

Использование свёрточных нейронных сетей для классификации видов искусственного освещения

М.И. Сорокин¹, Д.Д. Жданов¹

vergotten@gmail.com|ddzhdanov@mail.ru

¹Университет ИТМО, Санкт-Петербург, Россия.

Актуальность данной темы обусловлена бурным развитием систем виртуальной, смешанной и дополненной реальности. Проблема заключается в формировании естественных условий освещения объектов виртуального мира в реальном пространстве. Реально наблюдаемая картина всегда содержит источники освещения, и когда в реальный мир добавляется виртуальный объект, он должен быть освещен естественным образом, т.е. со стороны источников света и отбрасывать тень в противоположную сторону, другими словами соответствовать освещению реального мира. Задача заключается в анализе и классификации источников освещения для дальнейшего построения соответствующих оптических параметров среды с использованием технологии свёрточных нейронных сетей, которые себя хорошо зарекомендовали в области классификации. В данной работе предложен подход использования свёрточных нейронных сетей для классификации типа освещения с использованием библиотеки машинного обучения Keras.

Ключевые слова: классификация, освещение, свёрточные нейронные сети.

Using convolutional neural networks to classify types of artificial lighting

M.I. Sorokin¹, D.D. Zhdanov¹

vergotten@gmail.com|ddzhdanov@mail.ru

ITMO University, St. Petersburg, Russia.

The relevance of this topic is due to the rapid development of virtual, mixed and augmented reality systems. The problem lies in the formation of natural conditions for lighting objects of the virtual world in real space. The actual picture, that is observed always contains lighting sources, and when a virtual object is added to the real world, it should be illuminated in a natural way, i.e. from the side of light sources and cast the shadow in the opposite direction, in other words, correspond to the illumination of the real world. The task is to analyze and classify the light sources for further construction of the corresponding optical parameters of the medium using the technology of convolutional neural networks that have already proved themselves in the field of classification. In this paper, an approach is proposed for the use of convolutional neural networks to classify the type of illumination using the Keras machine learning library.

Keywords: classification, illumination, convolutional neural networks.

1. Введение

Главное отличие систем дополненной реальности от систем виртуальной – это то, что в системах дополненной реальности слои изображения не просто симулируются, а накладываются на реальный мир, и, следовательно, основная проблема адаптации синтезированных объектов – освещение.

Освещение играет очень важную роль в представлении систем дополненной реальности, ведь неправильное освещение негативно влияет на восприятие, и неправильно освещенный объект может ввести наблюдателя в заблуждение, например, из-за отсутствия теней, ориентированных естественным образом.

Проблема заключается в дискомфорте восприятия синтезированных изображений виртуальной реальности с реальными объектами. Человеческий мозг находит несоответствие с тем, что он видит, и тем, что он ожидает увидеть. Например, некорректно воспринимаются тени от объектов: они излишне контрастны (без полутонов) или имеют неверную ориентацию. В работе [1] предложен метод вставки синтезированных изображений в уже существующие фотографии путем не только анализа освещения, но и геометрии помещения.

Как известно, освещение бывает двух видов: естественное и искусственное. Искусственное же в свою очередь можно разделить по его расположению на следующие виды: боковое, потолочное, напольное и

комбинированное. Каждый вид имеет свои оптические параметры и свойства, такие как угол падения и интенсивность, что является ключевым моментом в образовании теней. Поэтому очень важно правильно определить тип освещения.

Актуальность анализа освещения определяется бурным развитием систем виртуальной, дополненной и смешанной реальности. В данной работе предлагается подход классификации видов искусственного освещения с использованием свёрточных нейронных сетей.

Стоит заметить, что впервые свёрточная нейронная сеть была представлена Алексом Крижевски [2] на соревнованиях по распознаванию изображений ImageNet в 2012 году, где с огромным отрывом опередила своих конкурентов. В дальнейшем, свёрточная нейронная сеть уже использовалась повсеместно, лишь меняя свою архитектуру для достижения более лучших результатов.

2. Состояние проблемы

На данный момент свёрточные нейронные сети активно используются для различных задач, связанных непосредственно с анализом и обработкой изображений, будь то классификация или распознавание каких-либо отдельных участков. К основным преимуществам свёрточных нейронных сетей (CNN) следует отнести удобное распараллеливание вычислений, устойчивость к сдвигу изображения и обучению при помощи

классическому методу обратного распространения ошибки. Из недостатков - большое количество настраиваемых параметров, т.е. для решения какой-либо задачи следует попробовать различные слои и параметры для выбора лучшего решения, к таким параметрам относятся: размерность ядра свёртки, степень уменьшения размерности, использование слоёв субдискретизации, выбор функции активации и т.д. Более подробная информация представлена в официальной документации [3].

Естественно, что свёрточная нейронная сеть хорошо подходит для сегментации изображений. Например, в работе [4] используется свёрточная нейронная сеть для сегментации биомедицинских изображений.

В работе [5] используются свёрточные нейронные сети для оценки состояния окружения на открытом воздухе. Обучаясь на панорамных изображениях и анализируя такие параметры, как природные условия, расположение и яркость солнца и неба, сеть строит предсказание, как должна располагаться тень объекта при заданных условиях.

Аналогичным образом в работе [6] также идет анализ объектов на открытом воздухе. Здесь сеть не только анализирует положение неба, но и делит изображение пополам и анализирует нижнюю часть на предмет теней, анализируя которые, обучается и строит свои собственные предсказания.

В работе [7] происходит анализ небольшого участка помещения, по ко которому может быть корректно сформировано освещение всей видимой части сцены. В данном алгоритме сеть на вход получает изображение, по которому строится маска освещения, и на выходе получается панорама изображения с прогнозируемым освещением.

3. Архитектура CNN

В качестве архитектуры свёрточной нейронной сети в данной работе было решено взять архитектуру VGG 16 NET, которая успешно использовалась в работах [8,9], состоящая из трех слоев свертки размерности 3x3 и размером ядер подвыборки 2x2. Эти слои необходимы для составления карт признаков. После свёрточных слоев следуют полносвязные слои, которые используются для классификации. Далее идет слой исключения (dropout) с параметром 0.5, который необходим для предотвращения обучения, путем исключения из работы определенного процента нейронов. Так как в данной работе классифицируются изображения на два типа освещения, потолочное и боковое (wall, ceiling), то используется "бинарная функция ошибки" (binary crossentropy), оптимизатор - "адам", выходная функция активации "sigmoid" - для плавного изменения значений от 0 до 1, а в качестве метрики обучения - "аккуратность" (accuracy).

В качестве функции активации между слоями выбрана функция "ReLU", которая хорошо себя зарекомендовала в работе [10]. Функцию ReLU можно представить в виде формулы:

$$f(x) = \max(0, x),$$

которая реализует простой пороговый переход в нуле.

Слои дискретизации работают следующим образом: на его вход подается некий массив с информацией, и нужно процедить эту информацию, т.е. уменьшить размерность. На рисунке 1 изображен пример дискретизации.

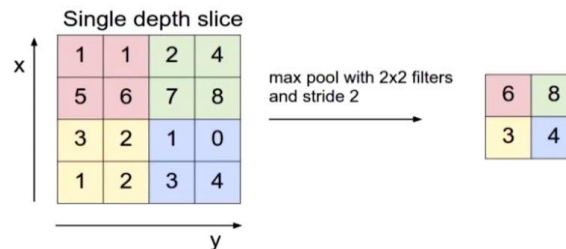


Рис. 1. Операция субдискретизации.

Слои дискретизации необходимы, т.к. использование этих слоёв имеет свои плюсы. Во-первых, уменьшается размерность, а чем меньше размерность, тем больше скорость обучения. Во-вторых, побеждают наиболее активные нейроны. И в-третьих, вырабатывается инвариантность к небольшим сдвигам.

4. Обучение

Так как обучение происходит "с учителем", необходимо для начала сформировать набор изображений каждого вида освещения, в данном случае нейронная сеть обучена определять два вида освещения - потолочное (ceiling) и боковое (wall). Всего набор данных состоит из 1560 изображений каждого класса размерности 200 на 300 пикселей, из которых 15% выделены на тестовый набор изображений и 15% на проверочный набор. Итого, каждый класс состоит из 1092 тренировочных изображений и по 234 тестовых и проверочных. Структура каталогов представлена на рисунке 2, где "train" - набор данных для тренировки, "validation" - для проверки, и "test", соответственно набор данных для тестирования. При составлении набора данных, изображения были зеркально отражены и немного наклонены в разные стороны, что необходимо для увеличения размерности набора данных и выработки инвариантности к наклону.

```
lighting_classification_cnn/
---train/
|---ceiling/
|---wall/
---validation/
|---ceiling/
|---wall/
---test/
|---ceiling/
|---wall/
```

Рис. 2. Структура каталогов наборов данных. Изображения из набора данных представлены на рисунке 3.



Рис. 3. Набор данных потолочного освещения

Обучение нейронной сети происходило при помощи библиотеки машинного обучения "keras", в качестве

бэкенда использовался "tensorflow". При обучении использовался метод "model.fit_generator", которому передавалось на вход два параметра "train_generator" - набор данных с изображениями для обучения и "val_generator" - набор данных для проверки. Обучение выполнялось в течении 30 эпох. На рисунке 4 изображена история обучения, где "acc" - точность на обучающем наборе данных, а "val_acc" - точность на проверочном наборе данных.

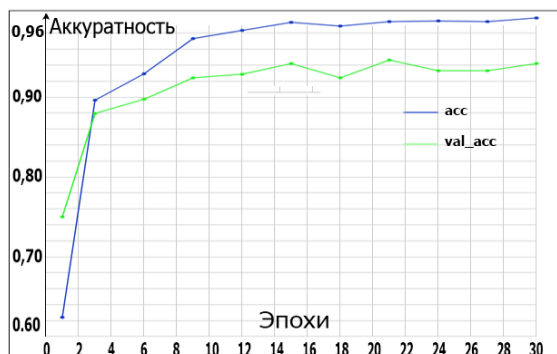


Рис.4. История обучения нейронной сети по эпохам

Как видно из графика, составленного по значениям аккуратности после каждой итерации эпохи, аккуратность сети на тестовых данных равняется 96.43%. Об этом так же говорит функция model.evaluate_generator, представленная на рисунке 5.

```
In [14]: scores = model.evaluate_generator(test_generator,
In [15]: print("Аккуратность на тестовых данных: %.2f%%" %
Аккуратность на тестовых данных: 96.43%
```

Рис.5. Функция оценки аккуратности нейронной сети

5. Пример работы

После того, как сеть обучилась, можно для проверки ее работоспособности запустить на обработку изображение, которое не участвовало в обучении и тестировании. На рисунке 6 изображен процесс подачи изображения на вход нейронной сети, а на рисунке 7 его классификация к классу wall (боковое).



Рис. 6. Подача изображения нейронной сети

```
In [31]: print(prediction)
print(classes[np.argmax(prediction)])

[[0.]]
wall
```

Рис 7. Классификация изображения к классу wall

6. Заключение

В последнее время, свёрточные нейронные сети набирают огромную популярность в решении различных задач, таких как поиск, классификация или обработка изображений. Задачи по анализу освещения не являются исключением. Свёрточные нейронные сети отлично себя показали в этих и смежных задачах. С каждым годом разрабатываются новые техники и рекомендации по более эффективному обучению, появляются новые архитектуры и новые наборы данных, что позволяет надеяться на возможность использования нейронных сетей не только для качественного, но и количественного анализа освещения.

В данной работе рассмотрен подход классификации видов освещения с помощью свёрточных нейронных сетей, который может использоваться для определения типа освещения с дальнейшей его обработкой и восстановления оптических параметров сцены. Хотя нейронная сеть в большинстве случаев работает корректно, однако еще многое предстоит усовершенствовать, так, например, возможно улучшить алгоритм определения источников освещения или разработать алгоритм для определения диффузных поверхностей. Помимо этого, следует дополнить набор данных тестовых изображений и изображений для тестирования.

7. Литература

- [1] Kevin Karsch, Varsha Hedau, David Forsyth, Derek Hoiem, Rendering synthetic objects into legacy photographs // Proceedings of the 2011 SIGGRAPH Asia Conference Article No. 157
- [2] Alex Krizhevsky, Geoffrey E. Hinton, Ilya Sutskever, ImageNet Classification with Deep Convolutional Neural Networks // Advances in neural information processing systems 2012, 1097-1105.
- [3] Convolutional Neural Networks (LeNet) - DeepLearning 0.1 documentation. DeepLearning 0.1. LISA Lab.
- [4] O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), pages 234–241, 2015.
- [5] Y. Hold-Geoffroy, K. Sunkavalli, S. Hadap, E. Gambaretto, and J.-F. Lalonde. Deep outdoor illumination estimation. In Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
- [6] J.-F. Lalonde, A. A. Efros, and S. G. Narasimhan. Estimating the natural illumination conditions from a single outdoor image. International Journal of Computer Vision, 98(2): 123–145, 2012.
- [7] M.-A. Gardner, K. Sunkavalli, E. Yumer, X. Shen, E. Gambaretto, C. Gagné, and J.-F. Lalonde. Learning to predict indoor illumination from a single image. ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia), preprints, 2017.
- [8] Song Han, Huizi Mao, William J. Dally, Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman

- Coding // Computer Vision and Pattern Recognition (cs.CV); Neural and Evolutionary Computing (cs.NE)
- [9] Jonathan Long, Evan Shelhamer, Trevor Darrell, Fully Convolutional Networks for Semantic Segmentation // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3431-3440
- [10] GE Hinton, V Nair, Rectified linear units improve restricted boltzmann machines // In Proceedings of the 27th International Conference on Machine Learning. (ICML-10), 807–814.

Об авторах

Сорокин Максим Игоревич, аспирант кафедры визуализации (базовой) факультета программной инженерии и компьютерной техники Университета ИТМО. Его e-mail vergotten@gmail.com.

Жданов Дмитрий Дмитриевич, к.ф.-м.н., заведующий кафедрой технологий визуализации Университета ИТМО. Его e-mail ddzhdanov@mail.ru.