

Сравнительное исследование методов естественной аугментации изображений

М.В. Порывай

Аннотация—В данной работе рассматриваются методы естественной аугментации изображений, т.е. те, результат применения которых близок к естественным воздействиям на объекты окружающей среды, с которыми модели машинного обучения могут столкнуться в промышленном применении: со стороны погодных условий; связанных с особенностями эксплуатации или неисправностями приборов восприятия окружающего мира (например, камер), и т.д.

Структурирование и анализ методов естественной аугментации изображений могут принести существенную практическую значимость в качестве пособия для исследователей методов машинного обучения, а также тех, кто занимается разработками моделей машинного обучения непосредственно для индустрии.

Известные на данный момент методы естественной аугментации изображений в данной работе было решено разделить на следующие группы: основанные на добавлении на изображение погодных артефактов, артефактов камеры или на подмене фона для основного объекта на изображении.

Также в статье рассмотрены существующие программные библиотеки для аугментации изображений: Albumentations, имеющая интеграцию с PyTorch; ImgAug, предоставляющая широкий спектр возможных естественных аугментаций; Augmentor, предоставляющая возможность добавлять на изображения разнообразные эффекты искажения камеры.

Для экспериментального исследования методов естественной аугментации изображений был выбран набор данных CIFAR-10; аугментации применялись по одной с вероятностью 50% каждая.

Ключевые слова—естественная аугментация изображений, погодные артефакты, артефакты камеры, подмена фона

I. Введение

В настоящее время методы машинного обучения применяются практически в любых областях науки и индустрии. Исследователи машинного обучения постоянно стремятся улучшить качество создаваемых ими моделей. Одним из основных способов улучшить качество модели машинного обучения является расширение обучающей выборки: чем разнообразнее набор входных данных, тем больше у модели шансов найти правильные зависимости между признаками. Классическим методом расширения обучающей выборки является применение аугментации входных данных.

Аугментация данных в общем случае — увеличение выборки входных данных за счет модификации этих данных и создания на их основе дополнительных данных. В данной работе было решено ограничить рассмотрение вариантов входных данных изображениями.

Среди современных методов аугментации изображений есть такие, которые направлены на защиту обучаемой модели в первую очередь от различных атак, например, состязательных. Однако известно, что даже при отсутствии попыток атак со стороны злоумышленников модель в реальных условиях может работать значительно хуже, чем на обучении и валидации. Это является следствием того, что обучающий набор данных всегда является лишь подмножеством генеральной совокупности, которая, вообще говоря, неизвестна.

В данной работе было решено сконцентрироваться на таких методах аугментации изображений, результат применения которых близок к различным естественным воздействиям на объекты окружающей среды, с которыми модели машинного обучения могут столкнуться в промышленном применении: со стороны погодных условий; связанных с особенностями эксплуатации или неисправностями приборов восприятия окружающего мира (например, камер), и т.д. Подобные методы аугментации в данной работе будут называться методами естественной аугментации изображений.

Существующие обзоры подобных методов аугментации изображений, как правило, недостаточно полны. Структурирование и анализ методов естественной аугментации изображений могут принести существенную практическую значимость. Например, если бы исследователи методов машинного обучения, а также те, кто занимается разработками моделей машинного обучения непосредственно для индустрии, могли бы на основе данного анализа более просто и, в то же время, осознанно принимать решения, какие из методов аугментации изображений лучше подойдут для их конкретного случая, это бы сэкономило им достаточное количество времени и затраченных ресурсов.

Оставшаяся часть статьи структурирована следующим образом. В разделе II обсуждаются различные методы естественной аугментации изображений. В разделе III рассмотрены существующие программные решения для аугментации изображений. Раздел IV посвящен описанию проведенного экспериментального исследования для сравнения методов естественной аугментации изображений. В разделе V приводятся текущие итоги исследования и идеи его дальнейшего развития.

II. Обзор методов естественной аугментации изображений

Для проведения аналитического обзора методов естественной аугментации изображений были выдвинуты следующие критерии:

- Возможность применять метод аугментации без дополнительных данных (имея только изображение).

Статья получена 15 сентября 2024.

Максим Викторович Порывай, МГУ им. М.В. Ломоносова, (email: maximporyvay@yandex.ru).

- Возможность применять метод аугментации «на лету».
- Наличие в открытом доступе кода или псевдокода метода.

Известные на данный момент методы естественной аугментации изображений в данной работе было решено разделить на следующие группы:

- Основанные на добавлении на изображение погодных артефактов.
- Основанные на добавлении на изображение артефактов камеры.
- Основанные на подмене фона для основного объекта на изображении.

Одними из классических методов естественной аугментации изображений являются методы, добавляющие на изображения разнообразные погодные артефакты: дождь, снег, туман, и т. п.

Авторы статьи [1] представили набор данных ImageNet-C: к изображениям из тестовой выборки ImageNet были применены различные преобразования, среди которых, в том числе, добавление на изображения эффектов снега, мороза, тумана и т. п. Несмотря на то, что набор данных ImageNet-C обычно используется не для обучения, а для проверки устойчивости модели на датасете ImageNet, открытый исходный код примененных авторами преобразований можно использовать для аугментации. Стоит отметить, что различных преобразований в статье достаточно много; к тому же эти аугментации градуируются по 5 степеням. Также эти аугментации делаются достаточно просто и вследствие этого также быстро, что позволяет использовать их при обучении «на лету». Однако из-за этого результирующие изображения часто получаются недостаточно реалистичными. Примеры изображений с примененными аугментациями из данной статьи представлены на рисунке 1 (сверху).

Авторы статьи [2] исследовали эффективность систем распознавания дорожных знаков, подверженных сбоям в работе камер, что является критической проблемой для безопасности и надежности автономных транспортных средств. Для того, чтобы повысить устойчивость к таким сбоям, в три популярных общедоступных набора данных, используемых для сравнительного анализа систем распознавания дорожных знаков, включили изображения, аугментированные добавлением 13 условий сбоев в работе камеры. Примеры изображений с примененными аугментациями из данной статьи представлены на рисунке 1 (снизу).

При этом нужно обратить внимание, что в реальной жизни объекты существуют в трехмерном пространстве, тогда как изображения являются двумерными сущностями. И из-за этого, в том числе, простые аугментации и могут часто являться не совсем реалистичными. Для того, чтобы получить более реалистичные результирующие изображения, авторы статьи [3] предлагают дополнительно исследовать глубину сцены на изображении (насколько далеко находятся от условной камеры различные объекты). Так, например, на рисунке 2 видно, что при использовании данной техники аугментация тумана выглядит значительно более реалистичной («туман» будто бы проникает вглубь сцены на изображении).

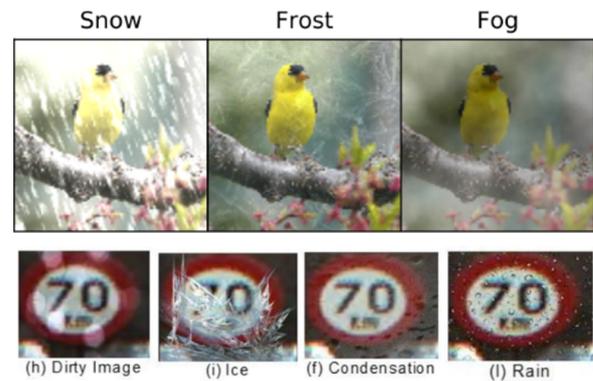


Рис. 1: Примеры аугментаций, основанных на имитации неблагоприятных погодных условий.



Рис. 2: Сравнение «обычной» 2D-аугментации тумана, и более реалистичной для объемной сцены версии.

Для аугментации путем генерации на изображениях погодных артефактов есть также некоторые интересные и более сложные техники. Так, например, авторы статьи [4] предприняли попытку генерировать искусственный дождь различной интенсивности с помощью математической модели, основанной на физических свойствах дождя. К сожалению, пока что метод на основе данной математической модели работает недостаточно быстро для того, чтобы проводить онлайн-аугментацию («на лету»). Хотя рендеринг на основе физики создает реалистичные полосы дождя и эффект тумана, он игнорирует основные характеристики дождя, такие как влажность, отражения, облака, и поэтому может не передавать общий вид дождливой сцены. Напротив, генеративные составительные сети (GAN) отлично справляются с обучением таким визуальным характеристикам, поскольку они являются сильными сигналами для дискриминатора в процессе обучения. Так, для сравнения авторы данной статьи также генерировали искусственные капли дождя на изображениях с помощью генеративно-составительной модели CycleGAN [5], работающей по принципу «image-to-image», и гибридной модели, где вначале изображение прогонялось через CycleGAN, а затем немного изменялось/дополнялось моделью на основе физики дождя. Сравнение полученных изображений приведено на рисунке 3. Методы, использующие CycleGAN лучше подходят для аугментации «на лету» из-за более высокой скорости работы.

Авторы статьи [6] также предложили генерировать погодные артефакты с помощью CycleGAN, но сконцентрировались на имитации загрязнения камеры на изображениях. На рисунке 4 (сверху) представлен первоначальный

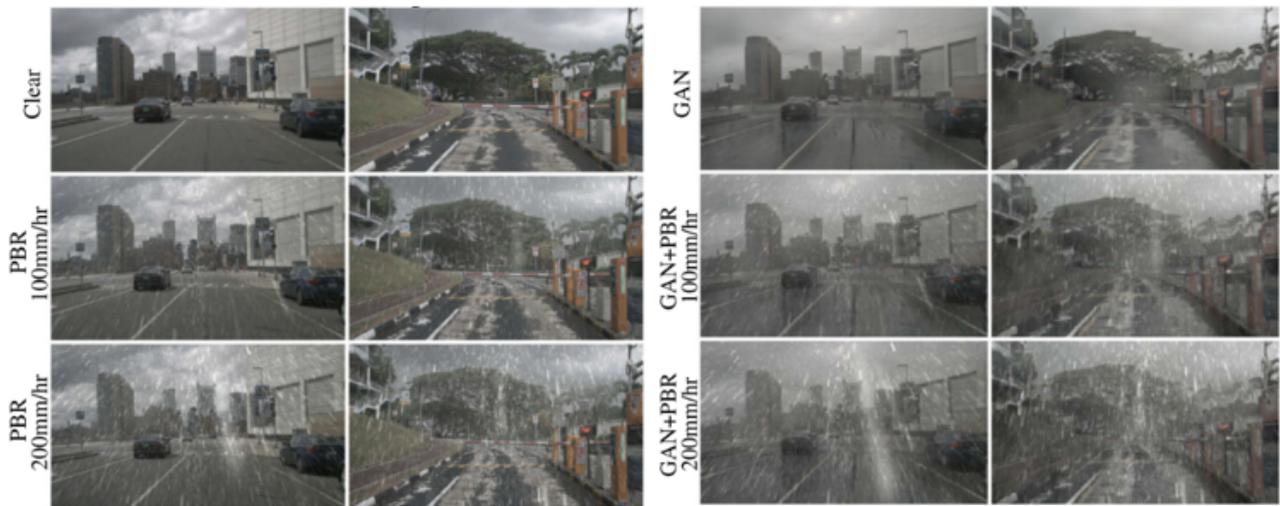


Рис. 3: Сравнение результатов генерации «искусственного» дождя разной интенсивности с помощью GAN, физической и гибридной моделей.



Рис. 4: Примеры генерации «искусственного» загрязнения камеры с помощью GAN.

пайплайн из соответствующей статьи: CycleGAN выдает изображение со сгенерированным «пятном», после чего извлекается только само пятно и накладывается на изначальное изображение. Это было сделано из-за того, что CycleGAN изменяет все пиксели исходного изображения, что не являлось целью работы авторов. Чтобы избавиться от необходимости проводить данные преобразования, модель обучили предварительно генерировать контур (силуэт) пятна, который затем подавался дополнительно к CycleGAN, чтобы тот изменял изображение только внутри данного контура. Примеры результатов работы измененной модели представлены на рисунке 4 (снизу).

Аугментацию изображений можно проводить также на основе добавления артефактов, связанных с эксплуатацией камер, захватывающих эти изображения.

Так, можно, например, имитировать размытие изображения в результате резких движений или увеличение масштаба изображения камеры, расфокусировку камеры и т. п. Это, в том числе, сделали авторы вышеупомянутой работы [1]. Авторы статьи [3] для добавления артефактов камеры также разработали преобразования, опирающиеся на анализ глубины сцены на изображении. Изображения с описанными преобразованиями представлены на рисунке 5.

На реальных изображениях с камер присутствует шум, но он значительно отличается от шума, генерируемого простыми моделями на основе вероятностных распределений. Шум реального изображения представляет собой

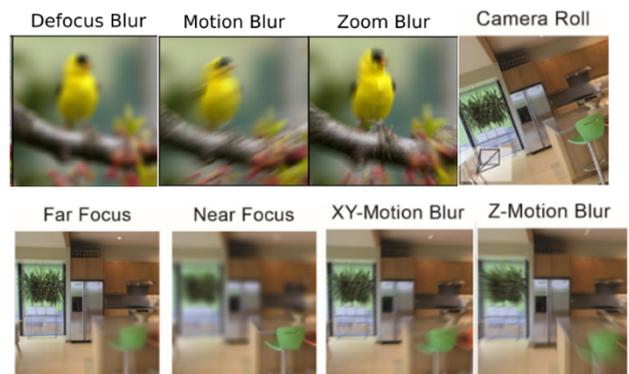


Рис. 5: Примеры аугментаций, основанных на имитации проблем с камерой.

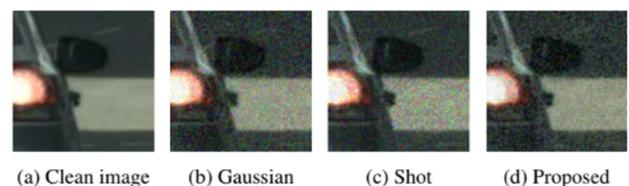


Рис. 6: Сравнение имитации шума с помощью простых моделей и более сложной модели, опирающейся на интенсивность пикселей изображения.

комбинацию множества типов шумов (например, фотонного шума, шума КТК, шума темного тока). Авторы работы [7] предлагают модель шума, которая учитывает наблюдаемое поведение камер. Разработанная ими модель шума зависит от интенсивности пикселей изображения. В соответствии с шумом изображения, наблюдаемым на реальных камерах, пиксели с низкой интенсивностью шумят сильнее, чем пиксели с высокой интенсивностью. Сравнение модели из вышеупомянутой статьи с более простыми моделями шума представлено на рисунке 6.

Также в статье [7] авторы предлагают рассмотреть методы аугментации, основанные на различных искривлениях изображений, получаем с камер. Так, одним из самых известных вариантов искривлений является «рыбий глаз». Пример представлен на рисунке 7.

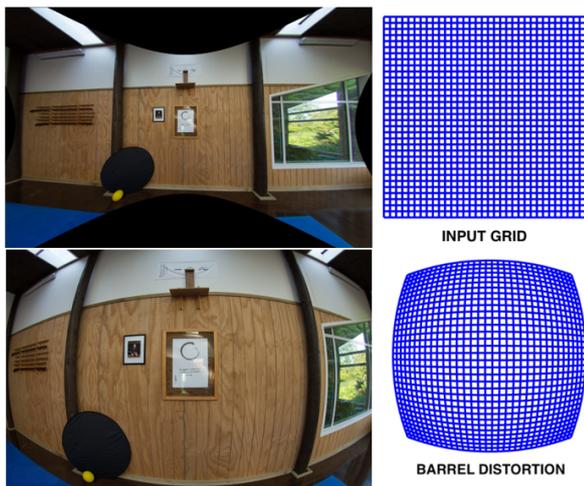


Рис. 7: Пример аугментации, основанный на имитации искривления изображения с камеры.

Таблица I: Сравнение методов естественной аугментации изображений в соответствии с предложенными критериями обзора.

Методы	Возможность применять методы без доп. данных	Возможность применять методы «на лету»	Наличие открытой реализации
Искажения из ImageNet-C	+	+	+
Искажения изображений дорожных знаков	+	+	+
Типичные 3D-искажения	-	+	-
Генерация дождя	-	-	-
Генерация пятен грязи	+	-	-
Реалистичный шум	+	+	-
Геометрические искривления	+	+	-
Замена фона	+	-	-

При обучении модели важно не допустить, чтобы она начала считать характеристики фона объекта его собственными признаками. Такое может произойти, если, например, в обучающей выборке все изображения самолетов будут на фоне безоблачного голубого неба. Или если все изображения артиллерийских орудий, бронемашин, танков будут на фоне коричнево-желтой земли. В такой ситуации правильнее всего дополнять обучающий набор подобными изображениями с разнообразным фоном. Но если это излишне трудозатратно, то можно прибегнуть к аугментации методом подмены фона.

Так, авторы статьи [8] предложили для задачи распознавания людей по камерам наблюдения добавлять в обучающий набор изображения с подмененным фоном — предварительно они сгенерировали определенное количество вариаций фона. Примеры представлены на рисунке 8. Однако необходимо заметить, что если добавить в обучающий набор слишком много вариантов фона для каждого изначального элемента, то точность модели может и не улучшиться — нужно экспериментально находить оптимальное число вариантов фона для каждого



Рис. 8: Примеры выделения силуэта основного объекта на изображении и последующей замены фона.



Рис. 9: Примеры аугментаций с помощью библиотеки Albumentations.

изображения.

Сравнение рассмотренных методов аугментации изображений в соответствии с предложенными критериями обзора представлено в Табл. I.

Более сложные методы аугментации ожидаемо требуют либо больше времени на их применение, либо дополнительных данных для работы. Также лишь малое количество методов имеет открытую реализацию — они и были рассмотрены в экспериментальном исследовании.

III. Обзор программных решений для аугментации изображений

Библиотека Albumentations [9] предоставляет определенное релевантное данной работе подмножество аугментаций. Примеры представлены на рисунке 9. Важно отметить, что данная библиотека имеет интеграцию с PyTorch [10] — главной современной библиотекой для глубокого машинного обучения.

Библиотека ImgAug [11] предоставляет значительно более широкий набор возможных аугментаций. В их число входят как всевозможные «базовые» аугментации, примеры которых представлены на рисунке 10, так и более релевантные данной работе аугментации, генерирующие погодные артефакты, как видно на рисунке 11.

Библиотека Augmentor [12] помимо простейших аугментаций предоставляет возможность добавлять на изображения разнообразные эффекты искажения камеры. Пример представлен на рисунке 12.

Сравнение рассмотренных библиотек для проведения аугментации изображений представлено в Табл. II.

Одним из основных недостатков при использовании этих библиотек в комбинации являются разные требования к предобработке изображений для загрузки в пайплайн аугментации. Было бы полезно разработать библиотеку с едиными требованиями к предобработке

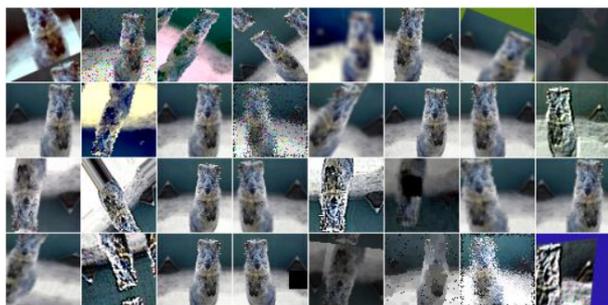


Рис. 10: Примеры «базовых» аугментаций с помощью библиотеки ImgAug.

входных данных, включающую в себя все методы естественной аугментации изображений из рассмотренных библиотек.

Таблица II: Сравнение библиотек для проведения аугментации изображений.

Библиотека	Преимущества	Недостатки
Albumentations	Имеется интеграция с PyTorch	Мало вариантов методов естественной аугментации изображений
ImgAug	Очень широкий выбор быстрых методов естественной аугментации изображений (с возможностью усиления/ослабления отдельных эффектов)	Отсутствие аугментаций методом замены фона
Augmentor	Возможность добавлять на изображения разнообразные эффекты искажения камеры	Мало других методов естественной аугментации изображений

IV. Экспериментальное исследование

В данном разделе представлено экспериментальное исследование методов естественной аугментации изображений. Основной целью экспериментального исследования было сравнение методов естественной аугментации изображений с открытой реализацией из опубликованных статей и открытых программных библиотек.

Для проведения экспериментального исследования был выбран набор данных CIFAR-10 [13]. Это один из наиболее широко используемых наборов данных для исследований в области машинного обучения. Набор данных CIFAR-10 содержит 60000 цветных изображений размером 32x32 дюйма 10 различных классов: самолеты, автомобили, птицы, кошки, олени, собаки, лягушки, лошади, корабли и грузовики. Примеры представлены на рисунке 13. Для каждого класса доступно 6000 изображений. Поскольку изображения в CIFAR-10 имеют низкое разрешение, этот набор данных позволяет быстро опробовать различные алгоритмы.

В качестве вычислительной среды для проведения экспериментов использовалась среда Google Colab со следующими характеристиками:

- Графический процессор: GPU NVIDIA Tesla T4
- ОЗУ графического процессора: 15 ГБ
- ОЗУ системы: 12.7 ГБ
- Объем дискового хранилища: 78.2 ГБ

В качестве классификатора для проведения экспериментального исследования была выбрана модель нейронной сети AlexNet [14].

Разбиение входного набора данных CIFAR-10 было произведено следующим образом:

- Тренировочная выборка: 45000 изображений
- Валидационная выборка: 5000 изображений
- Тестовая выборка: 10000 изображений

Были выбраны следующие ключевые параметры для обучения классификатора:

- Число эпох: 15
- Размер батча: 64
- Шаг обучения: 0.005
- Оптимизатор: стохастический градиентный спуск
- Сокращение веса: 0.005
- Импульс: 0.9
- Функция потерь: перекрестная энтропия

Было решено провести обучение и тестирование классификатора для задачи многоклассовой классификации изображений сначала без использования аугментаций, а затем последовательно с применением аугментаций на основе методов с открытым исходным кодом по одной. В запусках с использованием аугментаций преобразования применялись к каждому изображению с вероятностью 50%.

Были выбраны следующие виды аугментаций:

- Искажения из ImageNet-C: туман, мороз, снег (летащий); расфокусировка, размытие в движении, размытие при приближении.
- Искажения изображений дорожных знаков: запотевание, мороз, дождь; размытие.
- Библиотека ImgAug: туман, мороз, снег (лежащий/летащий), дождь; расфокусировка, размытие в движении, размытие при приближении.

Все запуски классификатора было решено проводить по 5 раз. Агрегация результатов по запускам проводилась с помощью взятия среднего.

Была поставлена задача получить для каждого набора запусков с конкретной аугментацией результаты по времени обучения модели и качеству работы классификатора по метрике accuracy [15]. Известно, что состязательное обучение обычно понижает качество классификации, но повышает устойчивость классификатора. Тем не менее, чем менее значительно такое понижение качества, тем лучше подход к состязательному обучению. Также важно, чтобы использование трансформаций входных примеров на основе выбранных аугментаций не занимало слишком много времени, так как это может сделать невозможным аугментирование «на лету».

В Табл. III, Табл. IV и Табл. V представлены результаты экспериментов с использованием аугментаций из ImageNet-C, аугментаций изображений дорожных знаков и аугментаций из библиотеки ImgAug соответственно.



Рис. 11: Примеры погодных аугментаций с помощью библиотеки Imgaug.

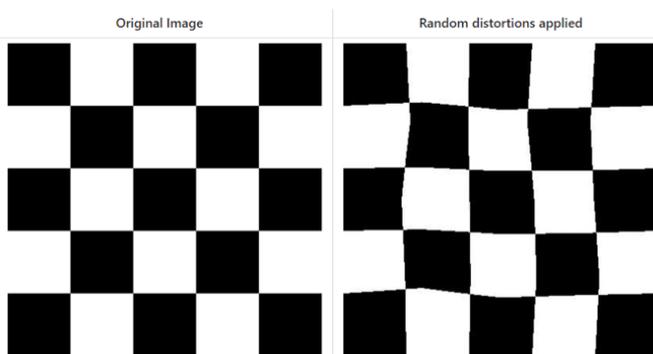


Рис. 12: Пример применения аугментации типа «искажения камеры» из библиотеки Augmentor.

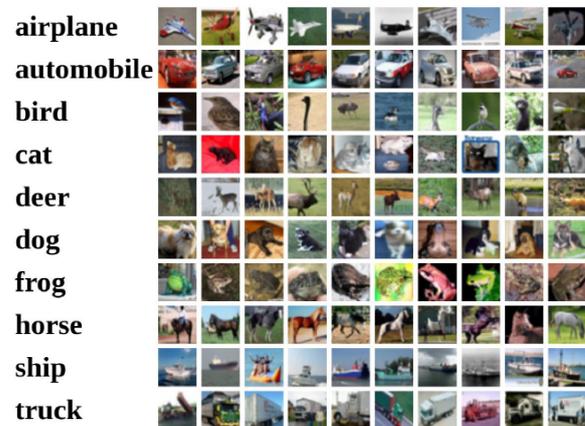


Рис. 13: Примеры изображений с их метками из датасета CIFAR-10.

Таблица III: Сравнение методов на основе искажений из ImageNET-C на предмет ухудшения точности и замедления обучения модели.

Используемый метод аугментации	Точность модели (avg), %	Время 1 эпохи обучения модели (avg), мин	Время 1 эпохи обучения модели (avg), относ. baseline
Без искажений (baseline)	80.61	1.69	1.0
Туман	79.42	2.86	1.69
Мороз	79.33	5.20	3.08
Снег	78.70	31.23	18.48
Расфокусировка	80.55	3.66	2.17
Размытие в движении	77.32	34.55	20.44
Размытие при приближении	79.21	37.13	21.97

Можно видеть, что ухудшение точности классификации для всех методов аугментации по отдельности достаточно малое и не превышает 3.29%. Это может быть следствием «простоты» входного набора данных.

При использовании большинства методов аугментации время обучения заметно не увеличилось, однако для некоторых методов аугментации из статей про рас-

познавание дорожных знаков и ImageNet-C увеличение времени обучения было критически большим (в 18 и более раз). Это может быть вызвано неоптимальностью авторских реализаций этих методов аугментации, а также ошибками в экспериментальном исследовании.

Необходимо заметить, что методы аугментации из библиотеки ImgAug показали себя в среднем более быстрыми, больших отклонений по времени среди них не было.

Таблица IV: Сравнение методов на основе искажений изображений дорожных знаков на предмет ухудшения точности и замедления обучения модели.

Используемый метод аугментации	Точность модели (avg), %	Время 1 эпохи обучения модели (avg), мин	Время 1 эпохи обучения модели (avg), относ. baseline
Без искажений (baseline)	80.61	1.69	1.0
Запотевание	79.61	9.91	5.86
Мороз	79.50	4.13	2.44
Дождь	80.04	9.33	5.52
Размытие	78.43	33.10	19.59

Таблица V: Сравнение методов на основе искажений из библиотеки ImgAug на предмет ухудшения точности и замедления обучения модели.

Используемый метод аугментации	Точность модели (avg), %	Время 1 эпохи обучения модели (avg), мин	Время 1 эпохи обучения модели (avg), относ. baseline
Без искажений (baseline)	80.61	1.69	1.0
Туман	80.34	5.39	3.19
Мороз	78.68	2.74	1.62
Снег лежащий	79.80	2.36	1.40
Снег летящий	79.10	3.56	2.11
Дождь	78.09	5.94	3.51
Расфокусировка	80.01	4.42	2.62
Размытие в движении	78.50	6.46	3.82
Размытие при приближении	79.69	5.94	3.51

V. Заключение

В работе был проведен анализ современных методов естественной аугментации изображений, которые имитируют различные естественные воздействия на окружающую среду. Рассмотрение таких методов оказалось актуальным, так как они способны улучшить устойчивость и точность моделей машинного обучения в реальных условиях эксплуатации.

В ходе работы были изучены различные подходы к естественной аугментации, такие как наложение погодных артефактов, артефактов камеры и замена фона основного объекта на изображении; оценены возможности доступных программных решений, таких как Albumentations, ImgAug и Augmentor; проведено экспериментальное исследование.

Исследование показало, что применение методов естественной аугментации при состязательном обучении может не оказывать значительного влияния на точность классификации, однако может сильно замедлять обучение в случае неоптимальной реализации алгоритмов преобразования изображений.

Основная идея разработки унифицированной библиотеки для всеобъемлющей естественной аугментации представляется перспективной, поскольку она позволит стандартизировать обработку входных данных и упростит интеграцию в существующие системы машинного обучения. Это также способствует более глубокому пониманию влияния различных типов аугментаций на конечное качество моделирования.

В заключение, данная работа выявила полезность разработки унифицированной программной библиотеки для естественной аугментации изображений, вбирающей в себя все лучшее из аналогов и предоставляющей единые требования к предобработке данных для подачи их в пайплайн аугментации.

БИБЛИОГРАФИЯ

[1] Hendrycks D., Dietterich T. G. Benchmarking neural network robustness to common corruptions and perturbations // ArXiv. — 2019. —

URL: <https://api.semanticscholar.org/CorpusID:56657912>, Retrieved: Sep, 2024.

- [2] Robust traffic sign recognition against camera failures / M. Atif, A. Ceccarelli, T. Zoppi et al. // IEEE Open Journal of Intelligent Transportation Systems. — 2022. — Vol. 3. — P. 709–722.
- [3] 3d common corruptions and data augmentation / O. Kar, T. Yeo, A. Atanov, A. Zamir // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2022. — P. 18941–18952.
- [4] Rain rendering for evaluating and improving robustness to bad weather / M. Tremblay, S. S. Halder, R. de Charette, Lalonde J.-F. // International Journal of Computer Vision. — 2021. — 02. — Vol. 129. — P. 1–20.
- [5] Unpaired image-to-image translation using cycle-consistent adversarial networks / J.-Y. Zhu, T. Park, P. Isola, A. A. Efros // ArXiv. — 2020. — URL: <https://doi.org/10.48550/arXiv.1703.10593>, Retrieved: Sep, 2024.
- [6] Let's get dirty: Gan based data augmentation for camera lens soiling detection in autonomous driving / M. Uříčář, G. Sistu, H. Rashed et al. // 2021 IEEE Winter Conference on Applications of Computer Vision (WACV). — 2021. — P. 766–775.
- [7] Kamann C., Rother C. Benchmarking the robustness of semantic segmentation models // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). — 2020. — P. 8825–8835.
- [8] McLaughlin N., Del Rincon J. M., Miller P. Data-augmentation for reducing dataset bias in person re-identification // 2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). — 2015. — P. 1–6.
- [9] Albumentations: fast and flexible image augmentations. — URL: <https://albumentations.ai>. Retrieved: Sep, 2024.
- [10] Pytorch framework. — URL: <https://pytorch.org/>. Retrieved: Sep, 2024.
- [11] Imgaug library documentation. — URL: <https://imgaug.readthedocs.io/en/latest>. Retrieved: Sep, 2024.
- [12] Augmentor documentation. — URL: <https://augmentor.readthedocs.io/en/stable>. Retrieved: Sep, 2024.
- [13] Cifar-10 dataset. — URL: <https://www.cs.toronto.edu/~kriz/cifar.html>. Retrieved: Sep, 2024.
- [14] Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classification with deep convolutional neural networks // Communications of the ACM. — 2017. — Vol. 60. — P. 84–90.
- [15] Multi-class classification metrics. — URL: <https://www.evidentlyai.com/classification-metrics/multi-class-metrics>. Retrieved: Sep, 2024.

A Comparative Study for Natural Image Augmentation Methods

Maksim V. Poryvai

Abstract—This paper deals with natural image augmentation methods, i.e., those whose results are close to the natural environmental effects that machine learning models may encounter in industrial applications: from weather conditions; from the operation or malfunction of environmental perception devices (e.g. cameras); and so on.

Structuring and analyzing natural image augmentation methods can bring significant practical relevance as a guide for researchers of machine learning methods, as well as those developing machine learning models directly for industry.

The currently known methods of natural image augmentation in this paper have been divided into the following groups: those based on adding weather artifacts, camera artifacts or substituting the background for the main object in the image.

The paper also considers existing software libraries for image augmentation: Albumentations, which has integration with PyTorch; ImgAug, which provides a wide range of possible natural augmentations; Augmentor, which provides the ability to add a variety of camera distortion effects to images.

The CIFAR-10 dataset was selected for an experimental study of natural image augmentation techniques; augmentations were applied one at a time with a 50% probability each.

Keywords—natural image augmentation, weather artifacts, camera artifacts, background substitution

REFERENCES

- [1] Hendrycks D., Dietterich T. G. Benchmarking neural network robustness to common corruptions and perturbations // ArXiv. — 2019. — URL: <https://api.semanticscholar.org/CorpusID:56657912>, Retrieved: Sep, 2024.
- [2] Robust traffic sign recognition against camera failures / M. Atif, A. Ceccarelli, T. Zoppi et al. // IEEE Open Journal of Intelligent Transportation Systems. — 2022. — Vol. 3. — P. 709–722.
- [3] 3d common corruptions and data augmentation / O. Kar, T. Yeo, A. Atanov, A. Zamir // 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2022. — P. 18941–18952.
- [4] Rain rendering for evaluating and improving robustness to bad weather / M. Tremblay, S. S. Halder, R. de Charette, Lalonde J.-F. // International Journal of Computer Vision. — 2021. — 02. — Vol. 129. — P. 1–20.
- [5] Unpaired image-to-image translation using cycle-consistent adversarial networks / J.-Y. Zhu, T. Park, P. Isola, A. A. Efros // ArXiv. — 2020. — URL: <https://doi.org/10.48550/arXiv.1703.10593>, Retrieved: Sep, 2024.
- [6] Let's get dirty: Gan based data augmentation for camera lens soiling detection in autonomous driving / M. Uříčář, G. Sistu, H. Rashed et al. // 2021 IEEE Winter Conference on Applications of Computer Vision (WACV). — 2021. — P. 766–775.
- [7] Kamann C., Rother C. Benchmarking the robustness of semantic segmentation models // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). — 2020. — P. 8825–8835.
- [8] McLaughlin N., Del Rincon J. M., Miller P. Data-augmentation for reducing dataset bias in person re-identification // 2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). — 2015. — P. 1–6.
- [9] Albumentations: fast and flexible image augmentations. — URL: <https://albumentations.ai>. Retrieved: Sep, 2024.
- [10] Pytorch framework. — URL: <https://pytorch.org/>. Retrieved: Sep, 2024.
- [11] Imgaug library documentation. — URL: <https://imgaug.readthedocs.io/en/latest>. Retrieved: Sep, 2024.
- [12] Augmentor documentation. — URL: <https://augmentor.readthedocs.io/en/stable>. Retrieved: Sep, 2024.
- [13] Cifar-10 dataset. — URL: <https://www.cs.toronto.edu/~kriz/cifar.html>. Retrieved: Sep, 2024.
- [14] Krizhevsky A., Sutskever I., Hinton G. E. Imagenet classification with deep convolutional neural networks // Communications of the ACM. — 2017. — Vol. 60. — P. 84–90.
- [15] Multi-class classification metrics. — URL: <https://www.evidentlyai.com/classification-metrics/multi-class-metrics>. Retrieved: Sep, 2024.