

Обучение нейронных сетей классификации объектов на основе изображений, полученных с использованием компьютерного моделирования

Е. С. Емельянов, email: cap_emela@mail.ru
А. С. Дикарев, email: antonyladur@yandex.ru

Военный учебно-научный центр Военно-воздушных сил «Военно-воздушная академия имени профессора Н.Е. Жуковского и Ю.А. Гагарина»

***Аннотация.** Рассмотрен подход к обучению нейронных сетей классификации объектов с использованием в качестве «обучающих» изображений результатов компьютерного моделирования на основе физически корректных алгоритмов (Physically Based Rendering – PBR) рассеяния электромагнитных волн видимого диапазона. С использованием PBR рассчитаны изображения тестового объекта в виде легкового автомобиля в различных условиях освещенности. Полученные изображения использованы для обучения классификационных слоев нейронной сети «SqueezeNet». Результаты классификации объекта на реальных фотографиях автомобиля подтвердили эффективность данного подхода.*

***Ключевые слова:** нейронные сети, классификация объектов, физически корректный рендеринг.*

Введение

Одним из проблемных вопросов при обучении нейронных сетей для классификации объектов на изображениях (видео) является отсутствие достаточного количества обучающих данных (маркированных изображений). Наиболее качественными данными для обучения являются фотографии объектов, полученные в реальной фоновой обстановке. Однако, использование реальных изображений сопряжено с некоторыми особенностями.

Это обусловлено тем, что в видимом диапазоне длин волн интенсивность и цветовая гамма материалов поверхности объекта, регистрируемые фотокамерой, существенным образом зависят от фоновой обстановки – состояния атмосферы (облачность, влажность, температура) и времени суток (положения солнца). Фон, на котором наблюдается объект (песок, трава, снег) и др., существенным образом влияет на вероятность его обнаружения, что особенно актуально для

объектов, имеющих специальную маскирующую окраску. В этой связи получение достаточного количества изображений объекта в различных фоно-целевых условиях требует существенных временных затрат. Причем некоторые варианты фоно-целевой обстановки предполагают наличие соответствующих погодных условий или могут быть реализованы только в определенной местности. Наличие в обучающей выборке изображений объекта в наибольшем количестве вариантов фоно-целевой обстановки является предпочтительным. Критичным данное обстоятельство может быть при создании охранных систем или систем военного назначения.

Кроме того, необходимо учитывать то обстоятельство, что в ряде случаев отсутствует сам объект, который необходимо классифицировать на изображении. Такая ситуация чаще всего встречается в системах военного назначения при необходимости обнаруживать и классифицировать военную технику противника, которая недоступна.

1. Основная часть

Одним из подходов, позволяющим преодолеть указанные выше трудности, является использование современных методов компьютерного моделирования изображений на основе физически корректных алгоритмов расчета отражения и распространения электромагнитных волн видимого диапазона.

Такой подход нашел широкое применение в современных методах компьютерной графики. Наиболее часто его называют Physically Based Rendering (PBR) – физически корректный рендеринг. Под PBR подразумевается набор техник визуализации, в основе которых лежат математические модели, основанные на физической теории распространения света. В процессе PBR удастся получить изображения с физической правдоподобностью, что и требуется для качественного обучения нейронной сети. Для реализации PBR необходима 3D модель объекта и оптические характеристики материалов его поверхности. При отсутствии 3D модели она может быть создана по чертежам или эскизам объекта. В программных продуктах, в которых реализованы алгоритмы PBR оптические свойства большинства существующих материалов уже определены. При необходимости создания новых материалов с особенными оптическими свойствами существует множество гибких настроек.

Распространенным способом получения изображений объекта в различных условиях освещенности и на различных фонах является использование HDRi (High Dynamic Range Image) карты. HDRi карта – это панорамная фотография, которая охватывает все углы наблюдения с одной точки и содержит большое количество данных (32 бита в пикселе

на каждый цветовой канал). Применяя HDRI карту возможно получить качественное освещение деталей объекта, что может быть несущественно в обычных условиях, однако важно при обучении нейронных сетей, к которым предъявляются повышенные требования надежности классификации объектов. Карты HDRI могут быть получены для различных фоно-целевых условий и являться своего рода набором типовых условий наблюдения, которые могут быть выбраны как наиболее вероятные в той или иной ситуации.

Преимущества такого подхода состоит в том, что полученные наборы изображений являются априори промаркированными, т.к. рендеринг выполняется для известного объекта. При наличии набора HDRI карт для различного времени суток, погодных условий, времени года, окружающего фона (лес, городская застройка и др.) возможно получение изображений объекта в практически любой фоно-целевой обстановке. Существуют как коммерческие, так и бесплатные ресурсы с готовыми HDRI картами. Пример HDRI карты представлен на рис 1.



Рис. 1. HDRI карта (фоно-целевая обстановка «Город»)

Таким образом с использованием современных подходов компьютерной графики возможно получить практически неограниченный набор изображений требуемых объектов в требуемой фоно-целевой обстановке для обучения нейронной сети.

2. Результаты моделирования

Для проверки данного подхода к обучению нейронных сетей проведен численный эксперимент. На первом этапе был рассчитан набор

изображений для обучения нейронной сети. В качестве объекта выбран легковой автомобиль (рис. 2).



Рис. 2. Моделирование изображения (рендер) автомобиля в фоно-целевой обстановке «Город»

Прочет изображений проведен с помощью системы рендеринга «Corona Renderer». В процессе рендеринга удается получить результаты с физической правдоподобностью и фотографическим качеством. Время расчета одного изображения с разрешением 1920 на 1080 пикселей составило от 3-х до 5-ти минут (процессор – Corei5, ОЗУ – 8Гб, видеокарта – Nvidia 1050Ti). Как видно на рис. 2 автомобиль выглядит реалистично присутствуют тени, в зеркальных поверхностях отражаются объекты HDRI карты (брусчатка, фонарный столб, небосвод), освещение объекта «встроенного» в HDRI карту соответствует ожидаемому. Полученные таким образом изображения далее будем называть «рендер-изображения».

Всего было получено 50 рендер-изображений автомобиля, при различных ракурсах наблюдения и пространственных положениях на HDRI карте. Пример полученных рендер-изображений представлен на рис. 3.

Рендер-изображения использовались при обучении слоев классификации нейронной сети «SqueezeNet» [1], т.е. был реализован так называемый процесс передачи обучения. При переобучении SqueezeNet, для классификации автомобиля, были заменены два последних сверточных слоя и итоговый слой классификации сети [2]. В

качестве альтернативных объектов классификации выбраны представленные в [2] объекты: «фонарик карманный», «кубик Рубика», «кепка», «колода карт», «маркер».



Рис. 3. Примеры рендер-изображений в обучающей выборке

Для увеличения количества изображений к обучающей выборке были добавлены рандомизированные путем вращения, отражения в вертикальной плоскости и масштабирования рендер-изображений. Для контроля качества обучения (валидации) использовано 30 % изображений, исключенных из обучающей выборки.

Результаты обучения представлены на рис. 4. Сеть обучилась за 80 эпох, точность на обучающей выборке составила 93%, на валидационной 98%.

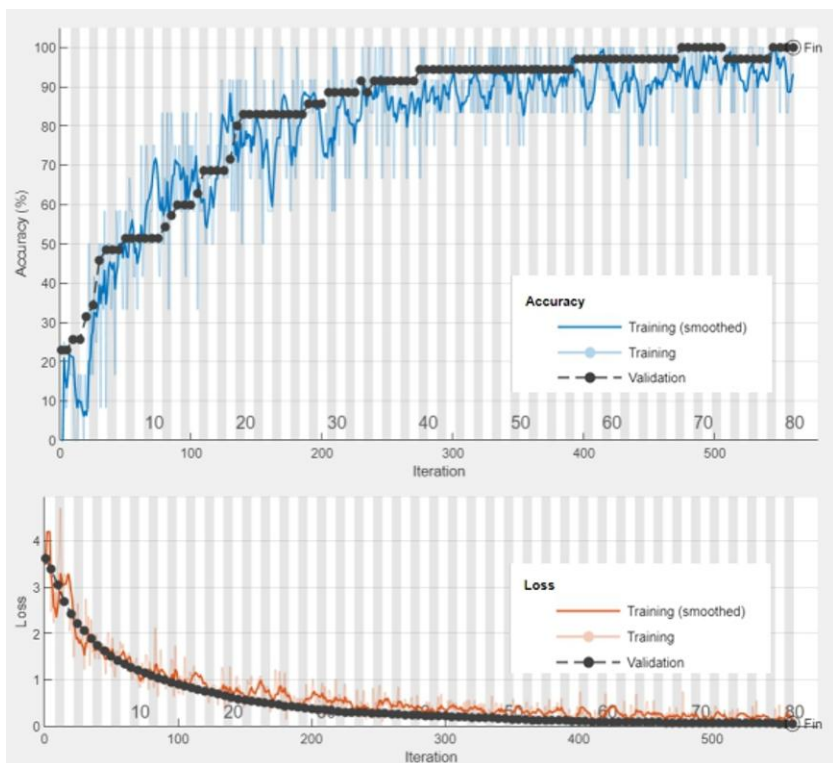


Рис. 4. Метрики обучения нейронной сети SqueezeNet.

Тестирование качества обученной сети проведено на реальных фотоизображениях легкового автомобиля. Результаты классификации приведены на рис. 5. Вероятность классификации на изображении автомобиля составила более 96 %, ложных классификаций нет.

Заключение

Таким образом, современные алгоритмы компьютерной графики на основе физически корректных алгоритмов могут быть использованы для получения набора изображений объекта в различной фоно-целевой обстановке и обучения на их основе нейронных сетей. Такой подход позволяет сократить временные затраты на получение требуемого для обучения набора изображений.



Рис. 5. – Результаты тестирования обученной сети SqueezeNet.

Список литературы

1. Iandola, Forrest N. "SqueezeNet." [Электронный ресурс] – Режим доступа : <https://github.com/forresti/SqueezeNet>
2. MathWorks Help Center [Электронный ресурс] – Режим доступа : <https://www.mathworks.com/help/deeplearning/ug/transfer-learning-with-deep-network-designer.html>