



NEURAL NETWORK APPROACH TO SOLVING THE PROBLEM OF SELF-ACTION OF WAVE FIELDS IN NONLINEAR MEDIA

E. Vasiliev, D. Bolotov, M. Bolotov, L. Smirnov

Lobachevsky State University,
603022, Nizhny Novgorod, Russia

DOI: 10.24412/2073-0667-2022-1-5-16

We consider the problem of propagation of optical impulses in media with Kerr nonlinearity. As a mathematical model describing an optical pulse propagation process, we chose a generalized parabolic equation, which in dimensionless variables has the form of a one-dimensional modified Nonlinear Schrödinger Equation. We trained a fully connected neural network with various optimization functions and did experiments with network configuration and hyperparameters optimization. The conducted experiments have shown the promise of using the quasi-Newtonian L-BGFS optimization function over first-order optimization functions in this problem.

The article presents the results of experiments on training the model using various optimization functions: SDG, RMSProp, Adam, L-BGFS, the last optimizer allows you to train the network about an order faster. We also consider the size of the training sample required to train the model. From the obtained training results, we can conclude that due to the randomly uniform selection of points from the area using the Latin hypercube, it is enough to make a train sample of 10–15 % of the dataset, this will correspond to a step of about 0.12 in z and 0.022 in τ compared to 0.039 in z and 0.008 in τ in a regular grid obtained by numerical methods.

In low-dimensional problems, the use of machine learning is not always appropriate, since training takes much more time than solving the problem using direct numerical simulation. However, as the complexity of the system increases, due to the increase in the number of unknown variables, a huge superiority of machine learning methods is expected due to fast calculation using an already trained network. Also open and interesting for future research is the issue of fast retraining of an already trained model for a problem with new parameters.

Key words: nonlinear Schrödinger equation, neural networks, deep learning, optimization functions.

References

1. George D., Huerta E. A. Deep Learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data // Physics Letters. 2018. B 778, P. 64–70. [Electron. Res.]: <https://doi.org/10.1016/j.physletb.2017.12.053>.
2. Gonoskov A., et al. Employing machine learning for theory validation and identification of experimental conditions in laser-plasma physics // Scientific Reports. 2019. N 9 1, P. 1–15. [Electron. Res.]: <https://doi.org/10.1038/s41598-019-43465-3>.
3. Ravi D., et al. Deep Learning for Health Informatics // IEEE J. of Biomedical and Health Informatics. 2017. N 21 1, P. 4–21. [Electron. Res.]: <https://doi.org/10.1109/JBHI.2016.2636665>.

4. Lachinov D., Vasiliev E., Turlapov V. Glioma Segmentation with Cascaded UNet // BrainLes 2018. LNCS 11384. 2018. [Electron. Res.]: https://doi.org/10.1007/978-3-030-11726-9_17.
5. Kumar M., Yadav N. Multilayer perceptrons and radial basis function neural network methods for the solution of differential equations: A survey // Computers and Mathematics with Applications. 2011. N 62, P. 3796–3811.
6. Hazanov E. A., Mironov S. Y., Muru. Z. Nelinejnoe szhatie sverhmoshchnyh lazernyh impul'sov: kompressiya posle kompressora // Uspekhi fizicheskikh nauk. 2019. N 189, P. 1173–1200. [Electron. Res.]: <https://doi.org/10.3367/UFNr.2019.05.038564>.
7. Ahmanov S. A., Vyslouh V. A., Chirkin A. S. Optika femtosekundnyh lazernyh impul'sov // M.: Nauka. Gl. red. fiz.-mat. lit. 1988.
8. Moxley F. I. Generalized finite-difference time-domain schemes for solving nonlinear Schrödinger equations // A Dissertation Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy/ 2013. [Electron. Res.]: <https://core.ac.uk/download/pdf/236621293.pdf>.
9. Stein M. Large sample properties of simulations using Latin hypercube sampling // Technometrics. 1987. N 29. P. 143–151. [Electron. Res.]: <https://www.jstor.org/stable/1269769>.
10. Glorot X., Bengio Y. Understanding the difficulty of training deep feedforward neural networks // Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. 2010. N 9, P. 249–256. [Electron. Res.]: <http://proceedings.mlr.press/v9/glorot10a.html>.
11. Robbins H., Monro S. A stochastic approximation method // The annals of mathematical statistics. 1951. Vol. 22. P. 400–407. [Electron. Res.]: <https://doi.org/10.1214/aoms/1177729586>.
12. Hilton G. Neural Networks for Machine Learning. Lecture 6a. Overview of mini-batch gradient descent. 2012. [Electron. Res.]: http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf.
13. Kingma D., Ba J. Adam. A Method for Stochastic Optimization // Cornell University Library. 2014. [Electron. Res.]: <https://arxiv.org/abs/1412.6980>.
14. Schraudolph N. N., Yu J., Gunter S. A Stochastic Quasi-Newton Method for Online Convex Optimization // International Conference on Artificial Intelligence and Statistics. 2007. P. 436–443. [Electron. Res.]: <http://proceedings.mlr.press/v2/schraudolph07a/schraudolph07a.pdf>.



НЕЙРОСЕТЕВОЙ ПОДХОД К РЕШЕНИЮ ЗАДАЧИ САМОВОЗДЕЙСТВИЯ ВОЛНОВЫХ ПОЛЕЙ В НЕЛИНЕЙНЫХ СРЕДАХ

Е. П. Васильев, Д. И. Болотов, М. И. Болотов, Л. А. Смирнов

Нижегородский государственный университет им. Н. И. Лобачевского,
603022, Нижний Новгород, Россия

УДК 517.957

DOI: 10.24412/2073-0667-2022-1-5-16

В работе рассматривается возможность применения технологий глубинного обучения для численного решения задачи о распространении оптических импульсов в средах с нелинейностью Керра. В качестве математической модели, описывающей процессы эволюции огибающей электромагнитного излучения, выбрано обобщенное параболическое уравнение, которое в безразмерных переменных имеет вид одномерного модифицированного нелинейного уравнения Шредингера. Была предложена постановка указанной проблемы, позволяющая действовать для расчетов методы искусственного интеллекта, и реализован один из возможных вариантов данного подхода с применением полносвязной нейронной сети для решения физических задач. При этом был проведен анализ различных алгоритмов подбора параметров, ответственных за передачу информации от слоя к слою такой сети в ходе ее обучения. Выполненные исследования показали, что наиболее перспективными с точки зрения скорости вычислений и адекватности предсказаний являются квази-ニュтоновские функции оптимизации, которые в стандартных библиотеках имеют аббревиатуру L-BGFS.

Ключевые слова: нелинейное уравнение Шредингера, нейронные сети, глубокое обучение, функции оптимизации.

Введение. В последние годы технологии слабого искусственного интеллекта и глубинного обучения, изначально создавшие большой прорыв в компьютерном зрении, активно применяются как в цифровой и промышленной индустрии, так и в различных областях наук, например, в физике [1, 2], биологии [3], медицине [4]. В свою очередь, в основе большого числа базовых математических моделей, достоверно описывающих фундаментальные явления и процессы в природе, лежат дифференциальные уравнения в частных производных. В связи с этим в настоящее время существенно возрастает интерес к использованию нейросетевого моделирования для анализа возможных решений такого рода задач с учетом различных вариантов начальных и граничных условий, а также потенциально нетривиальной геометрии окружения. Идея применения алгоритмов машинного обучения как альтернативы аналитическим методам исследования проблем, возникающим в

Работа выполнена при поддержке Проекта № 0729-2021-013, выполняемого в рамках Государственного задания на выполнение научно-исследовательских работ лабораториями, прошедшиими конкурсный отбор в рамках национального проекта „Наука и университеты“, в отношении которых принято решение Бюджетной комиссии Минобрнауки России (от 14.09.2021 № БК-П/23) о предоставлении из федерального бюджета субсидии на финансовое обеспечение государственного задания на выполнение научно-исследовательских работ.

нелинейной динамике и математической физике, не нова [5], однако ранее такой подход был ограничен доступными вычислительными ресурсами и существующими библиотеками. С распространением параллельных вычислений на графических ускорителях и развитием таких инструментов как TensorFlow, Torch резко увеличилась целесообразность и востребованность программных реализаций и развития идей, позволяющих задействовать средства интеллектуальной обработки больших массивов данных при изучении долговременной эволюции полей и классификации режимов их распространения в композитных средах.

В данной работе для апробации поиска направлений модернизации основных концепций использования искусственных нейронных сетей при построении решений уравнений в частных производных рассмотрена задача о самовоздействии волновых пучков и пакетов в нелинейных и неоднородных средах [6]. Математическое описание подобных волновых процессов обычно проводят в рамках параболического приближения, которое неоднократно подтверждало свою эффективность в различных областях классической физики, в частности, в физике плазмы, гидродинамике и нелинейной оптике, а также при изучении распространения радиоволн в ионосфере и анализе передачи электромагнитного излучения по системам, состоящим из открытых резонаторов или волноводов. Стоит особо отметить, что выбор и актуальность обсуждаемой проблемы обусловлены дополнительно еще одним обстоятельством. Несмотря на принципиальное различие в самой природе явлений, если принять во внимание коллективные эффекты, то в квантовой механике и теории конденсированных сред нередко приходят к аналогичным по своей структуре уравнениям. Наиболее наглядно это можно продемонстрировать в случае вырожденных квантовых газов (бозе-частиц) в приближении среднего поля для описания возбуждения эволюции и взаимодействия нелинейных когерентных локализованных образований (светлых и темных солитонов, одиночных квантовых вихрей, вихревых пар и вихревых колец) в бозе-эйнштейновском конденсате среднего поля удается ввести комплексную волновую функцию, играющую роль параметра порядка и удовлетворяющую уравнению Гросса-Питаевского, которое в безразмерных единицах идентично нелинейному уравнению Шредингера (НУШ) при наличии в нем нелинейного потенциала, связанного с внешними неоднородностями.

Ниже представлен ряд результатов по использованию полно связанных искусственных нейронных сетей для предсказаний наиболее вероятного сценария распространения оптических волн сквозь нелинейный кристалл: произведено сравнение построенного прогноза с данными, получаемыми путем прямого численного моделирования с помощью принятых стандартных схем расчета.

1. Постановка задачи. Для большей конкретики нами рассматривается задача о распространении оптических импульсов в средах с нелинейностью Керра [6]. При этом в качестве математической модели, описывающей процессы самовоздействия интенсивного лазерного излучения при прохождении через такой кристалл, выбрано обобщенное параболическое уравнение, которое в безразмерных переменных имеет вид модифицированного НУШ [6]:

$$i\frac{\partial\psi}{\partial z} - \frac{1}{2}\Delta_{x,y}\psi - \delta(x,y)\psi - \frac{\alpha}{2}\frac{\partial^2\psi}{\partial\tau^2} - \beta\left(|\psi|^2\psi + i\gamma\frac{\partial(|\psi|^2\psi)}{\partial\tau}\right) = 0 \quad (1)$$

где комплексная функция $\psi(\tau,x,y,z)$ пропорционально медленно изменяющейся амплитуде огибающей нестационарного квазимохроматического поля в слабо неоднородном

образце диспергирующего материала, а $\delta(x,y)$ представляет собой относительное изменение его диэлектрической проницаемости. В свою очередь, постоянные α и β могут принимать значения ± 1 в зависимости от знака дисперсии групповой скорости и нелинейной составляющей показателя преломления, соответственно. В соотношении (1) учтен не только квазистатический кубический отклик среды на приложенное электромагнитное воздействие, отвечающий за первый член заключенного в скобки составного слагаемого в левой части равенства (1), но и эффект волновой нестационарности, который проявляет себя фактически в форме зависимости групповой скорости от интенсивности поля, т. е. как нелинейная дисперсия, и приводит к возникновению второго члена в окруженному скобками выражении. Коэффициент γ , участвующий там в качестве множителя, вычисляется по следующей формуле: $\gamma = k'_\omega(\omega_0)/\omega_0 k''_{\omega_0}(\omega_0)$, где $k(\omega)$ — волновой вектор, ω_0 — центральная (несущая) частота излучения. Стоит отметить, что данный эффект вызывает „самообострение“ фронта импульса, вследствие чего происходит формирование ударной волны огибающей. Подчеркнем также, что уравнение (1) допускает обобщение на векторные поля, а также на другие типы нелинейностей. Кроме того, для тех исследований, где это необходимо, в (1) можно принять во внимание старшие производные по τ , т. е. более высокие порядки дисперсии.

Для моделирования процесса распространения оптического импульса в нелинейном кристалле в рамках параболического приближения на основе НУШ (1) разработан целый ряд численных схем [7, 8]. Одним из наиболее популярных и эффективных алгоритмов является метод расщепления с использованием быстрого преобразования Фурье [8]. Однако данная задача остается весьма трудоемкой и затратной с точки зрения вычислений, особенно когда в ходе расчетов появляются все более мелкие характерные масштабы, как по пространственным координатам из-за самофокусировки, так и по связанной со временем переменной τ за счет самообострения. Основная сложность состоит в том, что в какой-то момент размеры возникающей неоднородности становятся сравнимы с расстояниями между узлами (ячейками) расчетной (чаще всего эквидистантной) сетки, с которыми к тому же должен быть согласован элементарный шаг эволюции. В итоге, при попытках избежать нефизических результатов (в том числе, развития численных неустойчивостей) и повысить разрешающие возможности в ходе моделирования значительно возрастает объем обрабатываемых данных и вместе с тем резко увеличивается количество необходимых итераций используемой процедуры. Поэтому существует необходимость в разработке альтернативных методов расчета динамики волновых полей не только в приложении к нелинейной оптике, но и для других областей физики. В настоящее время перспективным направлением здесь представляется использование достижений слабого искусственного интеллекта, которые позволяют делать с высокой степенью точности достоверный прогноз о виде функции и выявлять скрытые закономерности в наборе исходных правил на основе сформулированных ограничений на класс искомых зависимостей и априорной информации о семействе возможных ответов для исследуемой проблемы. Данный подход снимает ограничения, которые имеются у классических алгоритмов поиска решений уравнений в частных производных, где проводится процедура дискретизации непрерывных переменных, т. е. их замены на упорядоченный ряд дискретных точек и формирование сетки узлов в многомерном пространстве.

С целью апробации, развития и тестирования идей нейросетевого моделирования (интеллектуальных вычислений) далее нами рассматривается случай распространяющегося в практически однородной среде локализованного лазерного излучения с большой попереч-

ной апертурой, в связи с чем можно пренебречь линейными дифракционными эффектами и считать структуру поля в плоскости x,y заданной и неизменной. Тогда уравнение (1) существенным образом упрощается и сводится к одномерному модифицированному НУШ:

$$i \frac{\partial \psi}{\partial z} - \frac{\alpha}{2} \frac{\partial^2 \psi}{\partial \tau^2} - \beta \left(|\psi|^2 \psi + i \gamma \frac{\partial (|\psi|^2 \psi)}{\partial \tau} \right) = 0 \quad (2)$$

Для его решения с помощью методов глубинного обучения прежде всего требуется составить функционал качества $G(\psi(z,\tau))$, который определяет степень выполнения моделью своего назначения и позволяет путем его минимизации осуществить подбор параметров искусственной нейросети (ее весов и смещений). В выделенном нами классе задач подобные функционалы ошибки следует выбирать, исходя из физических соображений и фундаментальных принципов и законов природы (например, на основе законов сохранения). Кроме того, стоит действовать и дополнительные свойства и особенности изучаемой системы, такие как наличие симметрии и инвариантность относительно специфических преобразований (в частности, сдвигов, смещений, поворотов и т. д.). В качестве первого приближения к построению универсальных схем и рецептов по применению искусственного интеллекта для предсказания режимов эволюции волновых полей нами были рассмотрены процессы самовоздействия оптических импульсов на базе НУШ (2). При этом формирование функционала качества $G(\psi(z,\tau))$ выполнялось прежде всего на базе дифференциальных соотношений, вытекающих из (2):

$$\begin{aligned} G_{PDE}(U_n, V_n) = & \frac{1}{N} \sum_{n=1}^N \left(\frac{\partial U_n}{\partial z} - \frac{\alpha}{2} \frac{\partial^2 V_n}{\partial \tau^2} - \beta \left((V_n^2 + U_n^2) V_n + \gamma \frac{\partial ((V_n^2 + U_n^2) U_n)}{\partial \tau} \right) \right)^2 + \\ & \frac{1}{N} \sum_{n=1}^N \left(\frac{\partial V_n}{\partial z} + \frac{\alpha}{2} \frac{\partial^2 U_n}{\partial \tau^2} + \beta \left((V_n^2 + U_n^2) U_n - \gamma \frac{\partial ((V_n^2 + U_n^2) V_n)}{\partial \tau} \right) \right)^2, \end{aligned} \quad (3)$$

где U_n и V_n — соответственно реальная и мнимая части комплексной функции $\psi(z,\tau) = U(z,\tau) + iV(z,\tau)$ в случайно взятых точках (z_n, τ_n) , которые лежат внутри счетной области. Кроме того, в $G(\psi(z,\tau))$ необходимо добавить две дополнительные компоненты, которые будут отвечать за согласование предсказания искусственной нейронной сети $\psi(z,\tau)$ с начальными и краевыми условиями:

$$G_{IC}(U_m, V_m) = \frac{1}{M} \sum_{m=1}^M \left((U_m - U_m^{(0)})^2 + (V_m - V_m^{(0)})^2 \right), \quad (4)$$

$$G_{BC}(U_k, V_k) = \frac{1}{K} \sum_{k=1}^K (U_k^2 + V_k^2). \quad (5)$$

Здесь пары U_m, V_m и U_k, V_k также представляют собой комбинацию реальной и мнимой частей $\psi(z,\tau)$, взятых соответственно в случайных M точках ($z = 0, \tau_m$) и в K случайных точках ($z_k, \tau = \pm \Gamma$). Видно, что (4) дает среднеквадратичное отклонение результатов нейросетевой модели от истинного начального распределения $\psi(z = 0, \tau) = \psi^{(0)}(\tau) = U^0(\tau) + iV^0(\tau)$, а (5) позволяет оценить близость $\psi(z, \tau = \pm \Gamma) = U(z, \tau = \pm \Gamma) + iV(z, \tau = \pm \Gamma)$ к границам $\tau = \pm \Gamma$ их расчетной области. Отметим, что нулевые граничные условия были выбраны для определенности. Их при необходимости можно заменить на любой другой тип условий. Таким образом, окончательно получаем следующий функционал качества:

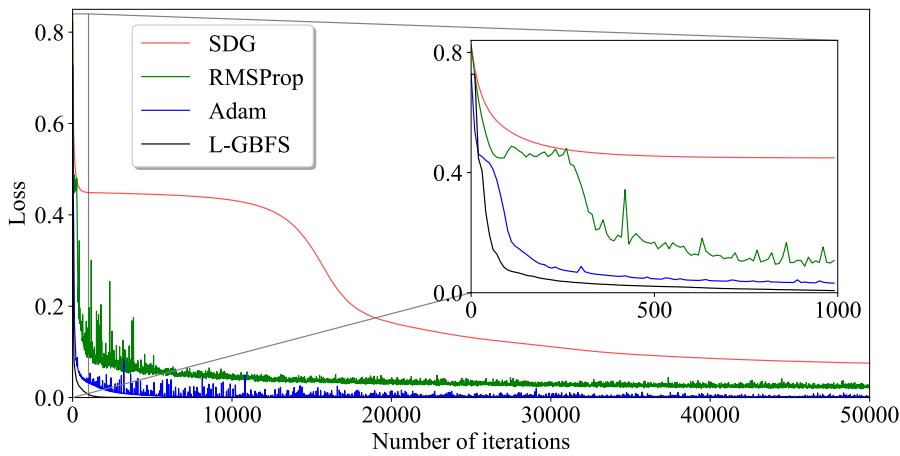


Рис. 1. Значения функции потерь во время обучения модели с помощью различных функций оптимизации для первых 50000 шагов. В графиках выводится значение функции потерь на каждой 10-й итерации

Таблица 1

Значения функции потерь после проведенного обучения и нормализованная среднеквадратическая ошибка для обученных разными оптимизаторами моделей на исходных данных

Оптимизатор	Loss	Ошибка U	Ошибка V	Ошибка ψ
SGD	0.07529	0.8698683	1.163877	0.4125057
RMSProp	0.02175	1.156291	1.197228	0.3141282
Adam	0.0004897	0.06765208	0.1103277	0.01818167
L-BGFS	0.00001284	0.004083095	0.006000219	0.001985626

$$G = G_{PDE}(U_n, V_n) + G_{IC}(U_m, V_m) + G_{BC}(U_k, V_k). \quad (6)$$

Способ генерации наборов из N , M и K случайных точек будет обсуждаться ниже. Отметим, что в данной работе мы ограничились случаем, когда базовая структура нейронной сети представляет собой полносвязный граф. Отдельно стоит выделить вопрос о ширине и глубине конфигурации из внутренних (скрытых) слоев такой модели, а также о функциях активации отдельных элементов.

2. Вычислительный эксперимент. 2.1. *Определение оптимального варианта алгоритма обратного распространения ошибки для обучения нейронной сети.* Исходными данными для обучения является численное решение данной системы 2, они представлены с дискретизацией на равномерной сетке: 256 точек по координате z от -5 до 5 , 201 точка по времени τ от 0 до $\pi/2$, в каждой полученной точке сетки вычислены действительная и мнимая части функции $\psi(z, \tau)$, таким образом, всего имеется 51456 точек. Для обучения в качестве тренировочной выборки используется выборка латинского гиперкуба (Latin hypercube sampling, LHS) [9], для равномерной выборки случайных значений из сетки. В тренировочную выборку S включены 50 случайных точек по z для $\tau = 0$ и по 50

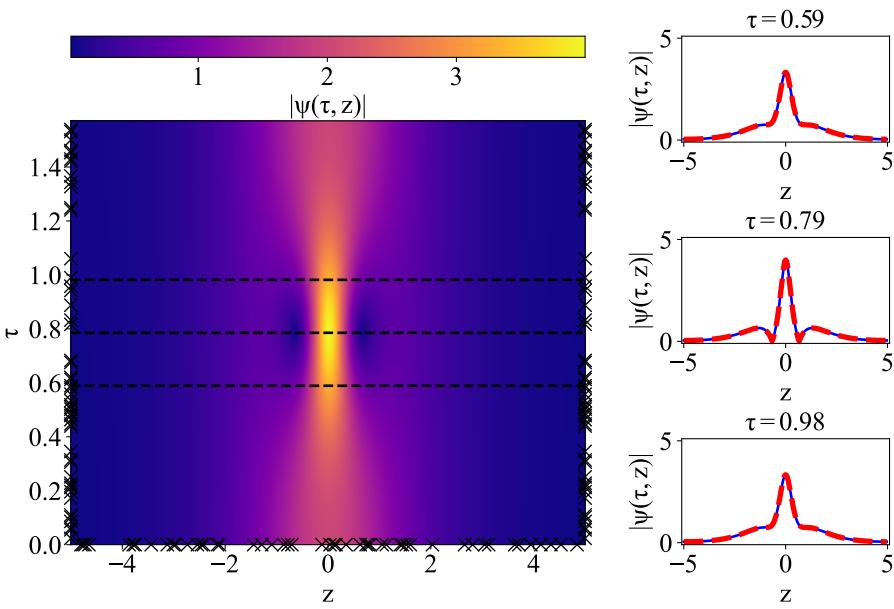


Рис. 2. Слева: график предсказанных значений функции $|\psi(z, \tau)|$ на двумерной сетке $z \in [-5, 5]$ и $\tau \in [0, \pi/2]$. Справа: сравнение предсказанных с помощью нейронной сети и полученных в рамках прямого численного моделирования решений $|\psi(z, \tau)|$, соответствующих трем временным снимкам.

Сплошная линия — результаты численного моделирования, прерывистая линия — результаты предсказания нейронной сети

случайных точек из интервала значений моментов времени $[0, \pi/2]$ для $z = -5$ и $z = 5$, а также 20000 тысяч точек внутри границ.

В качестве модели для экспериментов по обучению используется полносвязная сеть с двумя входами (параметры z и τ), 4 скрытыми слоями по 100 нейронов в каждом слое, и выходной слой с двумя выходами (действительный и мнимый корень НУШ), в качестве функции активации используется гиперболический тангенс. Для начальной инициализации весов используется алгоритм Ксавье [10], начальные веса имеют нормальное распределение и при этом позволяют хорошо обучаться с гиперболическим тангенсом в качестве функции активации.

Проведены эксперименты по использованию в качестве функции оптимизаторов следующих алгоритмов: стохастический градиентный спуск (SGD) [11], RMSProp [12], Adam [13], L-BGFS [14].

Представленные ниже эксперименты по созданию и обучению сети были выполнены в программной библиотеке для машинного обучения TensorFlow v1.14. На рис. 1 представлены графики обучения модели с помощью различных оптимизаторов.

По результатам обучения были получены следующие погрешности относительно заданных значений (таблица 1). Из данной таблицы можно увидеть, что методы SGD и RMSProp показывают наихудшие результаты обучения. Метод Adam справился значительно лучше двух предыдущих, однако ошибка на реальных данных является слишком высокой и не позволяет использовать обученную модель для решения НУШ. Оптимизатор L-BGFS за значительно меньшее число итераций позволяет достичь точности 10^{-3} степени, что показывает перспективность его применения для обучения искусственных нейронных сетей, решающих предсказывающих решение НУШ.

Таблица 2

Значение ошибки для моделей,
обученных на различном размере обучающей выборки

Количество точек	Ошибка U	Ошибка V	Ошибка ψ
500	0,532398	0,762146	0,119605
1000	0,096608	0,182720	0,026675
2000	0,046927	0,081595	0,010848
3000	0,011139	0,019412	0,004159
5000	0,017167	0,027406	0,004724
10000	0,007420	0,013037	0,002648
15000	0,006562	0,010187	0,002382
20000	0,005604	0,008395	0,002236
51456 (все точки)	0,005444	0,008823	0,001897

Таблица 3

Значение ошибки для моделей с различным количеством
внутренних слоев и нейронов

Архитектура сети	Ошибка U	Ошибка V	Ошибка ψ
[2, 32, 32, 2]	0,02664	0,04244	0,00704
[2, 32, 32, 32, 2]	0,00364	0,00578	0,00202
[2, 32, 32, 32, 32, 2]	0,00337	0,00463	0,00166
[2, 32, 32, 32, 32, 32, 2]	0,00230	0,00329	0,00145
[2, 64, 64, 2]	0,00571	0,00884	0,00201
[2, 64, 64, 64, 2]	0,00191	0,00257	0,00128
[2, 64, 64, 64, 64, 2]	0,00182	0,00239	0,00123
[2, 64, 64, 64, 64, 64, 2]	0,00211	0,00288	0,00121
[2, 128, 128, 2]	0,00547	0,00904	0,00181
[2, 128, 128, 128, 2]	0,00184	0,00260	0,00107
[2, 128, 128, 128, 128, 2]	0,00133	0,00162	0,00103
[2, 128, 128, 128, 128, 128, 2]	0,00148	0,00175	0,00115
[2, 32, 64, 2]	0,00605	0,00988	0,00213
[2, 32, 64, 128, 2]	0,00288	0,00393	0,00150
[2, 32, 32, 64, 128, 2]	0,00201	0,00256	0,00158
[2, 32, 64, 64, 128, 2]	0,00202	0,00263	0,00138

2.2. *Определение количества точек для обучения.* Важным параметром обучения модели является размер обучающей выборки. С увеличением размера выборки повышается точность предсказания и линейно увеличивается время обучения. В таблице 2 представлены результаты обучения модели на различном размере подвыборки, начиная от 500 точек и заканчивая полным набором (51456). Из представленных данных можно сделать вывод, что, благодаря случайному-равномерному выбору точек из области с помощью латинского гиперкуба [9], для обучения достаточно сделать выборку 10–15 % от набора, это будет соответствовать шагу около 0.12 по z и 0.022 по τ по сравнению с 0.039 по z и 0.008 по τ в регулярной сетке, полученной численными методами.

2.3. Определение внутренней архитектуры нейронной сети. Открытым для изучения является вопрос внутренней архитектуры нейронной сети. В таблице 3 собраны результаты обучения моделей с различным количеством внутренних слоев и нейронов в них. Для решения НУШ 2 двух скрытых слоев или 32 нейронов в скрытом слое недостаточно, предлагается использовать 4–5 скрытых слоев с количеством скрытых нейронов в каждом слое, близком к 100. Также были исследованы варианты структуры с увеличением количества нейронов в более дальних скрытых слоях, но большого прироста от использования данной стратегии получить не удалось.

Заключение. В статье предлагается решение задачи о распространении оптических импульсов в средах с нелинейностью Керра в форме модифицированного НУШ 2 с помощью нейросетевого подхода: произведено сравнение данных, полученных в результате обучения нейронных сетей с данными, получаемыми путем прямого численного моделирования с помощью принятых стандартных схем расчета.

Проведенные исследования показали возможность применения машинного обучения для решения поставленной задачи. Были проведены эксперименты по использованию различных функций оптимизации для обучения модели, в том числе SGD, RMSProp, Adam, L-BGFS, полученные результаты показывают превосходство квази-ニュтоновского оптимизатора над оптимизаторами первого порядка в данной задаче.

На задачах малой размерности использование машинного обучения не всегда является целесообразным, так как обучение занимает значительно большее количество времени, чем решение задачи с помощью прямого численного моделирования. Однако, при увеличении сложности системы, вследствие увеличения количества неизвестных переменных ожидается превосходство методов машинного обучения из-за быстрого вычисления с помощью уже обученной сети. Также открытым и интересным для будущих исследований является вопрос быстрого дообучения уже обученной модели для задачи с новыми параметрами.

Список литературы

1. George D., Huerta E. A. Deep Learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data // Physics Letters. 2018. B 778, P. 64–70. [Electron. Res.]: <https://doi.org/10.1016/j.physletb.2017.12.053>.
2. Gonoskov A., et al. Employing machine learning for theory validation and identification of experimental conditions in laser-plasma physics // Scientific Reports. 2019. N 9 1, P. 1–15. [Electron. Res.]: <https://doi.org/10.1038/s41598-019-43465-3>.
3. Ravi D., et al. Deep Learning for Health Informatics // IEEE J. of Biomedical and Health Informatics. 2017. N 21 1, P. 4–21. [Electron. Res.]: <https://doi.org/10.1109/JBHI.2016.2636665>.
4. Lachinov D., Vasiliev E., Turlapov V. Glioma Segmentation with Cascaded UNet // BrainLes 2018. LNCS 11384. 2018. [Electron. Res.]: https://doi.org/10.1007/978-3-030-11726-9_17.
5. Kumar M., Yadav N. Multilayer perceptrons and radial basis function neural network methods for the solution of differential equations: A survey // Computers and Mathematics with Applications. 2011. N 62, P. 3796–3811.
6. Хазанов Е. А., Миронов С. Ю., Муру. ЖК. Нелинейное сжатие сверхмощных лазерных импульсов: компрессия после компрессора // Успехи физических наук. 2019. № 189, С. 1173–1200. [Электронный ресурс]: <https://doi.org/10.3367/UFNr.2019.05.038564>.
7. Ахманов С. А., Выслюух В. А., Чиркин А. С. Оптика фемтосекундных лазерных импульсов // М.: Наука. Гл. ред. физ.-мат. лит. 1988.

8. Moxley F. I. Generalized finite-difference time-domain schemes for solving nonlinear Schrödinger equations // A Dissertation Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy / 2013. [Electron. Res.]: <https://core.ac.uk/download/pdf/236621293.pdf>.
9. Stein M. Large sample properties of simulations using Latin hypercube sampling // Technometrics. 1987. N 29. P. 143–151. [Electron. Res.]: <https://www.jstor.org/stable/1269769>.
10. Glorot X., Bengio Y. Understanding the difficulty of training deep feedforward neural networks // Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. 2010. N 9, P. 249–256. [Electron. Res.]: <http://proceedings.mlr.press/v9/glorot10a.html>.
11. Robbins H., Monro S. A stochastic approximation method // The annals of mathematical statistics. 1951. Vol. 22. P. 400–407. [Electron. Res.]: <https://doi.org/10.1214/aoms/1177729586>.
12. Hilton G. Neural Networks for Machine Learning. Lecture 6a. Overview of mini-batch gradient descent. 2012. [Electron. Res.]: <http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture\slides\lec6.pdf>.
13. Kingma D., Ba J. Adam. A Method for Stochastic Optimization // Cornell University Library. 2014. [Electron. Res.]: <https://arxiv.org/abs/1412.6980>.
14. Schraudolph N. N., Yu J., Gunter S. A Stochastic Quasi-Newton Method for Online Convex Optimization // International Conference on Artificial Intelligence and Statistics. 2007. P. 436–443. [Electron. Res.]: <http://proceedings.mlr.press/v2/schraudolph07a/schraudolph07a.pdf>.



Васильев Евгений Павлович — младш. науч. сотр. научно-исследовательской лаборатории „Искусственного интеллекта и обработки больших массивов данных“, старший преподаватель кафедры математического обеспечения и суперкомпьютерных технологий института информационных технологий, математики и механики ННГУ им. Н. И. Лобачевского; e-mail: eugene.unn@gmail.com, тел. 8-831-462-33-20. Область научных интересов: научная визуализация, компьютерное зрение, глубокое обучение.

Evgennii Vasiliev is a Research Associate at the Laboratory of Artificial Intelligence and Big Data Processing of the Department of Control Theory and System Dynamics and lecturer of Department of Software and Supercomputer Technologies of Institute of Information Technology, Mathematics and Mechanics, Lobachevsky State University of Nizhny Novgorod. Research interests: scientific visualization, computer vision, deep learning.

Болотов Дмитрий Ильич — лаборант-исследователь научно-исследовательской лаборатории „Искусственного интеллекта и обработки больших массивов данных“ кафедры тео-

рии управления и динамики систем Института информационных технологий, математики и механики ННГУ им. Н. И. Лобачевского; e-mail: bolotovdm@gmail.com, тел. 8-831-462-33-20. Область научных интересов: искусственный интеллект, глубокое обучение, нелинейная динамика.



Dmitry Bolotov is a research assistant at the Laboratory of Artificial Intelligence and Big Data Processing of the Department of Control Theory and System of Institute of Information Technology, Mathematics and Mechanics, Lobachevsky State University of Nizhny Novgorod. Research interests: artificial intelligence, deep learning, nonlinear dynamics.



Болотов Максим Ильич — канд. физ.-мат. наук, младший науч. сотр. научно-исследовательской лаборатории „Искусственного интеллекта и обработки больших массивов данных“, старший преподаватель кафедры теории управления и динамики систем Института информационных технологий, математики и механики ННГУ им. Н. И. Лобачевского; e-mail:

maksim.bolotov@itmm.unn.ru, тел. 8-831-462-33-20. Область научных интересов: нелинейная динамика, пространственно-временные структуры, распределенные системы, динамические системы.

Maxim Bolotov is a PhD. in Physical and Mathematical Sciences, Research Associate at the Laboratory of Artificial Intelligence and Big Data Processing of the Department of Control Theory and System Dynamics and lecturer of Department of Control Theory and System Dynamics of Institute of Information Technology, Mathematics and Mechanics, Lobachevsky State University of Nizhny Novgorod. Research interests: nonlinear dynamics, space-time structures, distributed systems, dynamic systems.

Смирнов Лев Александрович – канд. физ.-мат. наук, заведующий научно-исследовательской лаборатории „Искусственного интеллекта и обработки больших массивов данных“ кафедры теории управления и дина-

мики систем Института информационных технологий, математики и механики ННГУ им. Н. И. Лобачевского; e-mail: lev.smirnov@itmm.unn.ru, тел. 8-831-462-33-20. Область научных интересов: нелинейные волны (солитоны и вортексы), неравновесная статистическая механика, турбулентность, высокопроизводительные вычисления.



Lev Smirnov is a PhD. in Physical and Mathematical Sciences, the Head of the Laboratory of Artificial Intelligence and Big Data Processing of the Department of Control Theory and System of Institute of Information Technology, Mathematics and Mechanics, Lobachevsky State University of Nizhny Novgorod. Research interests: non-linear waves (solitons and vortices), non-equilibrium statistical mechanics, turbulence, high-performance computing.

Дата поступления – 01.02.2022