

Распознавание оптических мод Эрмита-Гаусса с использованием нейронной сети

В.В. Прудовская

Самарский национальный исследовательский университет им. академика С.П. Королева
Самара, Россия
vvrnlk@gmail.com

Аннотация—В данной работе праведно исследование влияния аффинных преобразований на точность распознавания мод Эрмита-Гаусса. Показано, что наличие искажений снижает точность классификации более, чем на 10%.

Ключевые слова—нейронные сети, моды Эрмита-Гаусса, аффинные преобразования, машинное обучение.

1. ВВЕДЕНИЕ

Не так давно нейронные сети вызвали революцию в сфере исследований искусственного интеллекта, благодаря своим достижениям в таких областях, как интерпретация голоса в текст, а также распознавание образов на изображениях [1]. Благодаря своей гибкости нейронные сети смогли найти применение и в различных областях современной науки, в том числе физике и химии [2]. Одно из потенциальных приложений нейронных сетей – телекоммуникации [3]. Оптические коммуникации основаны на генерации, передаче и обнаружении состояний света для кодирования и передачи информации [4]. Для увеличения скорости передачи информации при помощи оптической связи были разработаны многочисленные протоколы [5]. Текущие препятствия для еще более высоких скоростей передачи битов носят чисто технический характер. Основной технической трудностью является точная классификация значений, обнаруженных на приемном конце коммуникационной платформы или на выходном слое установки. Традиционный метод распознавания требует сложного процесса оптического выравнивания и обеспечивает стабильно плохие результаты для сигналов, содержащих шум. Отличительной особенностью нейронных сетей, относительно традиционных методов распознавания изображений на выходном слое различных оптических установок, является возможность анализа изображения при большом количестве искажений.

Целью данной работы является исследовать возможность сверточной нейронной сети распознавать оптические моды Эрмита-Гаусса при наличии в наборе данных другого класса оптических мод – Лаггера-Гаусса, а также изучить влияние аффинных преобразований на точность распознавания.

2. ФОРМИРОВАНИЕ НАБОРОВ ДАННЫХ

Рассмотрим пучок Эрмита-Гаусса, комплексная амплитуда которого имеет вид [6]:

$$E_{n,m}(x, y, z) = i^{n+m} \left(\frac{w}{w(z)}\right)^2 H_n\left(\frac{\sqrt{2x}}{w(z)}\right) H_m\left(\frac{\sqrt{2y}}{w(z)}\right) \exp\left(-\frac{x+y^2}{w^2(z)} + \frac{ik(x+y^2)}{2R(z)}\right) \exp\left(-i(n+m+1)\arctg\left(\frac{z}{z_0}\right)\right) \quad (1)$$

где $w(z)$ – радиус гауссова пучка, $R(z)$ – радиус кривизны волнового фронта гауссова пучка, z_0 – длина Рэлея, k – волновое число света.

Для обучения нейронной сети были получены наборы изображений оптических мод - моды Эрмита-Гаусса в качестве основного класса, моды Лаггера-Гаусса – в качестве второго класса. Для формирования изображений была использована библиотека LightPipes языка Python. Полученные в результате работы программы изображения мод отличаются степенью и порядком. Также были определены неизменяемые параметры, такие как длина волны λ , размер области 15 мкм, размерность k и масштаб w . Отличие сгенерированных мод друг от друга заключается в их степени и порядке, именно поэтому другие параметры, не влияющие на итоговое изображение – постоянные величины. Результаты работы программы, генерирующей оптические моды, представлены на Рис. 1.

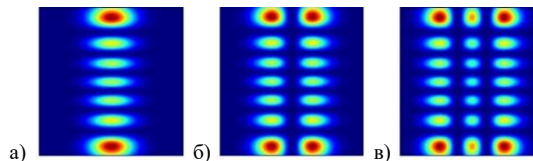


Рис. 1. Изображения мод Эрмита-Гаусса при заданных входных параметрах. а) Степень $l = 0$; б) Степень $l = 1$; в) Степень $l = 2$

Путем варьирования приведенных выше параметров были получены следующие наборы данных: четыре тысячи изображений мод Лаггера-Гаусса и столько же изображений мод Эрмита-Гаусса. Для оценки влияния аффинных преобразований на точность распознавания нейронной сети была реализована программа, осуществляющая преобразования. Программа была написана на языке Python с использованием библиотеки Keras. В результате работы программы получен набор данных из сдвинутых и повернутых под определенным углом изображений мод. Примеры изображений мод после применения аффинных преобразований представлены на Рис. 2.

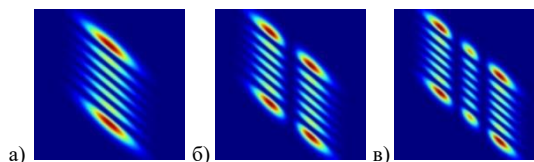


Рис. 2. Изображения мод Эрмита-Гаусса после применения к ним аффинных преобразований: а) Степень $l = 0$; б) Степень $l = 1$; в) Степень $l = 2$

3. МОДЕЛИРОВАНИЕ

Для однозначной классификации изображений мод Эрмита-Гаусса на языке Python была написана сверточная нейронная сеть. За основу обучающего набора была взята коллекция оптических мод Эрмита-Гаусса, отличающихся порядком и степенью. Нейронная сеть должна определять принадлежность поступившего на вход изображения к определенному классу.

Сверточная нейронная сеть основана на использовании последовательной модели Sequential. В структуре участвуют сверточные слои Conv2D с функцией активации ReLU, вспомогательные слои MaxPooling2D и Dropout, а также полносвязные слои Dense с функциями активации ReLU и Sigmoid. Вместе они формируют структуру сверточной нейронной сети, представленную на Рис. 3.

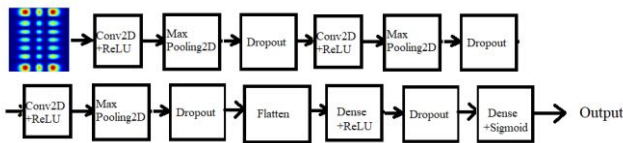


Рис. 3. Архитектура разработанной сверточной нейронной сети

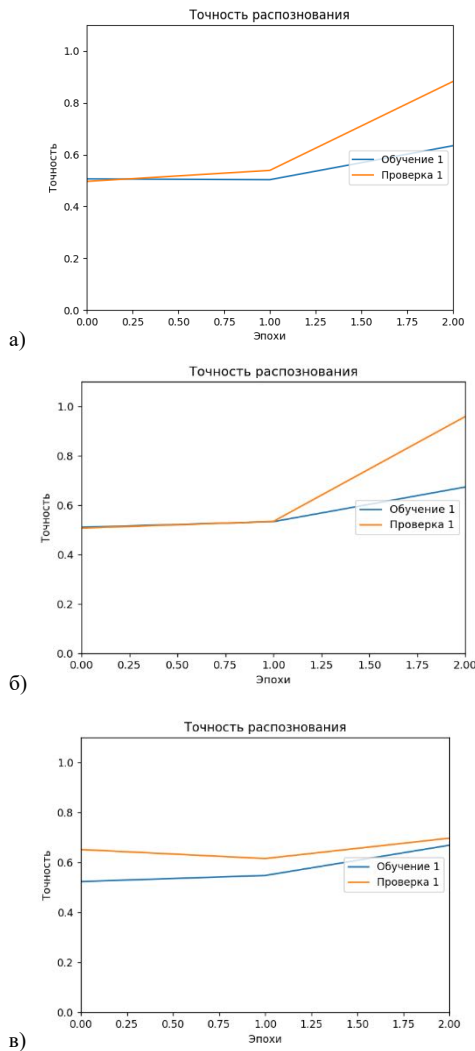


Рис. 4. Графики точности распознавания мод Эрмита-Гаусса на наборе изображений: а) без преобразований; б) с преобразованиями; в) из комбинации наборов с преобразованиями и без

Обучение и тестирование нейронной сети для оценки влияния аффинных преобразований на точность распознавания проводилось на разных наборах данных объемом 10 000 изображений. Было сформировано три набора данных: первый содержит только моды, сгенерированные с помощью библиотеки LightPipes, второй только изображения, подвергнутые преобразованиям, а третий набор является комбинацией. На Рис. 4 представлены графики точности распознавания, которые получились в результате проведенного исследования.

4. ЗАКЛЮЧЕНИЕ

В результате данной работы были сформированы наборы данных, содержащих изображения мод Эрмита-Гаусса и Лаггера-Гаусса, а также эти же изображения, подвергнутые аффинным преобразованиям. Была исследована зависимость точности распознавания смоделированной нейронной сетью от входных данных. Было установлено, что точность классификации мод Эрмита-Гаусса имеет высокий показатель, порядка 95% при распознавании изображений без аффинных преобразований. Также достаточно высокая точность достигается и при использовании смешанного набора данных – 88%. Точность классификации достигает меньшего значения в случае, когда все входные изображения подвергаются аффинным преобразованиям, однако она все равно составляет 69%. Данное исследование может применяться для дальнейших работ, направленных на увеличение пропускной способности оптических систем связи.

ЛИТЕРАТУРА

- [1] Harizi, R. Convolutional neural network with joint stepwise character/word modeling based system for scene text recognition / R. Harizi, R. Walha, F. Drira // Multimedia Tools and Applications. – 2021. – Vol. 80(7). – P. 34-45.
- [2] Ye, S. Symmetrical Graph Neural Network for Quantum Chemistry with Dual Real and Momenta Space / S. Ye, J. Liang, R. Liu, X. Zhu // The Journal of Physical Chemistry A. – 2020. – Vol. 124(34). – P. 6945-6953.
- [3] Yuan, X. Dual-Output Mode Analysis of Multimode Laguerre-Gaussian Beams via Deep Learning / X. Yuan, Y. Xu, R. Zhao, X. Hong, R. Lu, X. Feng, Y. Chen, J. Zou, C. Zhang, Y. Qin, Y. Zhu // Optics. – 2021. – Vol. 2(2). – P. 87-95.
- [4] Anbarasi, K. A review on channel models in free space optical communication system / K. Anbarasi, C. hermanth, R. Sangeetha // Optics & Laser Technology. – 2017. – Vol. 97(1). – P. 161-167.
- [5] Satrusalya, S. Review on free space optical communication / S. Satrusalya, L. Goswami // Materialstoday: proceedings. – 2021. – Vol. 45(6). – P. 1-25.
- [6] Котляр, В.В. Лазерные пучки Эрмита-Гаусса с орбитальным угловым моментом / В.В. Котляр, А.А. Ковалёв, А.П. Порфирьев // Компьютерная оптика. – 2014. – Т. 38, № 4. – С. 651-657. DOI: 10.18287/0134-2452-2014-38-4-651-657.