

АУГМЕНТАЦИЯ ДАННЫХ ДЛЯ РЕГУЛЯРИЗАЦИИ В НЕЙРОСЕТЕВЫХ АЛГОРИТМАХ ПОВТОРНОЙ ИДЕНТИФИКАЦИИ ЛЮДЕЙ ПО ВИДЕОДАНЫМ

С.А. Игнатьева¹, Р.П. Богущ¹, С.В. Абламейко^{2,3}

¹ *Полоцкий Государственный Университет имени Ефросинии Полоцкой*
ул. Блохина, д. 29, г. Новополоцк, Беларусь,
igniteva604@gmail.com, r.bogush@psu.by

² *Белорусский Государственный Университет*
п-т. Независимости, д. 4, г. Минск, Беларусь

³ *Объединенный институт проблем информатики НАН Беларуси*
ул. Сурганова, 6, г. Минск, Беларусь, ablameyko@bsu.by

Предлагается новый способ аугментации данных и подход к ее применению для регуляризации при обучении сверточной нейронной сети (СНС) для решения задачи повторной идентификации людей при видеонаблюдении. Для аугментации используется циклический сдвиг изображения по вертикали и горизонтали на случайное число пикселей и добавление уменьшенной копии другого изображения. Кроме этого, особенностью является применение двухэтапного обучения СНС. На первом этапе сеть обучается на аугментированных данных, а на втором продолжает обучение на исходных изображениях. Показано, что предлагаемый подход позволяет увеличить значения трех метрик mAP, Rank1 и mINP.

Ключевые слова: Ре-идентификация человека; сверточные нейронные сети; регуляризация; аугментация данных.

DATA AUGMENTATION FOR REGULARIZATION IN NEURAL NETWORK PERSON RE-IDENTIFICATION ALGORITHMS BY VIDEO DATA

S.A. Ihnatsyeva¹, R.P. Bohush¹, S.V. Ablameyko^{2,3}

¹ *Euphrosyne Polotskaya State University of Polotsk, Blokhina 29, Novopolotsk, Belarus,*
igniteva604@gmail.com, r.bogush@psu.by

² *Belarusian State University, pr. Nezavisimosti 4, Minsk, Belarus*

³ *United Institute of Informatics Problems, National Academy of Sciences of Belarus,*
Surganova 6, Minsk, Belarus, ablameyko@bsu.by

A new data augmentation method and an approach to its application for regularization in training a convolutional neural network (CNN) for solving the people re-identification problem in video surveillance are proposed. The method uses two key techniques. First, the cyclic shift is applied to the image pixels in two directions, namely vertically and horizontally. Secondly, subimage replacement of with a reduced copy of another image from

the package is applied. In addition, feature is use of two-stage training CNN. It is shown that the proposed approach allows increasing the values mAP, Rank1 and mINP.

Keywords: Person re-identification; convolutional neural networks; regularization; data augmentation.

Введение

Для повторной идентификации (ре-идентификации) человека в распределенных системах видеонаблюдения наиболее эффективным способом извлечения признаков является использование сверточных нейронных сетей (СНС) [1]. Однако, недостаточное количество обучающих данных может приводить к переобучению и неустойчивости обученной модели к новым данным. Для решения проблемы переобучения одним из способов является регуляризация данных. Под регуляризацией в машинном обучении понимают добавление ограничений к архитектуре нейронной сети, функции потерь или входным данным [2]. Техника регуляризации позволяет избежать СНС запоминания обучающих примеров и улучшить обобщающую способность. Примерами регуляризации являются: прореживание нейронных связей СНС; L1- и L2-регуляризация; аугментация данных, при которой часть изображения удаляется или заменяется другой информацией.

Под аугментацией понимают расширение обучающей выборки путем выполнения каких-либо преобразований изображения без привлечения дополнительных данных. Примерами аугментации данных является добавление таких преобразований изображения как поворот, отражение, изменение яркости, контраста и другие, позволяющие увеличить разнообразие обучающей выборки.

1. Анализ существующих подходов

При использовании аугментации данных для регуляризации могут применяться два подхода: преобразование выполняется с заданной степенью вероятности к тензору загружаемого изображения [3] или к тензору пакета изображений [2, 4, 5].

В [3] предлагается метод *random erasing*, который требует удаление фрагмента из исходного изображения. При загрузке пакета изображений на разных эпохах обучения одно и то же изображение может быть загружено как с исходном виде, так и с отсутствующими фрагментами. Это позволяет СНС повысить устойчивость к окклюзиям.

Применение преобразования к пакету изображений рассматривается в [2], где предлагается метод регуляризации с использованием аугмента-

ции CutMix. Суть заключается в том, что фрагмент одного из изображений в пакете накладывается на другое изображение из пакета. Применение CutMix позволяет повысить устойчивость к окклюзиям и более детально изучать отдельные признаки изображений. Схожая идея используется в алгоритме Mixed Single Thumbnail (MST) [4], только фрагмент изображения замещается на уменьшенную копию другого изображения из пакета. Это, повышает устойчивость к окклюзиям и низкому разрешению изображений. Одновременно два изображения обрабатывается и при использовании аугментации mixup [5], при которой два изображения, выбранные из пакета, объединяются с помощью линейной интерполяции. Mixup позволяет улучшить обобщающую способность обученной СНС и снижает чувствительность к запоминанию ошибочных меток.

В [6] предлагается подход, состоящий из трех этапов: На первом этапе модель обучается на исходных неаугментированных данных. На втором этапе определяются области изображения, оказывающие наибольшее влияние на точность предсказания. Для этого методом скользящего окна закрываются разные участки изображения и оценивается, как это влияет на результат. При обнаружении области изображения, без которой сеть не может верно определить класс объекта (идентификационную метку человека), эта область закрывается. На третьем этапе сеть продолжает обучение на изображениях, объединяющих как исходные изображения, так и изображения со скрытыми фрагментами. Такой подход к применению аугментации вынуждает сеть искать не только наиболее очевидные признаки, но и дополнительные особенности, которые могут повысить точность обученной модели.

2. Аугментация с использованием циклического сдвига пикселей в изображении и замещены фрагмента в нем

Предлагаемый метод аугментации (ЦЗФ) состоит из двух шагов: циклический сдвиг изображения по вертикали и горизонтали на некоторое количество пикселей; при формировании пакета, фрагмент изображения замещается другим, уменьшенным изображением. На рисунке представлен пакет из восьми изображений. Изображения (в), д), е), з)) смещены по вертикали и горизонтали. Для всех изображений выполнена замена фрагмента на уменьшенное изображение из пакета. На каждой итерации для выбора уменьшенного изображения используется генератор псевдослучайных чисел. Уменьшенные изображения содержат циклический сдвиг в примерах а), б), в), г). На примере ж) ни к самому изображению, ни к уменьшенному изображению, циклический сдвиг не применялся. При загрузке следующего пакета уменьшенные копии будут иметь другой раз-

мер и место расположения или отсутствовать, а циклический сдвиг будет применяться к иным изображениям.



а) б) в) г) д) е) ж) з)

Пример применения метода аугментации ЦЗФ.

Такой метод позволяет сети повысить устойчивость к местоположению признаков, входному разрешению изображения и окклюзиям, при этом не требует больших вычислительных затрат.

Для снижения значения функции потерь при обучении, вызванной неправдоподобностью обучающих данных по отношению к тестовым данным, к которым аугментация не применяется, предлагается подход, использующий циклический сдвиг пикселей и замещение фрагмента уменьшенной копией другого изображения с применением двухэтапного обучения (ЦЗФД), основная идея которого заключается в том, что обучение включает два этапа. На первом этапе сеть обучается на аугментированных данных, а на втором продолжает обучение на исходных изображениях. Это позволяет более близко подойти к локальному минимуму функции потерь и минимизировать ошибку. Таким образом, сначала сеть учится быть устойчивой к окклюзиям, разрешению изображения и независимости от местоположения признаков, а затем на более низкой скорости обучения корректируется, что позволяет извлекать более надежные признаки.

Предлагаемый нами подход к применению аугментации ЦЗФД заключается в следующем:

- скорость обучения всех слоев СНС в 0.1 раз меньше скорости обучения классификационного слоя, а после 40 эпохи скорость обучения снижается в 0.1 раз;
- первые 45 эпох обучения выполняются с применением аугментации ЦЗФ;
- в течении последующих 15 эпох обучение продолжается на данных без аугментации.

3. Результаты экспериментов

Для определения наиболее эффективного метода аугментации данных использовался алгоритм повторной идентификации [7] реализованный на фреймворке pyTorch. Обучение модели выполнено для ResNet-50 и DenseNet-121 без применения предварительного обучения, в течение 60 эпох со скоростью 0,07 для классификационного слоя, и 0,007 для всех остальных слоев СНС, уменьшенными после 40-й эпохи в 0,1 раз, с размером пакета 16, на наборе данных Market-1501. Вероятность применения метода аугментации при проведении экспериментов составляет 0,5. Для ЦЗФД, объединяющего применение аугментации и двухэтапное обучение, значения гиперпараметров не изменяются, но после 45 эпохи на вход СНС подаются исходные изображения. Результаты экспериментов приведены в таблице.

Сравнение методов аугментации для СНС ResNet-50 и DenseNet-121 на наборе данных Market-1501

Метод	ResNet-50			DenseNet-121		
	mAP	Rank1	mINP	mAP	Rank1	mINP
Baseline[6]	61,20	82,78	0,238	61,40	82,45	0,248
Random erasing[1]	63,38	84,26	0,255	64,39	83,31	0,276
CutMix[2]	64,43	85,24	0,274	64,87	84,14	0,277
MixUP[4]	62,34	83,81	0,246	64,43	85,12	0,275
MST[3]	65,65	84,94	0,286	66,25	84,74	0,299
ЦЗФ	65,67	85,84	0,294	66,23	84,44	0,308
ЦЗФД	67,65	86,31	0,326	67,56	85,72	0,328

Анализ полученных результатов, представленных в таблице, показывает, что использованием ЦЗФД позволяет значительно повысить точность повторной идентификации по всем трем метрикам оценки точности для двух рассмотренных архитектур относительно базовой модели Baseline, а также увеличить точность по сравнению с существующими методами.

Заключение

В работе предлагается подход, отличающийся тем, что совмещает циклический сдвиг пикселей изображения, замещение фрагмента уменьшенной копией другого и использует двухэтапное обучение. Во-первых, такой подход позволяет получить большее разнообразие данных, чем если два выполняемых преобразования применять только к одному изображению [3], или только к пакету изображений [2,4], и позволяет повысить устойчивость обученной СНС к местоположению признаков, низкому раз-

решению изображения, окклюзиям и переобучению. Во-вторых, использование двухэтапного обучения позволяет минимизировать значение функции потерь и повысить точность обученной модели. То, что на разных этапах обучения используются различные данные, не позволяет СНС запомнить обучающие примеры, тем самым предотвращая переобучение.

Библиографические ссылки

1. Ye S., Bohush R., Chen C., Zakharova I., Ablameyko S. Person Tracking and Re-Identification in Video for Indoor Multi-Camera Surveillance Systems *Pattern Recognition and Image Analysis*, 2020. Vol. 30, № 4 P. 827–837. DOI: 10.1134/S1054661820040136
2. Yun S., Han D., Oh S., Chun S., Choe J., Yoo Y.J. CutMix: Regularization Strategy to Train Strong Classifiers With Localizable Features. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019, P. 6022–6031. DOI:10.1109/ICCV.2019.00612
3. Zhong Z., Zheng L., Kang G., Li S., & Yang, Y. Random Erasing Data Augmentation // *AAAI*. 2020. DOI:10.1609/AAAI.V34I07.7000
4. Xie T., Cheng X., Wang X., Liu M., Deng J., Zhou T., Liu M. Cut-Thumbnail: A Novel Data Augmentation for Convolutional Neural Network // *Proceedings of the 29th ACM International Conference on Multimedia*, 2021. DOI:10.1145/3474085.3475302
5. Zhang H., Cissé M., Dauphin Y., Lopez-Paz D. mixup: Beyond Empirical Risk Minimization. URL: <https://arxiv.org/abs/1710.09412>, 2018; (дата обращения: 09.08.2022). DOI:10.48550/arXiv.1710.09412
6. Huang H., Li D., Zhang Z., Chen X., Huang K. Adversarially Occluded Samples for Person Re-identification. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018. P. 5098-5107. DOI:10.1109/CVPR.2018.00535
7. Person ReID baseline pyTorch. URL: https://github.com/layumi/Person_reID_baseline_pytorch (дата обращения: 09.08.2022)