

Повышение качества распознавания изображений подбором параметров сверточной нейронной сети

Черемисинова О.Н, студент
Ростовцев В.С., кандидат технических наук, доцент
ЭВМ ФГБОУ ВО «ВятГУ», г. Киров

Повышение качества распознавания образов на основе сверточной нейронной сети (СНС) при отсутствии общепринятой методики определения оптимальных параметров СНС остается актуальной задачей. Вариантами повышения качества распознавания изображений являются интуитивный выбор архитектуры и гиперпараметров СНС, предобработка обучающих примеров и использование аугментации [1,2]. В статье приводятся рекомендации и результаты применения методики выбора параметров СНС для повышения качества распознавания изображений.

Ключевые слова: распознавание изображений, сверточные нейронные сети, аугментация, предобработка, функция активации, кадрирование, регуляризация

Неотъемлемым этапом разработки модели нейронной сети является подбор ее параметров для улучшения результатов ее работы. Единая общепринятая методика повышения качества распознавания изображения с помощью СНС отсутствует, и в большинстве случаев, подбор параметров нейронной сети определяется опытом и интуицией разработчика.

В данной статье представлен один из способов повышения точности распознавания изображений путем выбора параметров сверточной нейронной сети.

Авторами была разработана сверточная нейронная сеть, выполняющая задачу распознавания изображений, представленных в формате *.jpeg. В качестве языка программирования был выбран язык Python с использованием библиотеки Keras.

Разработанная модель нейронной сети представлена на рис.1.

Был выбран размер изображений обучающей выборки 210x210, размер пакета 32, количество эпох 100, оптимизатор RMSprop, скорость обучения 0,001, в качестве функции ошибки использована кросс-энтропия.

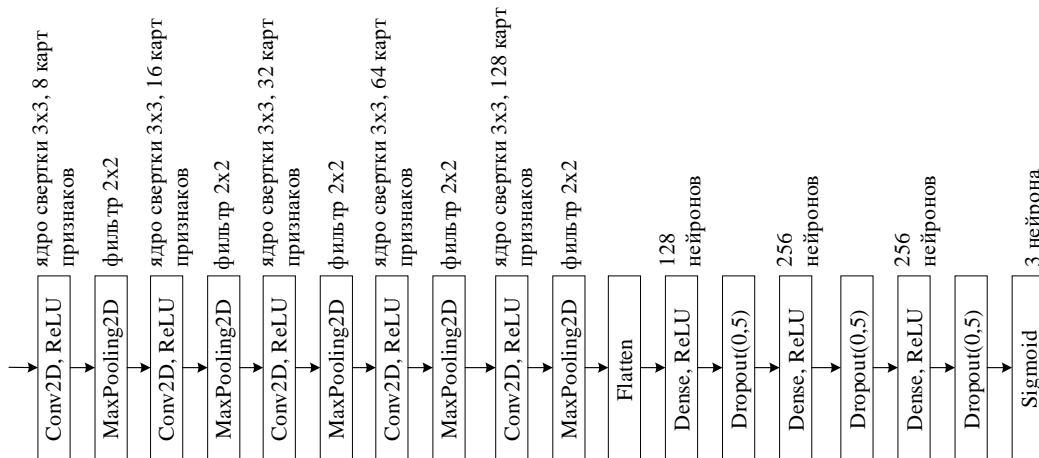


Рис.1. Архитектура сверточной нейронной сети

Точность определения класса данной моделью сверточной нейронной сети без соответствующего подбора параметров составила 88%, которую возможно повысить за счет настройки ряда параметров нейронной сети.

Сверточная нейронная сеть содержит большое число гиперпараметров, которые необходимо выбрать в процессе построения нейросети и от которых будет значительно зависеть результат обучения, например, количество слоев как сверточных, так и полносвязных, их размер, параметры фильтра сверточных слоев и другие.

К основным параметрам, влияющим на качество распознавания изображений, следует отнести: архитектуру сверточной нейронной сети, тип функции активации, параметр скорости обучения, количество эпох обучения, способ формирования обучающей

выборки с учетом нормализации данных и аугментации изображений. Была выполнена последовательная оптимизация перечисленных параметров. Дополнительно была добавлена локализация характерных признаков распознаваемых изображений.

Методом случайного поиска была произведена оптимизация гиперпараметров (количество сверточных слоев, количество каналов и размер ядра в них, количество слоев в полносвязной сети, скорость обучения, размер партии и количество эпох, размер входного изображения, использование регуляризации, функция активации). Случайный поиск позволяет существенно сократить время подбора параметров, заменяя их полный перебор на выборку нескольких значений случайным образом.

Другой возможный и перспективный метод - эволюционная оптимизация [3], основанная на принци-

пе работы эволюционных алгоритмов и обеспечивающая поиск оптимальных гиперпараметров нейронной сети.

Данный алгоритм прост в реализации, но очевидны его недостатки: при большом времени обучения сверточной нейронной сети на процесс подбора гиперпараметров потребуется много времени и ресурсов, а также необходим выбор операторов скрещивания, отбора, мутации и т.д. При разработке сверточной нейронной сети эволюционная оптимизация не использовалась из-за больших затрат времени обучения нейронной сети (до 1 часа). При больших значениях размера популяции эволюционных алгоритмов (более 100) потребуется их реализация с помощью программ и библиотек параллельного программирования.

Автоматический подбор гиперпараметров с помощью байесовской оптимизации [4] требует больших временных и ресурсных затрат. Кроме этого использование только байесовского подхода не даст такого значительного повышения точности распознавания, как использование комплексной оптимизации сверточной нейронной сети, рассматриваемое в данном дипломном проекте.

Была изменена выбранная функция активации с ReLU на ELU [5]. В использовавшейся функции активации ReLU при входном значении, меньшем нуля, выходное значение и производная также будут равны нулю, что приведет к отключению нейрона, то есть в дальнейшем нейрон больше не будет обновляться. ELU позволяет решить проблему «паралича» сети, возникшую в случае ReLU, а также нейроны получают возможность вырабатывать отрицательные значения. В результате ошибка на валидационных данных снижается и максимально приближается к ошибке на обучающих (тренировочных) данных.

Для заданных классов распознаваемых изображений лучшие результаты распознавания изображений были получены при использовании пяти сверточных слоев с распределением каналов по слоям 8-16-32-64-128 соответственно и двух слоев в полносвязной сети. Параметр скорости обучения, размер партии и размер входного изображения остались неизменными.

Также для каждого слоя была добавлена регуляризация Dropout [6], заключающаяся в отключении нейронов с заданной вероятностью. Исключенные нейроны не вносят вклад в процесс обучения, и их веса на данной эпохе не меняются.

В процессе обучения часто встречается такое явление как эффект переобучения нейронной сети. Он заключается в адаптации к тренировочным примерам, что выражается низкой ошибкой распознавания обучающих образов и ростом ошибки на валидационных данных. Причиной переобучения может быть избыточный размер нейронной сети, недостаточное количество обучающих примеров или чрезмерное количество эпох обучения. Кроме настройки этих параметров избежать эффекта переобучения сети помогает использование регуляризации, которая позволяет ограничивать веса нейронов.

Наблюдая за ходом обучения сверточной нейронной сети, было обнаружено, что примерно на 50 эпо-

хе ошибка на валидационной выборке максимально приближена к ошибке на тренировочных данных, и в дальнейшем не происходит ее существенного изменения. Сверточная нейронная сеть, обученная в течение 50 эпох, показала наилучшие результаты распознавания в сравнении с использованием 40 и 60 эпох.

Всю информацию о задаче нейронная сеть получает из обучающей выборки. Поэтому качество обучения нейронной сети непосредственно зависит от того, насколько репрезентативен тренировочный набор. Репрезентативность исходных данных включает в себя три аспекта: количество тренировочных примеров должно быть достаточно для обучения, разнообразие – охватываются все возможные комбинации входных и выходных данных, равномерность представления классов [7].

Данные на вход нейронной сети подавались последовательно по классам: сначала все изображения первого класса, потом второго и т.д. При таком подходе возможна адаптация нейронной сети к порядку их подачи. После подачи изображений со случайным чередованием по классам точность распознавания увеличилась на 3%. Кроме этого, было изменено соотношение обучающей и валидационной выборок на пропорцию 9:1.

Часто подбор большого количества примеров сложен, поэтому для решения проблемы нехватки данных можно создать новые примеры на основе имеющихся. Один из способов их искусственной генерации - аугментация.

Аугментация (augmentation, «раздутие») позволяет увеличить тренировочный набор данных и повысить точность распознавания за счет получения новых изображений с помощью случайного сдвига, масштабирования и растягивания [1,2].

Значительное повышение качества распознавания с помощью данного метода возможно только при подборе оптимального для данной задачи количества аугментированных изображений и применяемых к ним преобразованиям. Кроме этого, целесообразно будет применять к изображениям только те возможные искажения, которые могут встретиться в реальной жизни.

Использование для обучения исходных необработанных данных может негативно сказываться на результатах работы нейронной сети. Поэтому перед подачей изображений на обучение необходимо произвести их предобработку. Чаще всего используется нормализация данных, которая выравнивает вклад отдельных признаков, приводя значения, представленные в различных единицах измерения и разбросанные в большом интервале, к значениям из промежутка [0; 1], [-0,5; 0,5] или [-1; 1]. Экспериментально установлено, что для заданных примеров распознавания изображений наилучшие результаты показало использование нормализации [0; 1].

В задачах классификации нейронная сеть работает с одним объектом, для которого и определяет класс. Поэтому возможно добиться повышения точности распознавания, если и данные для обучения, и распознаваемое изображение будут содержать только один объект.

Для решения поставленной задачи был разработан программный модуль, основанный на библиотеке TensorFlow Object Detection API (https://github.com/tensorflow/models/tree/master/research/object_detection), которая способна распознать на изображении объект и определить область ее нахождения. К примерам обучающей и тестовой выборок применялся поиск объекта распознавания и удаление лишних областей изображения с помощью разработанного модуля.

Применение кадрирования исключительно к изображениям обучающей и валидационной выборок не смогло повысить точность определения класса. Но использование кадрирования также к примерам тестовой выборки позволило повысить точность распознавания изображений на 5%.

В табл. 1 представлены результаты обучения нейронной сети при отсутствии аугментации и при различном количестве добавления аугментированных изображений.

Данные табл. 1 свидетельствуют, что увеличение размера выборки за счет аугментированных изображений способно повысить точность распознавания изображений.

В табл. 2 представлены результаты обучения нейронной сети при использовании различных типов аугментации.

Для данной задачи хорошие результаты показывает использование в аугментации отражения, сдвига и масштабирования, а добавление параметра поворота изображения не увеличило точность распознавания.

После последовательного применения всех вышеперечисленных доработок (оптимизация гиперпараметров сверточной нейронной сети и параметров обучения, изменение способа формирования обучающей выборки, настройка параметров аугментации изображений, добавление локализации характерных признаков распознаваемых изображений) разработанная сверточная нейронная сеть прошла экспериментальную апробацию.

В ходе нее была определена доля правильно распознаваемых изображений. Тестирование проводилось на изображениях, не участвовавших в процессе обучения.

Результаты экспериментальной апробации работы сверточной нейронной сети представлены в табл. 3, а процентное соотношение верно и неверно распознанных изображений – в табл. 4.

Таблица 1. Результаты обучения

№	Аугментация	Ошибка на тренировочных данных	Ошибка на валидационных данных	Процент правильно распознанных тестовых изображений, %
1	Без аугментации	1,1921e-07	1,1921e-07	33
2	Аугментация с увеличением выборки в 2 раза	0,3419	1,1921e-07	87
3	Аугментация с увеличением выборки в 3 раза	0,3105	1,1921e-07	91

Таблица 2. Результаты обучения

№	Тип аугментации	Ошибка на тренировочных данных	Ошибка на валидационных данных	Процент правильно распознанных тестовых изображений, %
1	"Отражение и сдвиг"	0,0939	1,1921e-07	88
2	"Отражение и сдвиг и масштабирование"	0,2012	1,1921e-07	91
3	"Отражение и сдвиг и масштабирование и поворот"	0,3105	1,1921e-07	91

Таблица 3. Результаты распознавания изображений

Номер класса	общее количество изображений	Предсказанный класс		
		№1, изображений	№2, изображений	№3, изображений
№1	30	29	1	0
№2	30	0	29	1
№3	30	0	1	29

Таблица 4. Результаты распознавания изображений (в процентах)

Номер класса	общее количество изображений	Результат	
		верно распознано, %	неверно распознано, %
№1	30	97	3
№2	30	97	3
№3	30	97	3
Всего	90	97	3

Из представленных выше таблиц 3 и 4 видно, что по сравнению с изначальной моделью точность распознавания изображения на полученной модели нейронной сети увеличилась на 9%: с 88% до 97% за счет подбора параметров сверточной нейронной сети.

Такая точность является высокой для нейронной сети при наличии обучающей выборки небольшого размера (менее 1000 изображений каждого класса).

Для данной задачи с заданным исходным набором изображений (около 1000 изображений каждого класса), лучшие результаты показало использование нейронной сети с 5 сверточными и 2 полносвязными слоями, при этом сверточные слои имеют количество каналов 8-16-32-64-128 соответственно и 128 нейронов в скрытом слое полносвязной сети.

Использование функции активации ELU [4] ускорило процесс обучения и снизило ошибку распознавания. Сократив количество эпох обучения до 50 и добавив регуляризацию Dropout, удалось не только повысить точность распознавания изображений, но и избежать эффекта переобучения нейронной сети.

Основной вклад в повышение качества распознавания изображений внесла аугментация, включающая в себя увеличение объема обучающей выборки

за счет автоматически вводимых искажений (масштабирование, сдвиг, отражение, поворот); кадрирование изображений как обучающей и валидационной, так и тестовой выборок улучшило результат распознавания за счет использования только полезной информации и удаления избыточной информации из входных данных.

Трудностью для повышения точности распознавания является поиск и выбор параметров и методов, способных улучшить качество распознавания сверточной нейронной сетью. Затратным по времени этапом оптимизации оказался подбор гиперпараметров сверточной нейронной сети, на который ушло 1/3 времени оптимизации. Для локализации характерных признаков потребовался поиск подходящего решения, способного определить распознаваемый образ и его координаты на изображении, и написание программного модуля кадрирования.

Результаты проведенных экспериментов СНС, выбранная методика подбора параметров создает предпосылки для разработки сервисной программы определения оптимальных параметров сверточной нейронной сети на основе эволюционных алгоритмов с использованием библиотек параллельного программирования.

Литература:

1. Глубокое обучение для новичков: тонкая настройка нейронной сети. URL: <https://habr.com/ru/company/wunderfund/blog/315476/>.
2. Аугментация (augmentation, "раздутие") данных для обучения нейронной сети на примере печатных символов. URL: <https://itnan.ru/post.php?c=1&p=264677>.
3. Краткий курс машинного обучения или как создать нейронную сеть для решения скоринг задачи. URL: <https://habr.com/ru/post/340792/>.
4. Xcessiv. URL: <https://xcessiv.readthedocs.io/en/stable/>.
5. Activation Functions in Deep Learning (Sigmoid, ReLU, LReLU, PReLU, RReLU, ELU, Softmax). URL: <http://laid.delanover.com/activation-functions-in-deep-learning-sigmoid-relu-lrelu-prelu-rrelu-elu-softmax/>.
6. Dropout — метод решения проблемы переобучения в нейронных сетях. URL: <https://habr.com/ru/company/wunderfund/blog/330814/>.
7. Формирование обучающей выборки при использовании искусственных нейронных сетей в задачах поиска ошибок баз данных. URL: <https://pandia.ru/text/78/203/3396.php>.