalai.rahachou@gmail.com

Обучение нейронных сетей для распознавания людей по лицу и одежде с использованием triplet loss доказало свою эффективность на многочисленных датасетах. Тем не менее, на практике обучение таких сетей очень часто сталкивается с проблемой коллапса нейросетевого отображения, т.е. схождения обучения к локальному оптимуму, в котором отображение $f(x)$ входа нейронной сети на её выход принимает вид

$$f(x) \equiv 0. \quad (1)$$

Настоящая работа посвящена проблеме коллапса нейронных сетей, обучаемых с использованием triplet loss, и способам ее решения. В рамках работы были опробованы подходы, связанные с подбором learning rate, размера мини-батча и обучающей выборки, а также весов инициализации, и установлено, что никакие разумные значения этих параметров не предотвращают коллапс сети. Был сделан вывод о том, что наиболее важным гиперпараметром в этом смысле является способ выбора триплетов. В частности, было установлено, что для триплетов (x_i^a, x_i^p, x_i^n) , где x_i^a – изображение определенного класса, x_i^p – наиболее удаленное от x_i^a изображение того же класса (сложный положительный пример), и функции потерь

$$L = \text{Softplus} \left(\left\| f(x_i^a) - f(x_i^p) \right\|_2 - \left\| f(x_i^a) - f(x_i^n) \right\|_2 \right), \quad (2)$$

где $\text{Softplus}(z) = \ln(1 + e^z)$, критическое значение имеет как раз выбор отрицательных примеров x_i^n – изображений класса, отличного от класса изображений x_i^a и x_i^p .

Выбор сложных отрицательных примеров, которые находятся ближе всех к x_i^a , в теории является наиболее информативным и полезным, но на практике приводит к излишней сложности задачи, недостижимости нетривиальных оптимумов и коллапсу сети. Введение же ограничения

$$\left\| f(x_i^a) - f(x_i^p) \right\|_2 < \left\| f(x_i^a) - f(x_i^n) \right\|_2 \quad (3)$$

на отрицательные примеры позволяет, как показано в работе, сделать сложность задачи, с одной стороны, достаточно высокой для успешного выявления статистических различий в изображениях, и, с другой стороны, достаточно низкой, чтобы избежать коллапса отображения $f(x)$.

Список литературы

1. Hermans, A., Beyer, L., Leibe, B. In Defense of the Triplet Loss for Person Re-Identification [Electronic resource] / A. Hermans, L. Beyer, B. Leibe. – arXiv, 2017. – Mode of access: <https://arxiv.org/pdf/1703.07737.pdf>. – Date of access: 02.04.2018.