DOI: https://doi.org/10.30898/1684-1719.2023.2.7
УДК: 621.396.969.3:004.032.26

ИСПОЛЬЗОВАНИЕ ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСШИРЕНИЯ ОБУЧАЮЩИХ (ТЕСТОВЫХ) ВЫБОРОК

М.Л. Белокопытов

Военно-космическая академия имени А.Ф. Можайского
197198, Санкт-Петербург, ул. Ждановская, 13

Статья поступила в редакцию 21 декабря 2022 г.

Аннотация. Статья посвящена вопросам формирования исходного набора данных для обучения нейронных сетей. Представлены основные способы формирования и расширения обучающих (тестовых) выборок, в частности с помощью аугментации, методов математического моделирования, а также генеративно-состязательных нейронных сетей. Рассмотрен принцип работы GAN-сетей, описан алгоритм работы генератора и дискриминатора по методу минимаксной игры, методу Вассерштейна и методу «штрафных градиентов». Произведено цифровое моделирование GAN-сети по минимаксной модели, обученной на наборе данных MNIST. Сделаны выводы относительно возможности применения разработанной генеративно-состязательной нейронной сети с целью расширения банка обучающих (тестовых) выборок.

Ключевые слова: нейронные сети, распознавание, генеративное моделирование, обучающая выборка, тестовая выборка.

Автор для переписки: Белокопытов Марк Львович, hommer1990@mail.ru

Введение

В настоящее время нейронные сети, как один из инструментов решения трудноформализуемых задач, нашли широкое применение на практике при распознавании образов. Это связано с тем, что качество обнаружения и идентификации объектов с помощью нейронных сетей часто значительно превышает качество, обеспечиваемое классическими алгоритмами машинного зрения (такими, как корреляционные фильтры) [1].

Особую важность в процессе работы с нейронными сетями представляет этап их обучения. Именно от него зависит получаемая точность работы алгоритма распознавания. Для качественного обучения нейросети необходимо иметь подготовленный набор исходных данных. Однако зачастую возникает проблема, связанная с трудностью поиска наборов данных, особенно при решении частных задач, таких как распознавание и классификация элементов инфраструктуры военных и военно-промышленных объектов. В этих случаях сбор изображений и формирование из них наборов исходных данных является подготовительным обеспечивающим этапом.

Первичный набор исходных данных принято называть генеральной совокупностью или *dataset*. Из генеральной совокупности в последующем формируются выборки – конечные подмножества элементов генеральной совокупности, изучив которые можно получить представление об особенностях исходного множества.

Вероятностная модель порождения данных предполагает, что выборка из генеральной совокупности формируется случайным образом. Простая выборка является математической моделью серии независимых опытов и, как правило, используется для машинного обучения. При этом для каждого этапа обучения необходим свой набор данных, включающий обучающую выборку, тестовую выборку и проверочную (валидационную) выборку.

Способы формирования обучающих и тестовых выборок зависят от класса задачи, решаемой с помощью машинного обучения:

– для задач распознавания (классификации) данные следует разделять так, чтобы в полученных наборах количественное соотношение объектов разных классов было таким же, как в исходной генеральной совокупности;

– для задач регрессионного анализа необходимо обеспечить одинаковое распределение целевой переменной в полученных наборах, которые будут использоваться для обучения и контроля качества.

В задачах распознавания при построении предсказательных моделей на основе аппарата нейронных сетей *dataset* разбивается на обучающую (*training data*) и тестовую (*test data*) выборки. Обучающая выборка используется собственно для обучения той или иной модели, то есть для построения математических отношений между некоторой переменной-откликом и предикторами, тогда как тестовая выборка служит для получения оценки прогнозных свойств модели на новых данных, которые не были использованы для обучения модели. Как правило, обучающая выборка составляет 75…80% от объема *dataset*, хотя каких-то строгих правил в этом отношении не существует.

Зачастую, подготовка подобного рода набора исходных данных, в частности обучающей выборки, требует наличия и привлечения большого количества сил и материально-технических ресурсов при получении экспериментальных данных в требуемых объемах с априорным знанием характеристик наблюдаемой сцены, что на практике является затруднительным [2]. Кроме того, не всегда представляется возможным получение данных по всей номенклатуре объектов для всех возможных условий их наблюдения и параметров функционирования аппаратуры.

Для решения проблемы получения обучающей выборки достаточного объема, которая будет обладать высокой точностью и устойчивостью к условиям съемки, используют различные методы искусственного расширения *training data*. Наиболее часто встречающийся из этих методов – это синтезирование изображений интересующих объектов – аугментация реальных изображений с помощью вырезания объектов с дальнейшим наложением их на изображение фона со случайным поворотом, случайным зеркальным отражением или

случайным масштабированием в заданных пределах. Этот прием, используемый для создания дополнительных обучающих данных на основе уже имеющихся данных, позволяет увеличить исходную обучающую выборку, которая содержит ограниченное количество изображений объектов заданного класса. Наиболее распространенными вариантами аугментации в соответствии с [3] являются: отражение по горизонтали; случайное кадрирование; изменение цвета.

Следующий приём – использование методов математического моделирования. Он нашел широкое применение в тех случаях, когда не представляется возможным получить экспериментальные данные по всей номенклатуре объектов интереса для всех возможных условий их наблюдения и параметров функционирования датчиковой аппаратуры. Исследования по моделированию изображений были проведены в работах [4, 5], но наиболее подробно они изложены в работе [6]. В данной работе разработанная авторами модельно-ориентированная методика позволяет произвести расчет разноракурсных радиолокационных портретов квазистационарных объектов к которым относятся различные виды наземной, авиационной и морской техники на фоне подстилающей поверхности применительно к заданным параметрам и условиям радиолокационного наблюдения земной поверхности.

Существует еще один эффективный способ преумножения изображений для расширения обучающей выборки – это применение генеративно-состязательной нейронной сети (*GAN*), представляющей собой архитектуру, состоящую из генератора и дискриминатора, которые являются различными свёрточным нейросетями (СНС). СНС-генератор создает случайные новые экземпляры данных, СНС-дискриминатор оценивает их на подлинность, тем самым, принимает решение, относится ли экземпляр данных к набору обучающих данных или нет.

1. Исходные данные и постановка задачи

Генеративное моделирование – это задача машинного обучения, которая заключается в автоматическом обнаружении закономерностей и зависимостей во входных данных, которые можно было бы использовать для генерации на выходе новых примеров, которые могли бы быть непротиворечиво/правдоподобно присутствовать в оригинальном (исходном) наборе данных.

Генеративно-состязательные сети основаны на теоретико-игровом сценарии, в котором генерирующая сеть должна конкурировать с противником. Цель генеративной сеть состоит в том, чтобы генерировать такие образцы данных, которые будут приняты дискриминатором. Его противник, сеть дискриминатора, пытается различить выборки, взятые из обучающих данных, и выборки, взятые из генератора. Цель дискриминатора – определить, является ли изображение подлинным.

Модель генератора принимает случайный вектор фиксированной длины в качестве входных данных и генерирует выборку. Вектор берется случайным образом из гауссовского распределения и используется для начального процесса генерации. После обучения точки в этом многомерном векторном пространстве соответствуют точкам в исходной области, образуя сжатое представление распределения данных. Это векторное пространство называется скрытым пространством или векторным пространством, состоящим из скрытых переменных.

Модель дискриминатора берет пример из области входных данных (действительных или сгенерированных) и предсказывает двоичную метку класса реального или поддельного (сгенерированного) объекта. Реальный пример берется из обучающего набора данных. Сгенерированные примеры выводятся моделью генератора. Таким образом дискриминатор представляет собой стандартную классификационную модель.

Пример архитектуры модели GAN-сети представлен на рисунке 1.

Однако при использовании *GAN*-сетей для генерации изображений существуют определенные трудности с их обучением. Так, при обучении дискриминатора необходимо удерживать значения генератора постоянными и наоборот. То есть каждая сеть должна тренироваться против статичного «противника».

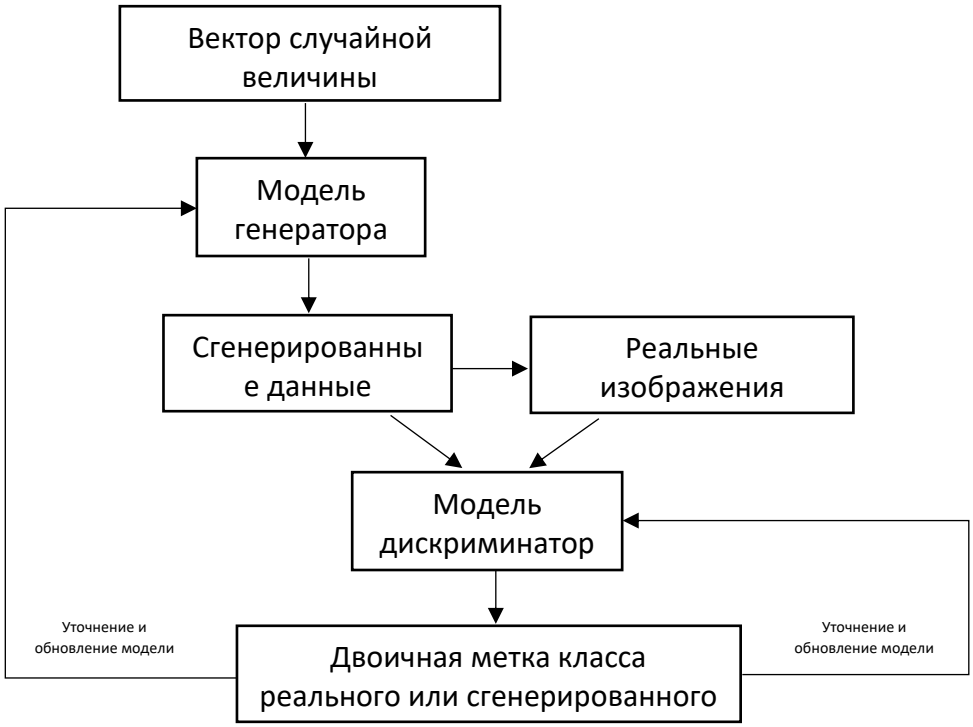


Рис. 1. Архитектура модели *GAN*-сети

Ещё одна трудность связана с неравномерностью обучения, например, когда дискриминатор слишком хорошо обучен, то он будет возвращать значения очень близкие к 0 или 1, тем самым осложняя работу генератора при чтении вектора градиента. Если же генератор хорошо обучен, то он будет использовать неточности дискриминатора, которые будут приводить к ложному срабатыванию [7].

2. Моделирование GAN-сетей для расширения обучающих (тестовых) выборок

Проиллюстрируем процесс обучения *GAN*-сетей. Пусть *D(x)* – выход дискриминатора, который представляет собой вероятность того, что *x* является реальным изображением. *G(z)* – выход генератора (данные, синтезированные генератором), где *z* – шум, поданный на вход генератору. Цель дискриминатора – максимизировать логарифм вероятности реальных данных и логарифм обратной вероятности синтетических данных:

$$\max_D \left\{ \log\left(D(x)\right) + \log\left(1 - D\left(G(z)\right)\right) \right\}. \quad (1)$$

Генератор, в свою очередь, пытается минимизировать логарифм обратной вероятности (полученной дискриминатором) синтетических данных и учится выдавать такие данные, которые бы с малой вероятностью были распознаны дискриминатором как синтезированные:

$$\min_G \left\{ \log\left(1 - D\left(G(z)\right)\right) \right\}. \quad (2)$$

Таким образом, окончательная функция потерь будет минимаксной игрой между двумя классификаторами, которую можно проиллюстрировать следующим образом:

$$\min_G \max_D \left\{ \log\left(D(x)\right) + \log\left(1 - D\left(G(z)\right)\right) \right\}, \quad (3)$$

которая теоретически сходится к дискриминатору, предсказывающему все с вероятностью 0,5.

Однако на практике минимаксная игра часто приводит к тому, что сеть не сходится, поэтому важно тщательно настроить процесс обучения.

Для стабилизации процесса обучения *GAN*-моделей можно использовать метод Вассерштейна. Таким образом, решается оптимизационная задача следующего вида (4):

$$\min_G \max_{D \in \xi} \left\{ D(x) - D\left(G(z)\right) \right\}, \quad (4)$$

где D – класс функций 1-Липшица ($f \in \xi$, если $\forall x, y \in R \exists L : |f(x) - f(y)| \le L|x - y|$).

В этом случае модель дискриминатора обучается в несколько раз чаще генератора и уже не предсказывает вероятность попадания события в конкретный класс (как в задаче бинарной классификации), а «оценивает» событие по реальным меткам класса. Другими словами, выход модели дискриминатора не подвергается активации и поэтому интерпретируется не как вероятность, а как количественная оценка входных данных. Функция потерь здесь определена как расчет среднего прогнозируемого значения по реальным и сгенерированным данным.

Для того чтобы учесть условие, что дискриминатор должен принадлежать к классу функций Липшица, требуется, чтобы веса дискриминатора были небольшими и находились в рамках какого-то интервала, который бы задавался гиперпараметрами [8].

Для еще большей стабилизации обучения генератора можно использовать метод «штрафных градиентов» (*gradient penalty*). Он заключается в дополнительном обучении модели генератора на интерполированных данных между реальными и сгенерированными данными и учитывает это обучение в обновлении весов дискриминатора [9].

В настоящей статье в качестве примера при построении *GAN*-сети использовалась минимаксная модель. В качестве набора исходных данных при обучении *GAN*-сети использовался простейший набор данных *MNIST*, который находится в открытом доступе и содержит 60000 изображений рукописных цифр от 0 до 9 размером 28 на 28 пикселей (рисунок 2).

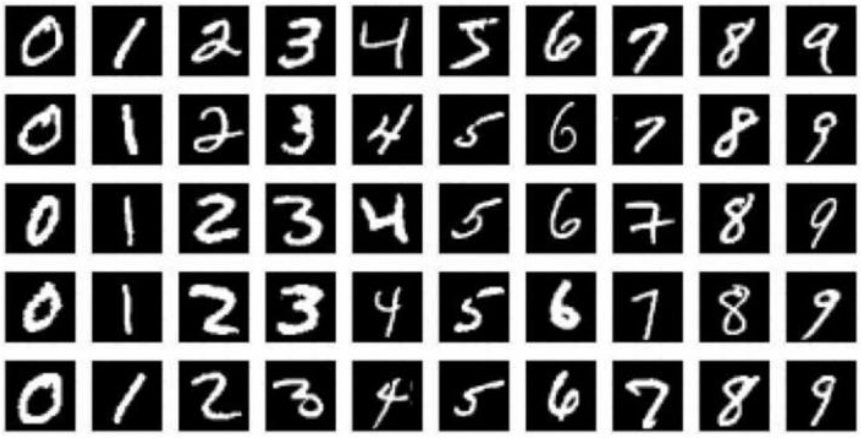
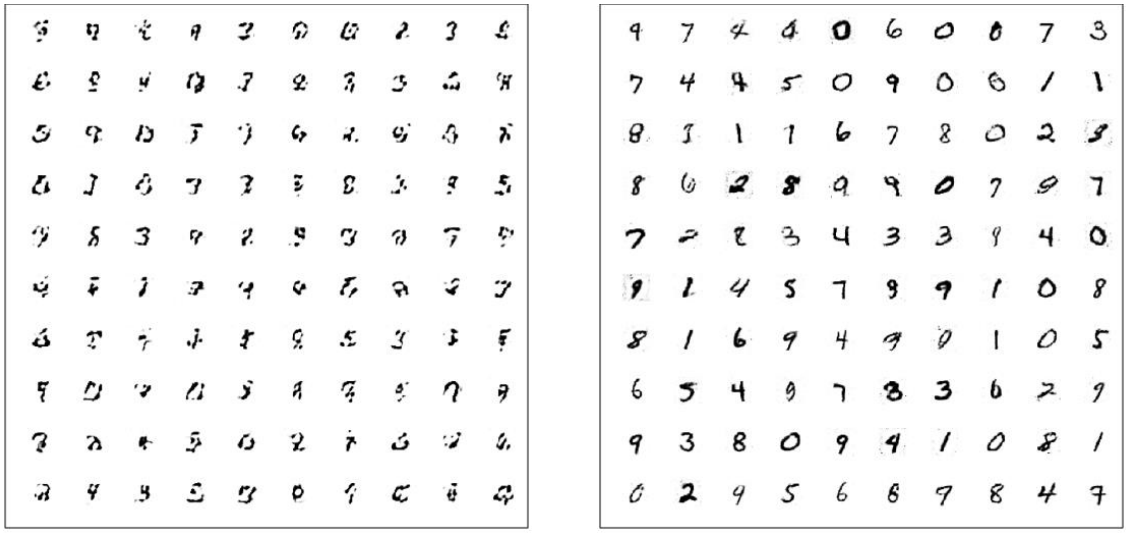


Рис. 2. Пример данных из набора *MNIST*

Программный код написан на высокоуровневым языке программирования общего назначения *Python* с использованием библиотеки *PyTorch* (включая *torchvision*). Визуализация результатов, сгенерированных *GAN*-сетью, была построена с использованием библиотеки *Matplotlib*.

Из-за простоты чисел две архитектуры – дискриминатор и генератор – были построены из полностью связанных слоев. Обучение СНС происходило 200 эпох на оборудовании *Nvidia GeForce GTX* 1080 8 *Gb* и заняло примерно 1 час.

Примеры сгенерированных изображений представлены на рисунках 3а (соответствует 100 эпохам обучения СНС) и 3б (соответствует 200 эпохам обучения СНС).



а б

Рис. 3. Пример изображений, сгенерированных *GAN*-сетью

На рисунке 4 представлена зависимость значения функции потерь от количества эпох обучения.

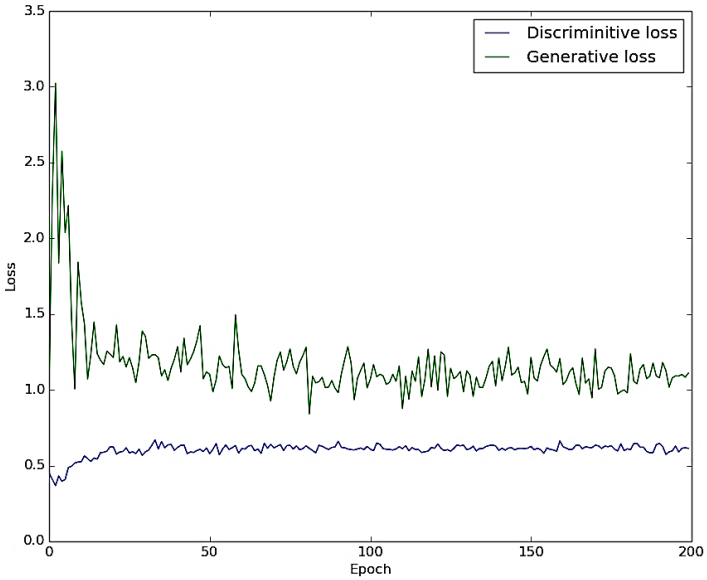


Рис. 4. Зависимость значения функции потерь
от количества эпох обучения

Сгенерированные таким образом изображения в последующем возможно использовать для расширения банка обучающих (тестовых) выборок, необходимых для качественного обучения СНС.

Заключение

Исходя из всего вышеизложенного, можно сделать вывод о том, что подготовка набора исходных данных, включающего в себя обучающую и тестовую выборки, является одним из ключевых этапов при обучении нейронных сетей и решении задач распознавании образов. Качественно подобранная и структурированная информация напрямую связана с точностью работы алгоритма распознавания.

В условиях применения нейросетей для задач обнаружения и идентификации элементов инфраструктуры военных и военно-промышленных объектов задача подготовки обучающей и тестовой выборок приобретает ещё большую актуальность. Прежде всего она обусловлена трудностью получения экспериментальных данных в требуемых объемах с априорным знанием характеристик наблюдаемой сцены.

Для решения данной проблемы в настоящее время используют различные методы искусственного расширения набора исходных данных. Однако, среди

этих методов моделирование изображений с использованием генеративно-состязательных нейросетей, на взгляд автора, является наиболее востребованным и перспективным направлением.

Литература

1. Малыхина Г.Ф., Меркушева А.В. Элементы статистической концепции обучения нейронной сети и прогнозирование ее функционирования. *Научное приборостроение*. 2005. Т.15. №1. С.29-45.
2. Белокопытов М.Л., Шленских Д.А., Морозов С.В., Сирота С.В. Мониторинг объектов морского судоходства по аэрокосмическим данным дистанционного зондирования в СВЧ диапазоне с применением нейросетевых технологий. *Журнал радиоэлектроники* [электронный журнал]. 2022. №4. https://doi.org/10.30898/1684-1719.2022.4.2
3. Окунев С.В. Рассмотрение способов формирования наборов данных для обучения нейронных сетей. *Вестник науки и образования*. 2020. Т.3 №2. С.16-19.
4. Ананьин Э.В., Андрющенко М.С. Методы исследований радиолокационной сигнатуры при разработке малозаметных образцов военной техники. *Вопросы оборонной техники. Серия 16: технические средства противодействия терроризму*. 2015. №87-88. С.102-108.
5. Тертышник В.В., Зиновьев В.В. Методика расчета радиолокационных портретов аэродинамических объектов. *Международная научно-техническая конференция, посвященная 15-летию кафедры радиотехники*. Саратов. 2004. С.11-24.
6. Филиппских Е.Э., Попов А.В., Галкин Ф.А. Модельно-ориентированная методика расчета радиолокационных портретов (сигнатур) металлических объектов. *Журнал радиоэлектроники* [электронный журнал]. 2022. №3. https://doi.org/10.30898/1684-1719.2022.3.5
7. Гарсия Глория Буэно. *Обработка изображений с помощью OpenCV*. Москва, ДМК Пресс. 2015. 387 с.

8. Jonathan Hui. GAN – Wasserstein GAN [web]. *Medium*. Дата обращения: 11.11.2022. URL: https://jonathan-hui.medium.com/gan-wasserstein-gan-wgan-gp-6a1a2aa1b490
9. Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, Aaron Courville. Improved Training of Wasserstein GANs [web]. *NIPS 2017*. Дата обращения: 21.11.2022. URL: https://proceedings.neurips.cc/paper/2017/file/892c3b1c6dccd52936e27cbd0ff683d6-Paper.pdf

Для цитирования:

Белокопытов М.Л. Использование генеративно-состязательных нейронных сетей для расширения обучающих (тестовых) выборок. *Журнал радиоэлектроники* [электронный журнал]. 2023. №2. https://doi.org/10.30898/1684-1719.2023.2.7