

Отбор информативных элементов для обучения легкого сверточного нейросетевого классификатора в условиях сильного дисбаланса обучающей выборки*

Д.Е. ИВАНОВ^{1,II}, Д.В. ПОЛЕВОЙ^{1,III,IV}, Д.Л. ШОЛОМОВ^{1,V}

^I Национальный исследовательский технологический университет «МИСиС», г. Москва, Россия

^{II} ООО «Когнитивные технологии», г. Москва, Россия

^{III} Институт системного анализа Федерального исследовательского центра «Информатика и управление» Российской академии наук, г. Москва, Россия

^{IV} Московский физико-технический институт, г. Долгопрудный, Московская обл., Россия

^V Институт проблем передачи информации им. А.А. Харкевича Российской академии наук, г. Москва, Россия

Аннотация. В данной статье рассматривается задача балансировки обучающей выборки при обучении классификатора изображений на основе сверточной нейронной сети. Для обучения «легкого» классификатора в условиях сильного дисбаланса обучающей выборки предлагается активная схема отбора (прореживания) информативных элементов в процессе обучения. Экспериментальная проверка на примере задачи классификации изображений рукописных цифр и зон детекции дорожных знаков показывает стабильное преимущество предложенной схемы по сравнению со случайным отбором.

Ключевые слова: распознавание образов, сверточные нейронные сети, глубокое обучение, активное обучение, обучение по расписанию.

DOI: 10.14357/20790279180523

Введение

Несбалансированность данных – хорошо известная проблема в области статистического обучения [1]. Обширный обзор литературы и исследование этой тематики для классификаторов изображений на основе сверточных нейронных сетей могут быть найдены в [2]. Одним из вариантов учета дисбаланса является «взвешивание» на разных этапах процесса обучения. Альтернативой является балансировка выборок за счет изменения состава представителей. Дополнение выборки (oversampling) возможно при неравных, но сравнимых по количеству представителей классах за счет повторного включения копий элементов недопредставленных классов. Такое дополнение часто сочетается с аугментацией (augmentation) обучающей выборки. В случае сильного дисбаланса для наиболее многочисленных классов делается прореживание выборки (undersampling). Обычно дополнение и прореживание делаются рандомизированно в предположении «правильного» распределения примеров в исходных множествах.

В практических задачах часто «сырых» исходных данных существенно больше, чем требуется для обучения, а создание эталонной разметки дорогой процесс. Для разрешения противоречия между конкурирующими задачами минимизации стоимости разметки при максимизации качества обученного классификатора часто используют методы активного обучения, в которых после каждого раунда обучения модели очередное множество точек для разметки выбирается из еще не размеченного пула [3,4]. Также к активным схемам обучения классификаторов можно отнести обучение по расписанию [5,6] и различные варианты выбора элементов для обучения (sampling) [7,8].

Будем называть обучаемый классификатор «легким» если технические и временные ресурсы позволяют провести достаточное количество итераций его обучения (обычно, несколько сотен или тысяч). Далее рассмотрим активную схему отбора информативных элементов для обучения легкого сверточного нейросетевого классификатора в условиях сильного дисбаланса обучающей выборки за счет прореживания многочисленных классов.

* Работа выполнена при частичной финансовой поддержке РФФИ, грант № 17-29-03370.

1. Активный отбор информативных элементов для обучения

Для отбора информативных элементов в процессе обучения классификатора предлагается итеративно расширять сбалансированную обучающую выборку за счет формирования на каждой итерации дополнительного множества обучающих примеров с максимальным «разнообразием».

Первые N_0 элементов представителей каждого класса выбираются случайным образом из всего множества обучающих примеров. Так формируется выборка для итерации 0. Далее на каждой итерации повторяется следующий процесс.

1. Выборка, которая была получена на предыдущей итерации, используется для обучения новой сети.
2. Из оставшихся элементов множества обучающих примеров случайным образом выбирается некоторое подмножество, элементы которого подаются на вход полученной сети для получения характеристических векторов (например, значения активаций промежуточного слоя).
3. В полученном множестве характеристических векторов выбираются N_i элементов для каждого класса, и соответствующие им изображения добавляются в обучающую выборку.

Пространство характеристических векторов обозначим X . На X введем функцию схожести, между векторами $x, y \in X$, которую будем рассчитывать по формуле:

$$d(x, y) = \cos(x, y) = \frac{x \cdot y}{\|x\|_2 \|y\|_2} = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}}$$

Если Q_{ik} – множество элементов, отобранных на итерации i для класса k , то на итерации j множество Q_j всех отобранных элементов

$$Q_j = \bigcup_{i=0}^j \bigcup_k Q_{ik}$$

Рассмотрим две модификации базового алгоритма, которые отличаются способом отбора на шаге 3.

Подход 1.

Для каждого класса объектов k найдем δ_{jk} , при котором выполняется:

$$\begin{aligned} \exists Q_{jk} \subset X_k \setminus \bigcup_{i=0}^{j-1} Q_{ik} : |Q_{jk}| = \\ = N_j, \forall x, y \in Q_{jk} \quad d(x, y) \leq \delta_{jk}. \end{aligned}$$

Любое из возможных множеств Q_{jk} добавляем в обучающую выборку. Условие отбора элементов множества Q_{jk} означает их попарную удаленность между собой.

Подход 2.

Для каждого класса объектов k найдем δ_{jk} , при котором выполняется условие:

$$\begin{aligned} \exists Q_{jk} \subset X_k \setminus \bigcup_{i=0}^{j-1} Q_{ik} : |Q_{jk}| = N_j, \\ \forall x, y \in \bigcup_{i=0}^j Q_{ik} \quad d(x, y) \leq \delta_{jk}. \end{aligned}$$

Любое из возможных множеств Q_{jk} добавляем в обучающую выборку. Элементы набранного множества Q_{jk} должно быть не только попарно удалены между собой, но тоже должно выполняться и для ранее добавленных в обучающую выборку элементов.

2. Обучение классификатора рукописных цифр

Первый эксперимент проведен на примере обучения сверточной нейронной сети архитектуры LeNet-5 на эталонном множестве изображений MNIST DIGITS [9] (рис. 1).

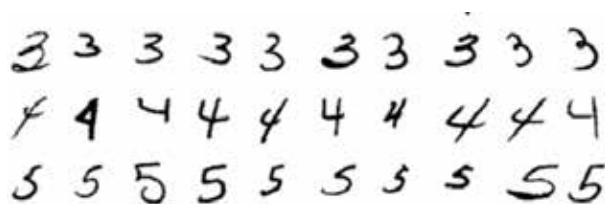


Рис. 1. Примеры изображений рукописных цифр набора данных MNIST DIGITS

Для архитектуры LeNet-5 в качестве характеристических векторов использовались активации предпоследнего полносвязного слоя (inner_product_1). В ходе эксперимента размер обучающей выборки для каждого класса увеличился с 50 до 1000 элементов. Качество оценивалось как показатель правильности (ассигасу) для обученных сетей при распознавании на тестовой выборке.

Для снижения влияния стохастической природы метода обучения сетей для каждого подхода итеративная процедура обучения запускалась 10 раз, а показатели качества усреднялись. Для контроля обучения сетей той же архитектуры на случайных выборках такого же размера проводилось 10 раз с усреднением качества. Потеря качества при сокращении обучающей выборки оценивалась сравнением с качеством сети, обученной на полном обучающем множестве данных MNIST DIGITS. Результаты эксперимента (табл. 1) показывают, что при использовании обоих подходов получаются выборки, которые дают результат обучения лучше, чем случайные выборки аналогичного размера.

Табл. 1

Сравнение усредненных показателей правильности (ассигасы) для сетей архитектуры Lenet-5, обученных на построенных выборках и на случайных

Элементов (на класс)	Подход 1	Подход 2	Случайная выборка
50	0.911	0.909	0.912
100	0.959	0.957	0.939
200	0.977	0.977	0.958
300	0.983	0.982	0.967
400	0.986	0.984	0.971
500	0.987	0.986	0.975
600	0.988	0.986	0.976
700	0.989	0.986	0.979
800	0.990	0.986	0.980
900	0.990	0.986	0.981
1000	0.991	0.986	0.982
6000			0.993

При этом заметно преимущество первого подхода. Более того, правильность при обучении по методу 1 на последних итерациях (0.991) приближается к результатам обучения сети на полном обучающем множестве (0.993).

3. Обучение классификатора дорожных знаков

При разработке систем помощи водителю (ADAS) [10] и беспилотных автомобилей одной из подзадач является детектирование и классификация дорожных знаков. Возможный вариант решения этой задачи состоит в использовании трехэтапной схемы, когда на первом этапе зоны знаков выделяются на ректифицированном кадре [11] бы-

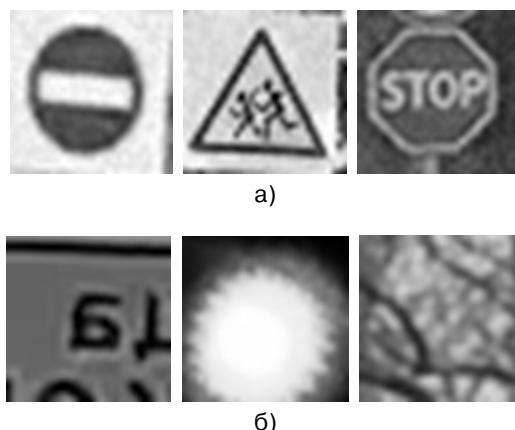


Рис. 2. Примеры детекций дорожных знаков: а) корректные, б) ошибочные

стрым детектором с высокой полнотой, но низкой точностью (рис. 2).

Результаты такого детектирования распознаются фильтрующим нейросетевым классификатором. Низкая точность детектора может компенсироваться локальным уточнением зоны знака в сторону коррекции масштаба и повторного распознавания. Сеть-фильтр принимает решение, является ли зона детекции знаком и надо ли немного изменить границы зоны для знака. Затем изображения знаков подаются на финальный классификатор, который определяет точный тип знака.

Особенностью задачи обучения фильтрующего классификатора является существенный дисбаланс представителей классов, когда количество ошибочных детекций во много раз превышает количество знаков. Вторая особенность состоит в использовании кадров видеопоследовательности, как первичного источника данных. Зоны детекций на соседних кадрах последовательности могут содержать идентичные или очень близкие изображения. В таких условиях особенно остро требуется отбирать информативные примеры для обучения.

Для экспериментальной проверки алгоритма активного отбора информативных элементов в процессе обучения (подход 1) рассмотрена задача обучения сверточной классифицирующей сети. Общее количество в исходном наборе данных объектов типа «знак» 31235, типа «не знак» 1052012. Для отбора информативного множества для типа «не знак» на каждой итерации по предложенной выше схеме из еще не рассмотренных элементов типа «не знак» случайно выбиралось подмножество размера 1000 элементов, из которого уже активно отбиралось 50 элементов для расширения обучающего множества. Итеративная процедура заканчивалась при достижении 30000 элементов

типа «не знак» в обучающей выборке. Для сравнения отбирался случайный набор обучающих примеров с такой же численностью по классам. На каждом из наборов обучалась нейронная сеть, а качество распознавания измерялось как правильность (ассигасу) на полном исходном наборе данных. При использовании активной схемы отбора общая правильность составила 0.99831 против 0.99078 при случайном наборе, т.е. отобранные при помощи предложенной схемы примеры служат более информативным материалом для обучения, чем набранные случайным образом.

Стоит отметить, что рассматриваемый активный отбор информативных элементов для обучения нейросетевого классификатора может использоваться совместно с методами аугментации обучающей выборки для выбора «более полезных» для обучения вариантов искажения исходных примеров, и естественным образом может включаться в схемы параллельного обучения классификаторов. Экспериментальная проверка эффективности такого комбинирования оставлена для дальнейших исследований.

Заключение

Для прореживания обучающих выборок предложена схема активного отбора информативных элементов в процессе обучения легкой сверточной классифицирующей сети. Проведенные эксперименты на открытом массиве данных MNIST DIGITS и собственном наборе для классификации зон детекций дорожных знаков показали стабильное преимущество предложенной схемы по сравнению с отбором элементов, проводимым случайным образом.

Литература

1. He H., Garcia E.A. Learning from imbalanced data // IEEE Transactions on Knowledge and Data Engineering – 2009. – V. 21 – № 9 – С. 1263–1284.
2. Buda M., Maki A., Mazurowski M.A. A systematic study of the class imbalance problem in

- convolutional neural networks / – 2017. – С. 1–23.
3. Settles B. Active learning literature survey // Madison.–University of Wisconsin. –2010 – 11 с.
4. Dasgupta S. Analysis of a greedy active learning strategy // Advances in Neural Information Processing Systems.–Saul.–2014.
5. Elman J.L. Learning and development in neural networks: the importance of starting small // Cognition. –1993.–V.48.–№1.– С.71– 99.
6. Bengio Y., Louradour J., Collobert R. et. al. Curriculum learning // Proceedings of the 26th Annual International Conference on Machine Learning, ser. ICML '09, Montreal, Quebec, Canada.–ACM. –2009. – С. 41–48.
7. Huang C., Li Y., Loy C.C., Tang X. Learning Deep Representation for Imbalanced Classification // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, С. 5375-5384.
8. Yu H., Zhang Z., Qin Z., Wu H., Li D., Zhao J., Lu X. Loss Rank Mining: A General Hard Example Mining Method for Real-time Detectors // – 2018. [Электронный ресурс] arXiv.org URL: <https://arxiv.org/abs/1804.04606> (дата обращения: 06.07.218).
9. LeCun Y., Bottou L., Bengio Y., et. al. Gradient-based learning applied to document recognition // Proc. IEEE.–86(11). –1998. – С. 2278–2324.
10. Prun V.E., Postnikov V.V., Sadekov R.N., Sholomov D.L. “Development of Active Safety Software of Road Freight Transport, Aimed at Improving Inter-City Road Safety, Based on Stereo Vision Technologies and Road Scene Analysis” // Proceedings of the Scientific-Practical Conference “Research and Development – 2016”, Springer, Cham, pp.209-218. – 2017. ISBN 978-3-319-62869-1
11. Prun V., Polevoy D., Postnikov V. Forward rectification: spatial image normalization for a video from a forward facing vehicle camera // Proc. SPIE 10341, Ninth International Conference on Machine Vision (ICMV 2016), T. 10341, С. 103410W

Иванов Денис Евгеньевич. ООО «Когнитивные технологии», г. Москва, Россия. Младший разработчик. Национальный исследовательский технологический университет «МИСиС», г. Москва, Россия. Баклавр. E-mail: mr.salixnew@gmail.com

Полевой Дмитрий Валерьевич. Институт системного анализа Федерального исследовательского центра «Информатика и управление» Российской академии наук, г. Москва, Россия. Старший научный сотрудник, кандидат технических наук. Национальный исследовательский технологический университет «МИСиС», г. Москва, Россия. Доцент. Количество печатных работ: более 20. Область научных интересов: цифровая обработка изображений, компьютерное зрение, искусственный интеллект. E-mail: dvpsun@gmail.com

Шоломов Дмитрий Львович. Институт проблем передачи информации им. А.А. Харкевича Российской академии наук г. Москва, Россия. Старший научный сотрудник, кандидат технических наук. Национальный исследовательский технологический университет «МИСиС», г. Москва, Россия. Старший преподаватель. Количество печатных работ: более 20. Область научных интересов: цифровая обработка изображений, компьютерное зрение, искусственный интеллект. E-mail: sholomov@list.ru

Selection of informative elements for the training of a lightweight convolutional neural network classifier in the conditions of a strong imbalance of the training sample

D.E. Ivanov^{I,II}, D.V. Polevoy^{I,III,IV}, D.L. Sholomov^{I,V}

^I National university of science and technology "MISIS", Moscow, Russia

^I Cognitive Technologies, Moscow, Russia

^I Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russia

^I Moscow Institute of Physics and Technology, Dolgoprudny, Moscow region, Russia

^I The Institute for Information Transmission Problems of Russian Academy of Sciences, Moscow, Russia

Abstract. In this paper we consider the task of balancing a training sample in the training of an image classifier based on a convolutional neural network. An active scheme of selection (thinning out) of informative elements in the learning process is proposed to teach the lightweight classifier in the conditions of a strong imbalance of the training samples. An experimental verification using the example of the problem of classifying images of handwritten figures and zones of road sign detection shows a stable advantage of the proposed scheme in comparison with random selection.

Keywords: *pattern recognition, convolutional neural networks, sampling, active learning, curriculum learning.*

DOI: 10.14357/20790279180523

References

1. He H., Garcia E.A. Learning from imbalanced data // IEEE Transactions on Knowledge and Data Engineering – 2009. – V. 21 – № 9 – P.1263–1284.
2. Buda M., Maki A., Mazurowski M.A. A systematic study of the class imbalance problem in convolutional neural networks / – 2017. – P.1–23.
3. Settles B. Active learning literature survey // Madison.–University of Wisconsin. –2010. – 11 p.
4. Dasgupta S. Analysis of a greedy active learning strategy // Advances in Neural Information Processing Systems.–Saul.–2014.
5. Elman J.L. Learning and development in neural networks: the importance of starting small // Cognition.–V.48.–№1.–P.71– 99.–1993.
6. Bengio Y., Louradour J., Collobert R. e.a. Curriculum learning // Proceedings of the 26th Annual International Conference on Machine Learning, ser. ICML '09, Montreal, Quebec, Canada.–ACM.–2009.–P. 41–48.
7. Huang C., Li Y., Loy C.C., Tang X. Learning Deep Representation for Imbalanced Classification // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, –2016. P. 5375-5384.
8. Yu H., Zhang Z., Qin Z., Wu H., Li D., Zhao J., Lu X. Loss Rank Mining: A General Hard Example Mining Method for Real-time Detectors // – 2018. Available at: <https://arxiv.org/abs/1804.04606> (accessed July 6, 2018).
9. LeCun Y., Bottou L., Bengio Y., e.a. Gradient-based learning applied to document recognition // Proc. IEEE.–86(11).–P. 2278–2324.–1998.
10. Prun V.E., Postnikov V.V., Sadekov R.N., Sholomov D.L. "Development of Active Safety Software of Road Freight Transport, Aimed at Improving Inter-City Road Safety, Based on Stereo Vision Technologies and Road Scene Analysis" // Proceedings of the Scientific-Practical Conference "Research and Development – 2016", Springer, Cham, pp.209-218. – 2017. ISBN 978-3-319-62869-1
11. Prun V., Polevoy D., Postnikov V. Forward rectification: spatial image normalization for a video from a forward facing vehicle camera // Proc. SPIE 10341, Ninth International Conference on Machine Vision (ICMV 2016), T. 10341, C. 103410W

D.E. Ivanov Bachelor of Applied Mathematics. Cognitive Technologies, Moscow, Russia. Developer. E-mail: mr.salixnew@gmail.com

D.V. Polevoy. Ph.D. Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Moscow, Russia, Senior researcher. National university of science and technology ”MISiS”, Moscow, Russia, Associate professor. Number of the printed works: more than 20. Range of scientific interests: image processing, computer vision, artificial intelligence, deep learning. e-mail: dvpsun@gmail.com

D.L. Sholomov. Ph.D. The Institute for Information Transmission Problems of Russian Academy of Sciences, Moscow, Russia. Senior researcher. National university of science and technology ”MISiS”, Moscow, Russia, Senior lecturer. Number of the printed works: more than 20. Range of scientific interests: image processing, computer vision, artificial intelligence. e-mail: sholomov@list.ru