

Аугментация данных на основе вейвлет-фильтрации при обучении нейронных сетей

К.А. Алимагадов¹, С.В. Умняшкин¹

¹ Национальный исследовательский университет «МИЭТ», пл. Шокина, д. 1, Зеленоград, Москва, 124498, Россия

Аннотация

Исследование посвящено изучению влияния аугментации данных с применением вейвлет-фильтрации на точность работы нейросетевого детектора при распознавании зашумлённых изображений. Предлагается добавлять шум к примерам обучающего набора с последующей их обработкой фильтром Винера в области дискретного вейвлет-преобразования (ДВП). При генерации входных обучающих изображений используется модель белого аддитивного гауссового шума, с помощью которого выполняется их искажение. Помимо винеровской фильтрации для шумоподавления рассматриваются методы пороговой обработки в области ДВП. На основе полученных экспериментальных данных проводится сравнение результатов работы детектора при различных значениях среднеквадратического отклонения шума. Также был рассмотрен вариант, когда нейронная сеть обучалась на зашумлённых изображениях без применения последующей фильтрации. Демонстрируется, что предложенный подход позволяет добиться большей точности распознавания, по сравнению с аугментацией, не использующей подавление шумов.

Ключевые слова

Аугментация данных, машинное обучение, нейронные сети, вейвлет-фильтрация.

Data Augmentation Based on Wavelet Filtration During Neural Network Training

К.А. Alimagadov¹, S.V. Umnyashkin¹

¹ National Research University of Electronic Technology, Shokin sq., 1, Zelenograd, Moscow, 124498, Russia

Abstract

The research is devoted to studying the influence of data augmentation by means of wavelet filtration on recognition accuracy of noised images by the neural network detector. It is proposed adding noise to samples of the training dataset and then processing them by Wiener filter in the domain of discrete wavelet transform (DWT). Generating of input training images uses the model applying white additive Gaussian noise to distort them. In addition to the Wiener filtering, we considered some thresholding methods in the domain of DWT for images denoising. Based on the obtained experimental data, we compare the results of the detector recognition for the different values of noise standard deviation. Also, we considered the case, when the neural network was trained on noised-augmented images without subsequent filtering. It is shown that the proposed approach allows higher accuracy of recognition to be achieved, than the augmentation, which doesn't use noise suppression.

Keywords

Data augmentation, machine learning, neural networks, wavelet filtration.

ГрафиКон 2023: 33-я Международная конференция по компьютерной графике и машинному зрению, 19-21 сентября 2023 г., Институт проблем управления им. В.А. Трапезникова Российской академии наук, г. Москва, Россия

EMAIL: alimagadovk@yandex.ru (К.А. Алимагадов); vrinf@micee.ru (С.В. Умняшкин)

ORCID: 0009-0008-3420-3048 (К.А. Алимагадов); 0009-0000-1932-8709 (С.В. Умняшкин)



© 2023 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

1. Введение

Аугментация данных широко применяется при обучении современных нейронных сетей, в том числе предназначенных для решения задач компьютерного зрения [1, 2]. Она позволяет увеличить объём обучающей выборки и тем самым снижает риск переобучения модели, не требуя при этом дополнительного сбора новых обучающих примеров.

Настоящая работа посвящена изучению аугментации, используемой при обучении сетей с целью повышения их устойчивости к возникновению шумов на входных изображениях (в смысле незначительного снижения точности распознавания). Для решения этой задачи были рассмотрены несколько случаев:

1. аугментация с добавлением шума в процессе обучения;
2. аугментация с добавлением шума и последующей фильтрацией в процессе обучения (при использовании обученной сети на вход подаются отфильтрованные изображения);
3. обучение без аугментации.

Первый вариант является одной из классических техник аугментации и реализован в ряде специальных библиотек [1, 3], применяемых при обучении нейросетевых моделей.

Предложенный нами второй вариант основывается на гипотезе о том, что извлечение признаков на отфильтрованных изображениях является более простой задачей. Это связано с тем, что фильтрация приводит к снижению доли шумов, сохраняя при этом полезный сигнал, несущий в себе информацию, используемую при распознавании. Предполагается, что сеть, обученная с такой аугментацией, будет демонстрировать большую точность на данных с подавленными шумами, по сравнению с вариантом, когда и при обучении, и при непосредственной работе сети на вход подаются зашумлённые изображения.

Третий случай рассматривается для того, чтобы относительно него оценить прирост точности для первого и второго вариантов.

Для подавления шумов были выбраны методы фильтрации в базе дискретного вейвлет-преобразования (ДВП). Использование вейвлет-фильтрации для решения этой задачи основывается на результатах предыдущих исследований [4, 5]. Было установлено, что данный подход обеспечивает более качественную обработку в смысле роста мер SNR [6] и SSIM [7], чем другие рассмотренные методы.

2. Модель шума и методы вейвлет-фильтрации

В процессе обучения для аугментации данных использовалась модель белого аддитивного гауссового шума, функция плотности распределения которого имеет вид:

$$\rho(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (1)$$

где μ – математическое ожидание (будем полагать, что $\mu = 0$), σ – среднеквадратическое отклонение (СКО) шума. При обучении сети значение σ будет генерироваться для каждого изображения как реализация непрерывной случайной величины, имеющей равномерное распределение $\sigma \in U_{a,b}$. Это необходимо для того, чтобы нейронная сеть могла извлекать признаки при различных уровнях шума, не подстраиваясь под конкретное значение СКО.

Для подавления шума были рассмотрены несколько методов фильтрации в области ДВП:

- жёсткая пороговая фильтрация [4];
- мягкая пороговая фильтрация [4];
- винеровская фильтрация [4];
- винеровская фильтрация с применением нейронных сетей [5].

Проверка гипотезы об аугментации с подавлением шумов проводилась на примере решения задачи распознавания автомобильных номеров с применением нейронной сети SqueezeDet [8, 9]. Номер считается распознанным, если верно определены все составляющие его символы. Обучающий набор данных состоит из 45000 изображений, тестовый – из 2500.

Определим качество шумоподавления для тестовых изображений указанными выше фильтрами с помощью усреднённых значений мер SNR и SSIM, вычисленных по всему тестовому набору. Метод фильтрации, который продемонстрирует наилучшие результаты, будем использовать в дальнейшем для формирования примеров при обучении нейронной сети.

Результаты обработки изображений тестового набора данных при 3-х уровнях шума: низком ($\sigma = 7$), среднем ($\sigma = 15$) и высоком ($\sigma = 30$) – представлены в таблице 1. Наибольшие значения в каждой строке выделены жирным шрифтом.

Таблица 1 – Результаты фильтрации тестовых изображений

Мера оценки качества	Зашумлённое изображение	Жёсткий пороговый фильтр	Мягкий пороговый фильтр	Фильтр Винера	Фильтр Винера (с нейр. сетями)
$\sigma = 7$					
SNR (средн.) [дБ]	24,07	26,86	27,25	28,95	28,56
SSIM (средн.)	0,81	0,92	0,92	0,94	0,93
$\sigma = 15$					
SNR (средн.) [дБ]	17,86	22,25	22,46	24,32	24,23
SSIM (средн.)	0,59	0,84	0,84	0,87	0,86
$\sigma = 30$					
SNR (средн.) [дБ]	12,17	18,15	18,35	20,03	20,19
SSIM (средн.)	0,36	0,73	0,73	0,76	0,75

На основе полученных средних значений SNR и SSIM можно сделать вывод, что винеровская фильтрация в области ДВП лучше справляется с подавлением шумов, чем пороговые методы. Помимо этого, ещё одним преимуществом фильтра Винера является то, что для формирования его коэффициентов в случае белого шума достаточно знать только оценку СКО шума [4], в отличие от методов, требующих эмпирического подбора порога фильтрации. Исходя из этого, винеровская фильтрация в области ДВП является более предпочтительной для генерации примеров при аугментации. Отметим, что модификация, использующая нейронные сети для формирования коэффициентов фильтра, превосходит по SNR свой аналог без применения сетей только в случае высокого уровня шума.

Поскольку фильтрация будет применяться ко всем изображениям обучающего набора каждую эпоху обучения, то в целях экономии времени при аугментации мы будем использовать вариант фильтра Винера без применения сетей, в том числе и для $\sigma = 30$. При генерации обучающих примеров применялись и другие виды аугментации (сдвиг, поворот, изменение масштаба, яркости, контраста и т.д.), затем добавлялся гауссов шум и выполнялось его подавление (см. рисунок 1).

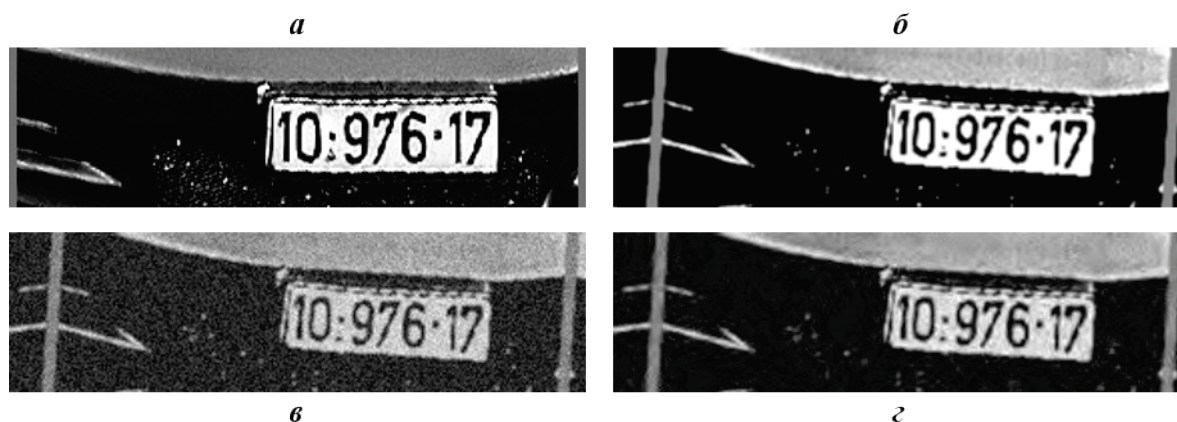


Рисунок 1 – Пример изображения (а), применение к нему поворота, изменения яркости и контраста (б), добавление шума (в), результат шумоподавления фильтром Винера (г)

3. Параметры обучения и результаты экспериментов

Обучение сетей [8] выполнялось с помощью метода Adam [10] со значениями параметров: $\alpha = 10^{-4}$; $\beta_1 = 0,9$; $\beta_2 = 0,999$; $\epsilon = 10^{-7}$. Размер одного батча был выбран равным 40, количество эпох обучения выбрано равным 350. Параметры равномерного распределения генерируемых значений СКО $U_{a,b}$: $a = 0$; $b = 30$.

Для сравнения качества работы сетей рассматриваются точность (доля верно распознанных номеров), доля номеров, распознанных с ошибками (хотя бы один символ номера распознан, при этом необязательно верно), и доля пропущенных номеров (ни один символ номера не обнаружен). Результаты работы сетей, полученные на тестовом наборе изображений, зашумлённых при 4-х уровнях шума ($\sigma \in \{0; 7; 15; 30\}$), представлены в таблице 2. Первый столбец таблицы содержит информацию об используемом типе аугментации и методе подавления шумов на тестовом наборе данных. Прочерк “-” обозначает то, что в данном случае аугментация/фильтрация не использовались. Лучшие результаты в каждом столбце выделены жирным шрифтом.

Таблица 2 – Результаты распознавания тестовых изображений при $\sigma \in \{0; 7; 15; 30\}$

Тип аугментации, метод фильтрации	Точность [%]	Доля номеров, распознанных с ошибками [%]	Доля пропущенных номеров [%]
$\sigma = 0$			
- , -	91,9	4,97	3,13
Шум, -	92,01	5,33	2,66
Шум + фильтрация, -	92,08	4,43	3,49
$\sigma = 7$			
- , -	80,74	12,1	7,16
Шум, -	91,72	5,54	2,74
Шум + фильтрация, Винер	92,01	4,61	3,38
Шум + фильтрация, Винер с сетями	92,01	4,61	3,38
$\sigma = 15$			
- , -	52,63	21,63	25,74
Шум, -	91,83	5,4	2,77
Шум + фильтрация, Винер	91,43	5,15	3,42
Шум + фильтрация, Винер с сетями	91,14	5,4	3,46
$\sigma = 30$			
- , -	12,17	19,73	68,1
Шум, -	88,44	8,82	2,74
Шум + фильтрация, Винер	88,44	8,1	3,46
Шум + фильтрация, Винер с сетями	89,27	7,24	3,49

По данным таблицы 2 можно видеть, что в большинстве случаев точность детектора, обученного на отфильтрованных изображениях, выше, чем у его аналога, обученного на зашумлённых изображениях. Для того чтобы получить более наглядное представление о зависимости точности обученных моделей от уровня шумов, были также проведены эксперименты для значений $\sigma \in \{3,5; 11; 22,5; 45\}$ (см. таблицу 3). Особый интерес представляет собой случай $\sigma = 45$, поскольку это значение выходит за диапазон СКО шумов, использованных при обучении.

Таблица 3 – Результаты распознавания тестовых изображений при $\sigma \in \{3,5; 11; 22,5; 45\}$

Тип аугментации, метод фильтрации	Точность [%]	Доля номеров, распознанных с ошибками [%]	Доля пропущенных номеров [%]
$\sigma = 3,5$			
Шум, -	91,79	5,51	2,7
Шум + фильтрация, Винер	91,86	4,68	3,46
Шум + фильтрация, Винер с сетями	91,94	4,61	3,45
$\sigma = 11$			
Шум, -	91,83	5,36	2,81
Шум + фильтрация, Винер	91,86	4,68	3,46
Шум + фильтрация, Винер с сетями	91,86	4,72	3,42
$\sigma = 22,5$			
Шум, -	90,64	6,44	2,92
Шум + фильтрация, Винер	91	5,62	3,38
Шум + фильтрация, Винер с сетями	91,14	5,47	3,39
$\sigma = 45$			
Шум, -	83,84	13,64	2,52
Шум + фильтрация, Винер	82,97	13,57	3,46
Шум + фильтрация, Винер с сетями	84,92	11,84	3,24

Полученные результаты позволяют сделать вывод, что обучение детектора [8] с применением аугментации на основе вейвлет-фильтрации входных изображений позволяет достичь большей точности распознавания, по сравнению с аугментацией на основе зашумления. Так, в самом лучшем случае ($\sigma = 45$) доля верно распознанных тестовых номеров у модели, обученной на отфильтрованных данных, на 1,08 % больше, чем у модели, обученной на зашумлённых данных.

Принимая во внимание то, что в экспериментах с нейронной сетью, обученной на зашумлённых изображениях, наблюдается наименьшая доля пропусков и наибольшая доля номеров, распознанных с ошибками, можно предположить, что при такой аугментации сеть становится чрезмерно чувствительной, чтобы обнаруживать символы даже при высоких уровнях шума. Однако это приводит к увеличению числа ложных срабатываний и неверно классифицированных символов. Детектор, который и при обучении, и при непосредственной работе получает на вход отфильтрованные изображения, допускает меньше подобных ошибок.

4. Заключение

Аугментация данных с добавлением шума и последующей вейвлет-фильтрацией практически во всех экспериментах позволяет достичь точности распознавания, превосходящей результаты, полученные при аугментации, включающей только зашумление. При значениях СКО шума, превышающих $\sigma \in U_{a,b}$, которые использовались при обучении, эта разница становится наиболее заметной.

Предложенный подход требует предварительного шумоподавления для всех изображений, поступающих обученной нейронной сети на вход. При этом следует отметить, что объём вычислений, необходимый для выполнения используемой фильтрации [5], гораздо меньше количества операций, которые совершают современные нейросетевые модели, разрабатываемые для решения задач компьютерного зрения.

5. Список источников

- [1] Alumentations: Fast and Flexible Image Augmentations [Электронный ресурс] / A. Buslaev [et al.] // Information. 2020. № 2. URL: <https://www.mdpi.com/2078-2489/11/2/125> (дата обращения 12.06.2023).
- [2] Vochkovskiy A., Wang C.Y., Liao H.Y.M. YOLOv4: Optimal speed and accuracy of object detection [Электронный ресурс] // arXiv preprint arXiv:2004.10934. 2020. URL: <https://arxiv.org/abs/2004.10934> (дата обращения 15.06.2023).
- [3] imgaug [Электронный ресурс] / Jung A.B. [et al.] // 2020. URL: <https://github.com/aleju/imgaug> (дата обращения 15.06.2023).
- [4] Alimagadov K.A., Umnyashkin S.V. Application of Wiener filter to suppress white noise in images: wavelet vs Fourier basis // 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus). 2021. P. 2059–2063. doi: 10.1109/ElConRus51938.2021.9396470.
- [5] Алимагадов К.А., Умняшкин С.В. Подавление белого шума на изображениях на основе вивнеровской фильтрации в области дискретного вейвлет-преобразования с применением нейросетевых технологий // Изв. вузов. Электроника. 2022. Т. 27. № 6. С. 807–818. doi: 10.24151/1561-5405-2022-27-6-807-818.
- [6] Gonzalez R.C., Woods R.E. Digital image processing. 4th ed. New York, Pearson Education, 2018. 1168 p.
- [7] Image quality assessment: from error visibility to structural similarity / Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli // IEEE Transactions on Image Processing. 2004. Vol. 13. Iss. 4. P. 600–612. doi: 10.1109/TIP.2003.819861.
- [8] Squeezednet: Unified, small, low power fully convolutional neural networks for real-time object detection for autonomous driving / B. Wu [et al.] // Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017. P. 129–137.
- [9] Alimagadov K., Khamukhin A., Umnyashkin S. Symbol Sequence Recognition Using Viterbi Algorithm and Probability Approximation Table // 2023 25th International Conference on Digital Signal Processing and its Applications (DSPA). 2023. P. 1–5, doi: 10.1109/DSPA57594.2023.10113407.
- [10] Kingma D.P., Ba J.L. Adam: a method for stochastic optimization [Электронный ресурс] // arXiv preprint arXiv:1412.6980. 2014. URL: <https://arxiv.org/abs/1412.6980> (дата обращения 15.06.2023).