Сравнительный анализ нейронных сетей с комплексными функциями активации

Н. Ф. Алексиадис*, А. А. Сет‡

В докладе рассматривается сравнительный анализ нейронных сетей с разными комплексными функциями активации (CReLU, zReLU, ModReLU, Cardioid и ModTanh). Результаты анализа показали, что функция ModTanh обеспечивает наиболее высокую эффективность обучения.

Ключевые слова: нейронные сети, функции активации, скорость обучения, устойчивость модели.

1. Постановка задачи

Целью настоящего исследования является сравнительный анализ влияния комплексных функций активации на эффективность обучения при использовании одной и той же архитектуры комплексной многослойной нейронной сети.

Для достижения поставленной цели рассматриваем следующие вопросы: построение архитектуры изучаемой нейронной сети, анализ трёх методов инициализации комплексных весов, сравнение результатов различных комплексных функций активации по точности, скорости, устойчивости и выбору наилучшей функции активации.

Для эксперимента мы используем FFT преобразование данных из MNIST (см. https://www.kaggle.com/datasets/hojjatk/mnist-dataset, дата обращения: 10.09.2026).

Для сравнительного анализа комплексных функций активации был выбран комплексный многослойный персептрон [1, 2]. Общая структура

* *Алексиадис Никос Филиппович* — канд. физ.-мат. наук, доцент кафедры прикладной математики и искусственного интеллекта института информационных и вычислительных технологий Национального исследовательского университета «МЭИ», e-mail: aleksiadis@yandex.ru, ORCID: 0009-0006-2863-1148.

Aleksiadis Nikos Philippovich — PhD, associate professor, National Research University «MPEI», Institute of Information and Computation technologies, Chair of applied mathematics and artificial intelligence.

‡ *Сет Аунг Айар* — аспирант кафедры прикладной математики и искусственного интеллекта института информационных и вычислительных технологий Национального исследовательского университета «МЭИ», e-mail: aiara@mpei.ru.

Ayar Aung Set — PhD student, National Research University «MPEI», Institute of Information and Computation technologies, Chair of applied mathematics and artificial intelligence.

Опубликовано на условиях лицензии CC-BY 4.0
Published under CC-BY 4.0

построенной комплексной нейронной сети представлена на рис. 1.

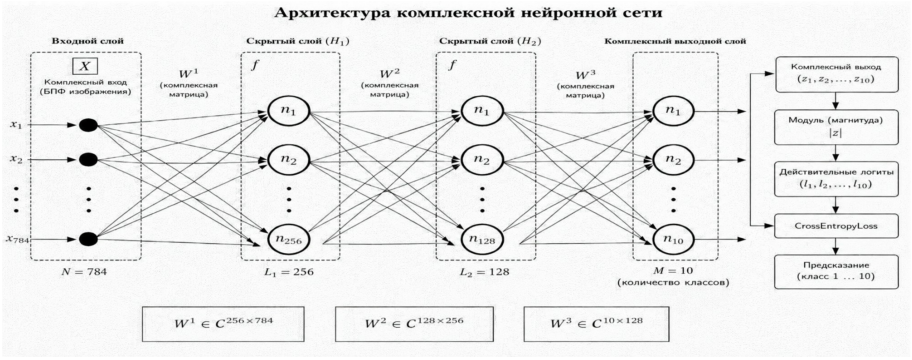


Рис. 1. Архитектура комплексной нейронной сети

В рамках настоящего исследования в качестве f рассматриваем 5 комплексных функций [3, 4] (где $ReLU(x) = max(0, x)$):

$CReLU = ReLU(Re(z)) + iReLU(Im(z)), ModTanh = \tanh(|z|)(z/|z|)$ при $z \neq 0$ и $ModTanh(0) = 0$, $ModReLU = ReLU(|z|+b)(z/|z|)$, при $z \neq 0$, $ModReLU(0) = 0$, где b – обучаемый параметр, $Cardioid = (1/2)(1 + \cos(\arg(z)))z$, $zReUL(z) = z$, если$Re(z) > 0$ и $Im(z) > 0$, иначе 0.

В выходном слое формируется комплексный вектор логитов $Z \in \mathbb{C}^{10}$.

Для перехода к пространству действительных чисел, необходимому для вычисления функции потерь, выполняется операция взятия модуля ($magnitudeoperation$) для каждого комплексного логита. В итоге комплексные логиты преобразуются в действительный вектор логитов. Полученный вектор логитов передаётся в функцию потерь $CrossEntropyLoss$, значения которой используются для вычисления градиентов методом обратного распространения ошибок.

2. Экспериментальная часть исследования

Экспериментальная часть исследования была выполнена с использованием языка программирования Python и библиотеки PyTorch, а обучение комплексной нейронной сети осуществлялось с использованием алгоритма обратного распространения ошибки (Backpropagation). Размер минипакета (batch size) установлен равным 64, что обеспечило оптимальный баланс между скоростью сходимости и вычислительной эффективностью. В процессе обучения эксперименты проводились на протяжении 100 эпох. В качестве функции потерь применялась перекрестная энтропия

(CrossEntropyLoss). Обновление весов выполнялось на основе градиентного метода оптимизации Adam.

3. Сравнительный анализ результатов

Первый этап экспериментального исследования был посвящен анализу влияния скорости обучения на эффективность и устойчивость комплексной нейронной сети.

Второй этап экспериментального исследования был посвящен анализу влияния различных методов инициализации весов на эффективность обучения комплексной нейронной сети.

Третий этап эксперимента был посвящен сравнительному анализу эффективности разных комплексных функций активации. Целью данного эксперимента являлось определение наиболее эффективной функции активации при одинаковой архитектуре сети и неизменных параметрах обучения. Полученные результаты представлены в таблице 1.

Табл.1. Сравнение эффективности функций активации

Функция активации	Тестовая точность (%)	Стандартное отклонение (%)	Потери на валидации
ModReLU	98.06	0.07	0.0622
ModTanh	98.14	0.02	0.0608
CReLU	97.97	0.05	0.0688
Cardioid	97.84	0.02	0.0728
zReLU	97.46	0.29	0.0989

4. Заключение

Исследование показало, что ModTanh демонстрирует наилучшие результаты из всех рассмотренных комплексных функций активации.

Список литературы

[1] C. Y. Lee, H. Hasegawa, S. C. Gao, “Complex-Valued Neural Networks: A Comprehensive Survey”, *IEEE/CAA Journal of Automatica Sinica*, **9** (2022), 1406–1426. DOI: 10.1109/JAS.2022.105743.

[2] C. Trabelsi, O. Bilaniuk, Y. Zhang, et al., “Deep Complex Networks”, International Conference on Learning Representations (ICLR) (Vancouver), 2018.

[3] M. Arjovsky, A. Shah, Y. Bengio, “Unitary Evolution Recurrent Neural Networks”, Proceedings of the 33rd International Conference on Machine Learning (ICML), 2016, 1120–1128.

[4] E. K. Cole, J. Y. Cheng, J. M. Pauly, S. S. Vasanawala, “Analysis of Deep Complex-Valued Convolutional Neural Networks for MRI Reconstruction and Phase-Focused Applications”, *Magnetic Resonance in Medicine*, **86** (2021), 1093–1109. DOI: 10.1002/mrm.28733.

Статья поступила 18 октября 2026 г.

Comparative analysis of neural networks with complex activation functions

N. P. Aleksiadis, A. S. Ayar

The report examines a comparative analysis of neural networks with different complex activation functions (CReLU, zReLU, ModReLU, Cardioid, and ModTanh). The analysis results showed that the ModTanh function provides the highest training efficiency.

Keywords: neural networks, activation functions, learning rate, model stability.

References

[1] C. Y. Lee, H. Hasegawa, S. C. Gao, “Complex-Valued Neural Networks: A Comprehensive Survey”, *IEEE/CAA Journal of Automatica Sinica*, **9** (2022), 1406–1426. DOI: 10.1109/JAS.2022.105743.

[2] C. Trabelsi, O. Bilaniuk, Y. Zhang, et al., “Deep Complex Networks”, International Conference on Learning Representations (ICLR) (Vancouver), 2018.

[3] M. Arjovsky, A. Shah, Y. Bengio, “Unitary Evolution Recurrent Neural Networks”, Proceedings of the 33rd International Conference on Machine Learning (ICML), 2016, 1120–1128.

[4] E. K. Cole, J. Y. Cheng, J. M. Pauly, S. S. Vasanawala, “Analysis of Deep Complex-Valued Convolutional Neural Networks for MRI Reconstruction and Phase-Focused Applications”, *Magnetic Resonance in Medicine*, **86** (2021), 1093–1109. DOI: 10.1002/mrm.28733.

Received on October 18, 2026