УДК 004.032.26

МОДЕЛИРОВАНИЕ ПРИМИТИВНОЙ МНОГОСЛОЙНОЙ ОПТИЧЕСКОЙ НЕЙРОННОЙ СЕТИ ДЛЯ ОПРЕДЕЛЕНИЯ ЦИФР

В.Н. Кречко[1]

[1]Самарский университет, г. Самара
e-mail: krechkovsevolod@mail.ru

Аннотация

В работе представлены результаты исследования дифракционных оптических нейронных сетей (D2NN) для классификации изображений. Разработана математическая модель многослойной оптической сети на основе дифракции Френеля и фазовой модуляции. Реализован программный прототип на Python с использованием PyTorch. Достигнута точность классификации 89,7% на наборе данных MNIST при энергопотреблении ~1 мкДж на операцию, что на два порядка меньше цифровых аналогов.

Ключевые слова: дифракционная оптическая нейронная сеть, дифракция Френеля, фазовая модуляция, машинное обучение, оптические вычисления.

Введение

Цифровые вычислительные системы приближаются к физическим пределам миниатюризации, сталкиваясь с проблемами тепловыделения и роста энергопотребления. Дата-центры уже потребляют 1–3% мировой электроэнергии, к 2030 г. прогнозируется 8–20% [1, 2]. Альтернативой выступают оптические вычисления, обеспечивающие отсутствие омических потерь, параллелизм и высокую скорость передачи данных [3]. Дифракционные оптические нейронные сети (D2NN), предложенные в 2018 г. [4], реализуют вычисления непосредственно в процессе распространения света через оптические элементы.

Цель работы – разработка и оценка эффективности D2NN для классификации изображений с учётом физической реализуемости.

1. Теоретический анализ физических принципов

В основе D2NN лежит скалярная теория дифракции. Оптическое поле описывается комплексной функцией:

$$U(x,y,z) = A(x,y,z) \cdot e^{i\phi(x,y,z)}. \quad (1)$$

В приближении Френеля для достаточно больших расстояний дифракционное поле описывается интегралом:

$$U(x,y,z) = \frac{e^{ikz}}{i\lambda z} \iint U_0(x_0,y_0) e^{\frac{ik}{2z}[(x-x_0)^2+(y-y_0)^2]} dx_0 dy_0 . \quad (2)$$

В частотной области распространение поля представляется в виде:

$$U(f_x,f_y,z) = U_0(f_x,f_y) \cdot H(f_x,f_y), \quad (3)$$

где функция передачи системы:

$$H(f_x,f_y) = e^{ikz} \cdot e^{-i\pi\lambda z(f_x^2+f_y^2)}. \quad (4)$$

Анализ существующих решений показал ограничения по точности классификации (91-94% на MNIST) и сложности физической реализации из-за большого размера слоёв [5].

2. Математическая модель и архитектура сети

Архитектура сети состоит из L последовательных слоёв, каждый из которых включает фазовый модулятор и дифракционное распространение:

$$U_{l+1} = D\big(P(U_l)\big), \tag{5}$$

где P – фазовая модуляция, D – дифракционное распространение.

Для решения задачи классификации MNIST разработана детекторная система из 10 детекторов, расположенных по окружности на выходном слое.

Таблица 1 – Параметры разработанной модели

Параметр	**Обозначение**	**Значение**	**Физический смысл**
Длина волны	λ	532 нм	Источник излучения
Размер пикселя	Δ	8 мкм	Разрешение SLM/детектора
Расстояние между слоями	d	5 см	База оптической системы
Размер слоя	N	100×100	Апертура дифракционного элемента
Число слоев	L	10	Глубина оптической обработки

3. Реализация и вычислительные эксперименты

Прототип реализован на Python (PyTorch, NumPy) с классами FresnelPropagator, PhaseModulator и OpticalNeuralNetwork. Обучение проводилось на MNIST (60 000 изображений) в течение 150 эпох (batch size=32). Достигнута точность 89,7% на тестовой выборке. По классам: максимальная для 0 (99,1%) и 1 (99,0%), минимальная – для 2 (95,2%).

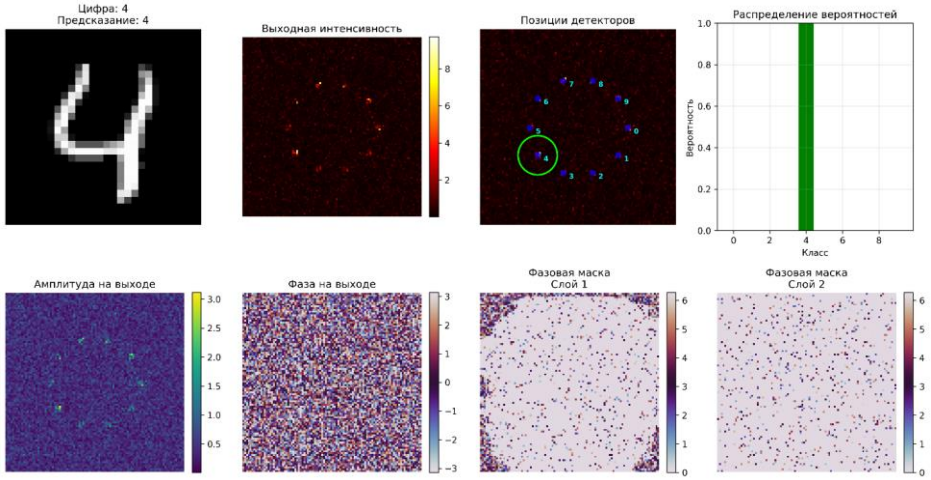

Рисунок 1 – Визуализация работы оптической нейронной сети на примере классификации цифры «4»: слева направо – исходное входное изображение, распределение интенсивности на выходном слое с расположением детекторов для 10 классов, гистограмма распределения вероятностей по классам; в нижней части – амплитуда и фаза выходного оптического поля, а также фазовые маски первого и второго обучаемых слоёв

Максимальная интенсивность сосредоточена в детекторе класса «4» (вероятность 99,8%). Фазовые маски эволюционируют от случайных паттернов к структурированным элементам. Матрица ошибок показывает путаницу между схожими цифрами: 2 и 4, 5 и 3, 9 и 4. Изначально модель без учёта юстировки показывала ~97,7% точности. Для обеспечения физической реализуемости параметры оптимизированы под толерантность к микросмещениям и угловым отклонениям при монтаже. В многолинзовой схеме даже малые сдвиги накапливают фазовые ошибки, снижая точность. Это подтверждает целесообразность использования архитектур с 1–2 линзами или однолинзовых многофокальных систем для практических задач.

4. Оценка физической реализуемости

Проверка критерия Френеля показала, что расстояние между слоями (5 см) удовлетворяет условию физической реализуемости. Теоретическая оценка разрешающей способности по теореме Аббе дала значение 33 мкм, практическая – 40 мкм.

Таблица 2 – Сравнение характеристик D2NN

Параметр	**D2NN (наша)**	**D2NN [4]**	**CNN (LeNet-5)**
Точность	89,7%	91,7%	99,2%
Число слоев	10	5	5
Размер слоя	100×100	200×200	—
Энергия/инференс	~1 мкДж	~2 мкДж	~100 мДж
Время инференса	~8 мс	~5 мс	~10 мс

Низкое энергопотребление (~1 мкДж) обусловлено пассивной природой вычислений: информация обрабатывается при свободном распространении когерентного света через статические фазовые маски без затрат на переключение транзисторов или передачу данных. В цифровых CNN основной расход энергии идёт на тактирование, активную коммутацию и работу памяти. Разница в два порядка физически обоснована и подтверждает перспективность D2NN для энергоограниченных систем.

Заключение

В ходе исследования получены следующие результаты:

1) проведён теоретический анализ физических принципов дифракции и фазовой модуляции в контексте оптических нейронных сетей;

2) разработана математическая модель дифракционной оптической нейронной сети с параметрами, обеспечивающими баланс между точностью и физической реализуемостью;

3) реализован программный прототип на Python с использованием PyTorch и NumPy;

4) достигнута точность классификации 89,7% на наборе данных MNIST;

5) подтверждена энергоэффективность системы (~1 мкДж на операцию).

Перспективными направлениями развития являются гибридные архитектуры D2NN + цифровая постобработка, использование поляризационной модуляции и интеграция с метаповерхностями.

Список использованных источников

1. Chen, H. All-optical machine learning using diffractive deep neural networks / H. Chen, J. Feng, M. Jiang // Engineering. – 2021. – V. 7, № 11. – P. 1483–1491.

2. Thompson, N. C. The energy footprint of ICT: A systematic review / N. C. Thompson, K. Kristensen, S. Boot // Renewable and Sustainable Energy Reviews. – 2022. – V. 169. – P. 112848.

3. Feldmann, J. Parallel convolutional processing using an integrated photonic tensor core / J. Feldmann, N. Youngblood, M. Karpov // Nature. – 2021. – V. 589. – P. 52–58.

4. Lin, X. All-optical machine learning using diffractive deep neural networks / X. Lin, Y. Rivenson, N. T. Yardimci // Science. – 2018. – V. 361. – P. 1004–1008.

5. Wu, J. Multi-wavelength diffractive deep neural networks / J. Wu, T. Zhou, L. Fang // Optics Express. – 2020. – V. 28, № 15. – P. 22312–22325.

6. Optical classification of images at different wavelengths using spectral diffractive neural networks / G. A. Motz, D. V. Soshnikov, L. L. Doskolovich, E. V. Byzov, E. A. Bezus, D. A. Bykov // Computer Optics. – 2025. – V. 49, № 2. – P. 187–199.

7. Design of Diffractive Neural Networks for Solving Different Classification Problems at Different Wavelengths / G. A. Motz, L. L. Doskolovich, D. V. Soshnikov, E. V. Byzov, E. A. Bezus, N. V. Golovastikov, D. A. Bykov // Photonics. – 2024. – T. 11, № 8. – C. 780.

MODELLING OF A PRIMITIVE MULTILAYER OPTICAL NEURAL NETWORK FOR DIGIT RECOGNITION

V.N. Krechko[1]

[1]Samara University, Samara
e-mail: krechkovsevolod@mail.ru

Abstract

The paper presents research results on diffractive optical neural networks (D2NN) for image classification. A mathematical model of a multilayer optical network based on Fresnel diffraction and phase modulation is developed. A software prototype was implemented in Python using PyTorch. A classification accuracy of 89.7% was achieved on the MNIST dataset with energy consumption of ~1 μJ per operation, which is two orders of magnitude lower than digital analogs.

Keywords: diffractive optical neural network, Fresnel diffraction, phase modulation, machine learning, optical computing.