Научная статья
УДК 004.932.2:535.42
EDN SQQGDD
DOI 10.17150/2713-1734.2026.8(3).301-312

А.М. Абалмасов
Воронежский государственный университет,
г. Воронеж, Российская Федерация
П.А. Головинский
Воронежский государственный университет,
г. Воронеж, Российская Федерация

Сходимость обучения в оптическом резервуарном компьютере с дифракционной сверткой и случайной фазовой маской

Аннотация. В работе даётся строгое теоретическое обоснование метода классификации изображений, использующего гибридный оптико-электронный подход на основе резервуарных вычислений. В отличие от многочисленных экспериментальных работ, демонстрирующих эффективность оптических резервуаров на случайных рассеивателях, до сих пор отсутствовало математически строгое объяснение причин, по которым дифракционная свертка на шероховатой поверхности порождает состояния, пригодные для линейной классификации. Предлагаемое обоснование опирается на представление о случайном проектировании в пространство высокой размерности. Показано, что отображение входного изображения в спекл-структуру интенсивности эквивалентно применению нелинейного случайного признакового отображения. Доказана теорема, согласно которой почти для всех реализаций случайной фазовой маски и при достаточном количестве детекторов ($N \geq M$) матрица состояний резервуара, составленная из M спекл-структур различных входов, имеет полный ранг с вероятностью 1. Это гарантирует возможность достижения сколь угодно малой ошибки обучения при использовании линейного классификатора с регуляризацией Тихонова. Результаты создают теоретическую основу для проектирования оптических резервуарных систем с предсказуемыми свойствами. Численное моделирование на базе лиц ORL подтверждает теоретические выводы. При полном ранге матрицы состояний (140 из 140) точность классификации на тестовой выборке достигает 95 %. Предложенный подход может быть распространён на другие задачи машинного обучения, где требуется быстрая обработка визуальной информации с низким энергопотреблением.

Ключевые слова. Резервуарные вычисления, оптический резервуарный компьютер, дифракция Френеля, случайная фазовая маска, гребневая регрессия, полный ранг, линейная классификация, распознавание лиц, спекл-структура, случайные проекции.

Информация о статье. Дата поступления: 8 июня 2026 г.; дата принятия к публикации: 13 октября 2026 г.; дата онлайн-размещения: 11 ноября 2026 г.

Для цитирования. Абалмасов А.М. Сходимость обучения в оптическом резервуарном компьютере с дифракционной свёрткой и случайной фазовой маской / А.М. Абалмасов, П.А. Головинский. — DOI 10.17150/2713-1734.2026.8(3).301-312. — EDN SQQGDD // System Analysis & Mathematical Modeling. — 2026. — Т. 8, № 3. — С. 301–312.

© Абалмасов А.М., Головинский П.А., 2026

ISSN 2713-1734

Original article

A.M. Abalmasov
Voronezh State University,
Voronezh, Russian Federation

P.A. Golovinsky
Voronezh State University,
Voronezh, Russian Federation

Training Convergence in an Optical Reservoir Computer with Diffractive Convolution and a Random Phase Mask

Abstract. This paper provides a rigorous theoretical justification for an image classification method employing a hybrid optoelectronic approach based on reservoir computing. Unlike numerous experimental studies demonstrating the effectiveness of optical reservoirs with random scatterers, a mathematically rigorous explanation of why diffractive convolution on a rough surface generates states suitable for linear classification has been lacking to date. The proposed justification relies on the concept of random projection into a high-dimensional space. It is shown that mapping an input image to an intensity speckle pattern is equivalent to applying a nonlinear random feature map. A theorem is proved, showing that for almost all realizations of a random phase mask and with a sufficient number of detectors ($N \geq M$), the reservoir state matrix composed of M speckle patterns from different inputs has full rank with probability 1. This guarantees achieving an arbitrarily small training error when using a linear classifier with Tikhonov regularization. The results provide a theoretical foundation for designing optical reservoir systems with predictable properties. Numerical simulation on the ORL face database confirms the theoretical findings. With a full-rank state matrix (140 out of 140), the classification accuracy on the test set reaches 95 %. The proposed approach can be extended to other machine learning tasks requiring fast, low-power visual information processing.

Keywords. Reservoir computing, optical reservoir computer, Fresnel diffraction, random phase mask, ridge regression, full rank, linear classification, face recognition, speckle pattern, random projections.

Article info. Received 8 June, 2026; Accepted 10 October, 2026; Available online 11 November, 2026.

For Citation. Abalmasov A.M., Golovinsky P.A. Training Convergence in an Optical Reservoir Computer with Diffractive Convolution and a Random Phase Mask. *System Analysis & Mathematical Modeling*, 2026, vol. 8, no. 3, pp. 301–312. (In Russian). EDN: SQQGDD. DOI: 10.17150/2713-1734.2026.8(3).301-312.

Введение

Распознавание визуальных образов остаётся ключевой задачей компьютерного зрения. Глубокие нейронные сети требуют значительных вычислительных ресурсов, что стимулирует развитие гибридных оптико-электронных систем [1–3]. Особый интерес представляют оптические резервуарные компьютеры (ОРК), реализующие парадигму резервуарных вычислений (Reservoir Computing, РК) [4]. В таких системах входной сигнал преобразуется фиксированным резервуаром (многомодовое волокно, рассеивающая среда, шероховатая поверхность), после чего обучается

только линейный выходной слой [3]. Несмотря на экспериментальные успехи, теоретическое понимание того, почему дифракция на случайном экране порождает состояния с достаточной разделяющей способностью, остаётся неполным. В большинстве работ [1–3] авторы ограничиваются демонстрацией работоспособности без строгого обоснования полного ранга матрицы состояний. Настоящая работа восполняет этот пробел, предлагая строгую теорию, в которой отображение входного изображения в спекл-структуру интерпретируется как случайное нелинейное проектирование в пространство высокой размерности. Доказывается, что для почти всех реализаций фазовой маски векторы состояний различных входов линейно независимы при числе детекторов, не меньшем числа обучающих примеров, что обеспечивает сходимость обучения. Теоретические основы построения подобных гибридных систем, сочетающих оптические методы обработки и алгоритмы экстремального обучения, детально рассмотрены в работе [5].

Теоретические основы и описание метода

Основу оптической части предлагаемой системы составляет дифракционное распространение когерентного монохроматического поля. Обозначим через (ξ, η) декартовы координаты в плоскости входного поля (плоскость $z = 0$), а через (x, y) — координаты в плоскости наблюдения на расстоянии z от входной плоскости. В скалярном приближении распространение поля $E(\xi, \eta)$ на расстояние z описывается интегралом Френеля [6]:

$$E(x,y;z) = \frac{e^{ikz}}{i\lambda z} \iint E(\xi,\eta) \exp\left\{\frac{ik}{2z}[(x-\xi)^2 + (y-\eta)^2]\right\} d\xi\, d\eta, \quad (1)$$

где $k = 2\pi/\lambda$ — волновое число, λ — длина волны, интегрирование ведётся по всей апертуре входного поля. Интеграл (1) представляет собой двумерную свёртку входного поля $E(\xi, \eta)$ с ядром Френеля:

$$h_z(x,y) = \frac{e^{ikz}}{i\lambda z} \exp\left(\frac{ik}{2z}(x^2 + y^2)\right), \quad (2)$$

так что $E(x, y, z) = E * h_z$. Свёрточная структура обеспечивает нелокальное перемешивание информации: каждая точка выходной плоскости зависит от всех точек входного поля. Входные изображения могут предварительно обрабатываться в стандартной 4f-системе [6] для пространственной фильтрации (выделения контуров и т.д.).

Центральным элементом резервуара является шероховатая поверхность, моделируемая случайным фазовым экраном (random phase screen), т.е. тонким экраном, модулирующим только фазу поля.

Физическое обоснование такой модели основано на разделении двух масштабов. Рассмотрим прозрачную пластину с показателем преломления n и переменной толщиной $d + h(\xi, \eta)$, где d — средняя толщина, а $h(\xi, \eta)$ — флуктуации высоты поверхности (профиль шероховатости). Оптическая длина пути через такую пластину равна $n[d + h(\xi, \eta)]$, и приобретаемая фаза составляет $kn[d + h(\xi, \eta)]$. Относительно плоского волнового фронта, прошедшего тот же путь в вакууме, набег фазы равен

$$\phi(\xi, \eta) = \frac{2\pi}{\lambda}(n - 1)\, h(\xi, \eta). \tag{3}$$

Ситуация с модулируемой фазой принципиально иная. Даже при относительно слабой шероховатости со среднеквадратичным отклонением высоты σ_h фазовые флуктуации (3) имеют среднеквадратичное отклонение $\sigma_\phi = 2\pi(n - 1)\, \sigma_h / \lambda$. При $n = 1{,}5$, $\lambda = 0{,}5$ мкм и $\sigma_h \gtrsim {}_1$ мкм отклонение фазы превышает 2π, и фаза оборачивается многократно, покрывая весь интервал $[0,2\pi)$. При отражении набег удваивается: $\phi = 4\pi h(\xi, \eta) / \lambda$.

Таким образом, математическая модель случайного фазового экрана формулируется следующим образом. Поле после экрана:

$$E_{\text{mod}}(\xi, \eta) = E_{\text{in}}(\xi, \eta)\, e^{i\phi(\xi, \eta)}, \tag{4}$$

где $\phi(\xi, \eta)$ — случайное поле, значения которого в дискретизированных точках экрана являются независимыми случайными величинами с абсолютно непрерывным распределением (в модели сильной шероховатости — равномерным на $[0,2\pi)$). Экран не изменяет амплитуду: $|E_{\text{mod}}| = |E_{\text{in}}|$. Случайность фазы обеспечивает нелинейное преобразование входного сигнала, необходимое для разделения классов.

Предлагаемый метод распознавания состоит из трёх последовательных этапов, реализуемых в гибридной оптико-электронной системе, схема которой показана на рис. 1. Пример реальных сигналов на каждом этапе приведён на рис. 2.

Этап 1. Предобработка в 4*f*-системе. Изображение лица, записанное в интенсивности $I_{\text{face}}(x, y)$, реализуется на амплитудном транспаранте и помещается во входную плоскость 4*f*-системы. Комплексная амплитуда на выходе транспаранта пропорциональна $\sqrt{I_{\text{face}}(x, y)}$. В фурье-плоскости системы устанавливается пространственный фильтр $H_{\text{edge}}(u, v)$, аппроксимирующий оператор Лапласа. На выходе 4*f*-системы формируется поле контуров

$$E_{\text{edge}}(x', y') = \mathcal{F}^{-1}\left\{\mathcal{F}\left\{\sqrt{I_{\text{face}}}\right\} \cdot H_{\text{edge}}\right\}(x', y'), \tag{5}$$

Этап 2. Дифракция на случайном фазовом резервуаре. Поле контуров E_{edge} освещает шероховатую поверхность (случайный фазовый экран). Согласно (4), после экрана поле приобретает случайную фазовую модуляцию: $E_{\text{mod}}(\xi, \eta) = E_{\text{edge}}(\xi, \eta)\exp[i\phi(\xi, \eta)]$. Далее поле распространяется на расстояние z в соответствии с интегралом Френеля (1), и в плоскости детектора регистрируется распределение интенсивности в виде спекл-структуры:

$$I_{\text{res}}(X, Y) = |(E_{\text{mod}} * h_z)(x, y)|^2. \tag{6}$$

Этап 3. Обучение и классификация. Интенсивность I_{res} растягивается в вектор $\boldsymbol{r} \in \mathbb{R}^N$ (N — число пикселей детектора). Для M обучающих изображений получаем матрицу состояний $\boldsymbol{R} = [\boldsymbol{r}_1, ..., \boldsymbol{r}_M] \in \mathbb{R}^{N \times M}$ и матрицу меток $\boldsymbol{Y} \in \mathbb{R}^{C \times M}$, где C — число классов (каждый столбец $\boldsymbol{Y}$ кодирует принадлежность соответствующего изображения к одному из C классов). Линейный классификатор $\boldsymbol{W} \in \mathbb{R}^{C \times N}$ ищется из условия минимизации функционала с регуляризацией Тихонова:

$$\boldsymbol{W}_{\text{out}} = \underset{\boldsymbol{W}}{\operatorname{argmin}}(\| \boldsymbol{WR} - \boldsymbol{Y} \|_F^2 + \beta \| \boldsymbol{W} \|_F^2), \tag{7}$$

где $\beta > 0$. Решение имеет вид

$$\boldsymbol{W}_{out} = \boldsymbol{YR}^T(\boldsymbol{RR}^T + \beta \boldsymbol{I})^{-1}. \tag{8}$$

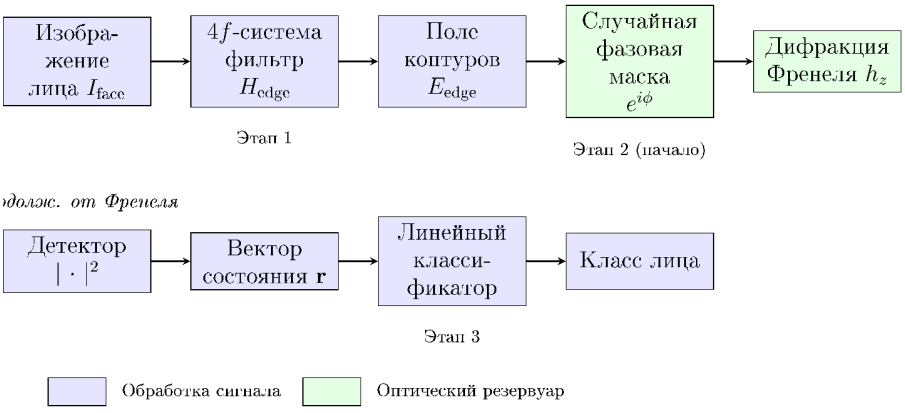


Рис. 1. Схема оптического резервуарного компьютера для распознавания лиц. Верхняя часть показывает формирование дифрагированного поля, нижняя — детектирование и линейную классификацию

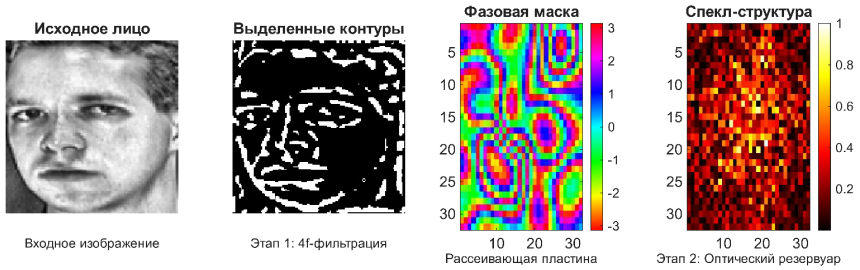


Рис. 2. Последовательные преобразования реального изображения лица из базы ORL в оптическом резервуарном компьютере. Слева направо: исходное изображение; контурное изображение после 4f-фильтрации (этап 1); фазовая маска рассеивающей пластины $e^{i\phi(\xi,\eta)}$; спекл-структура интенсивности I_{res} в плоскости детектора (этап 2)

Теоретическое обоснование

В стандартных резервуарных вычислениях эффективность достигается за счёт того, что фиксированная нелинейная динамическая система отображает входной сигнал в пространство состояний высокой размерности, где различные классы становятся линейно разделимыми. Для нашего резервуара отображение складывается из фазового умножения $\exp(i\phi\,(\xi,\eta))$, дифракционной свертки с ядром Френеля h_z (2) и квадратичного детектирования $|\cdot|^2$. Итоговое отображение $\mathcal{T}: I_{\text{edge}} \mapsto \boldsymbol{r}$ нелинейно и случайно; каждая компонента вектора $\boldsymbol{r}$ есть нелинейная функция входных данных со случайными параметрами ϕ). Это напоминает метод случайных признаков (random features) для аппроксимации ядер [7] или случайные нейронные сети [4]. Известно, что такие отображения с высокой вероятностью сохраняют разделимость классов, если размерность выходного пространства достаточно велика (теорема Ковера о линейной разделимости [8, 9]). Однако требуется доказать более сильное свойство — линейную независимость векторов состояний для различных входов, обеспечивающую достижение нулевой ошибки обучения при $\beta \to 0$. Если каждый вектор $\boldsymbol{r}_i$ получается применением одной случайной нелинейной функции к различным аргументам, то при $N \geq M$ эти векторы линейно независимы для почти всех реализаций случайности.

Для математического анализа необходимо перейти от непрерывных полей раздела 2 к конечномерным объектам. Предположим, что все поля дискретизированы на прямоугольной сетке. Соответствие между непрерывными (физическими) и дискретными (математическими) обозначениями устанавливается следующим образом:

– поле контуров $E_{\text{edge}}(\xi, \eta)$ (5), дискретизированное на сетке из P точек в плоскости маски, представляется вектором $\boldsymbol{e} \in \mathbb{C}^P$;

– случайная фаза $\phi(\xi, \eta)$ (3) представляется вектором $\boldsymbol{\phi} \in \mathbb{R}^P$, компоненты которого ϕ_j ($j = 1, ..., P$) являются независимыми случайными величинами с абсолютно непрерывным распределением (например, равномерным на $[0, 2\pi)$);

– умножение поля на фазовую маску $E_{\text{mod}} = E_{\text{in}} e^{i\phi}$ (4) записывается как $\boldsymbol{e}_{\text{mod}} = \boldsymbol{e} \odot e^{i\boldsymbol{\phi}}$, где $\odot$ обозначает покомпонентное (адамарово) произведение векторов;

– свёртка с ядром Френеля h_z (2) является линейным оператором и представляется матрицей $L \in \mathbb{C}^{N \times P}$, где N — число пикселей детектора (узлов сетки в плоскости наблюдения), а элемент L_{nj} есть дискретизация значения ядра $h_z(x_n - \xi_j, y_n - \eta_j)$; матрица L определяется исключительно длиной волны λ и расстоянием z и не зависит от $\boldsymbol{\phi}$;

– квадратичное детектирование $I_{\text{res}} = |E_{\text{mod}} * h_z|^2$ (6) записывается как $\boldsymbol{r} = |L(\boldsymbol{e} \odot e^{i\boldsymbol{\phi}})|^2 \in \mathbb{R}^N$, где квадрат модуля берётся покомпонентно.

Обозначим через $T(\boldsymbol{e}; \boldsymbol{\phi}) = |L(\boldsymbol{e} \odot e^{i\boldsymbol{\phi}})|^2$ полное отображение входного поля в вектор состояния детектора. Для M обучающих изображений с попарно различными входными полями $\boldsymbol{e}_1, ..., \boldsymbol{e}_M$ нам нужно исследовать матрицу состояний $R = [T(\boldsymbol{e}_1; \boldsymbol{\phi}), ..., T(\boldsymbol{e}_M; \boldsymbol{\phi})] \in \mathbb{R}^{N \times M}$.

Лемма 1. (Линейная независимость спекл-структур). Пусть $\boldsymbol{e}_1, ..., \boldsymbol{e}_M \in \mathbb{C}^P$ есть ненулевые и попарно различные комплексные векторы. Рассмотрим отображение $T(\boldsymbol{e}; \boldsymbol{\phi}) = |L(\boldsymbol{e} \odot e^{i\boldsymbol{\phi}})|^2$, где $L \in \mathbb{C}^{N \times P}$ — матрица полного ранга $\min(N, P)$, а $\boldsymbol{\phi} \in \mathbb{R}^P$ — вектор независимых непрерывных случайных величин, имеющих абсолютно непрерывное распределение (например, равномерное на $[0, 2\pi)$). Тогда с вероятностью 1 (относительно распределения $\boldsymbol{\phi}$) векторы $\boldsymbol{r}_1, ..., \boldsymbol{r}_M \in \mathbb{R}^N$ линейно независимы, если $N \geq M$.

Доказательство. Рассмотрим матрицу $\boldsymbol{R} \in \mathbb{R}^{N \times M}$, столбцами которой являются векторы $\boldsymbol{r}_i = T(\boldsymbol{e}_i; \boldsymbol{\phi})$, $i = 1, ..., M$. Условие линейной независимости столбцов эквивалентно тому, что матрица Грамма $\boldsymbol{G} = \boldsymbol{R}^T\boldsymbol{R} \in \mathbb{R}^{M \times M}$ имеет ненулевой определитель: $\det(\boldsymbol{G}) \neq 0$. Каждый элемент $\boldsymbol{G}_{ij} = \boldsymbol{r}_i^T \boldsymbol{r}_j$ является функцией от $\boldsymbol{\phi}$. Выразим $\boldsymbol{r}_i$ явно. Введём обозначение $\boldsymbol{w} = \boldsymbol{e} \odot e^{i\boldsymbol{\phi}} \in \mathbb{C}^P$, т.е. $w_j = e_j e^{i\phi_j}$ для $j = 1, ..., P$. Тогда n-я компонента вектора состояния ($n = 1, ..., N$) равна

$$r_n = |(L\boldsymbol{w})_n|^2 = (L\boldsymbol{w})_n \overline{(L\boldsymbol{w})_n} = \sum_{j=1}^{P} \sum_{l=1}^{P} L_{nj}\, \overline{L_{nl}}\, w_j \overline{w_l},$$

где L_{nj} — элемент матрицы L в строке n и столбце j. После подстановки $w_j = e_j e^{i\phi_j}$, $\overline{w_l} = \overline{e_l}\, e^{-i\phi_l}$ получаем

$$r_n(\boldsymbol{e}; \boldsymbol{\phi}) = \sum_{j=1}^{P} \sum_{l=1}^{P} L_{nj}\, \overline{L_{nl}}\, e_j \overline{e_l}\, e^{i(\phi_j - \phi_l)}. \tag{9}$$

Таким образом, каждый вектор $\boldsymbol{r}_i$ является линейной комбинацией (с коэффициентами $L_{nj}\,\overline{L_{nl}}\,(\boldsymbol{e}_i)_j\,\overline{(\boldsymbol{e}_i)_l}$, зависящими от входов $\boldsymbol{e}_i$ и матрицы L) комплексных экспонент $e^{i(\phi_j-\phi_l)}$, $j, l = 1, ..., P$.

Рассмотрим $\det(\boldsymbol{G})$ как функцию переменных $\boldsymbol{\phi}$. Из (9) видно, что каждый элемент $\boldsymbol{G}_{ij} = \sum_{n=1}^{N} r_n(\boldsymbol{e}_i; \boldsymbol{\phi}) r_n(\boldsymbol{e}_j; \boldsymbol{\phi})$ является конечной суммой членов вида $c \cdot e^{i(\phi_{j_1} - \phi_{l_1} + \phi_{j_2} - \phi_{l_2})}$ с числовыми коэффициентами c. Следовательно, $\det(\boldsymbol{G})$ представляет собой тригонометрический полином от переменных $\phi_1, ..., \phi_P$ — конечную линейную комбинацию произведений $e^{im_s\phi_s}$ с целочисленными показателями m_s. Введём подстановку $\zeta_s = e^{i\phi_s}$, $s = 1, ..., P$; тогда $\det(\boldsymbol{G})$ становится полиномом Лорана от комплексных переменных $\zeta_1, ..., \zeta_P$, ограниченным на тор $\mathbb{T}^P = \{|\zeta_s| = 1\}$.

Важнейшее наблюдение состоит в следующем. Если $\det(\boldsymbol{G})$ не является тождественным нулём как функция на $\mathbb{T}^P$, то множество его нулей $\{\boldsymbol{\phi} \in [0, 2\pi)^P : \det(\boldsymbol{G}(\boldsymbol{\phi})) = 0\}$ имеет лебегову меру нуль. Действительно, нетождественный полином Лорана от P комплексных переменных обращается в нуль на множестве комплексной коразмерности не менее 1. Его ограничение на компактный тор $\mathbb{T}^P$ является вещественно-аналитической функцией, множество нулей которой имеет меру нуль по теореме Лойясиевича [10; 11] (удобную формулировку см. в [11, теорема 4.5.1]).

Остаётся доказать, что $\det(\boldsymbol{G})$ не равен тождественному нулю, т.е. существует хотя бы одно значение $\boldsymbol{\phi}$, для которого $\det(\boldsymbol{G}) > 0$. Выберем $\boldsymbol{\phi} = \boldsymbol{0}$ (нулевая фаза). Тогда $e^{i\phi j} = 1$ для всех j, и отображение упрощается: $\boldsymbol{r}_i = |L\boldsymbol{e}_i|^2$. Покажем, что можно подобрать входы $\boldsymbol{e}_i$ так, что соответствующие $\boldsymbol{r}_i$ линейно независимы. Поскольку L имеет ранг $\min(N, P)$ и $N \geq M$, существуют M строк матрицы L, образующих матрицу $L' \in \mathbb{C}^{M \times P}$ ранга M. Выбрав $\boldsymbol{e}_i$ с непересекающимися носителями (т. е. $(\boldsymbol{e}_i)_j \neq 0$ только при $j \in S_i$, где $S_i \cap S_k = \emptyset$ при $i \neq k$), получим, что компоненты $|L\boldsymbol{e}_i|^2$ и $|L\boldsymbol{e}_k|^2$ зависят от непересекающихся множеств входных значений, что обеспечивает их линейную независимость при подходящем выборе амплитуд. (Конкретный пример: если первые M столбцов L линейно независимы, положим $(\boldsymbol{e}_i)_i = \delta_{ij}$; тогда $r_n(\boldsymbol{e}_i) = |L_{ni}|^2$, и столбцы матрицы $[|L_{ni}|^2]_{n,\,i}$ линейно независимы при условии, что столбцы L попарно различны, что выполнено для ядра Френеля при различных пространственных точках.) Итак, $\det(\boldsymbol{G}(\boldsymbol{0})) > 0$, что означает нетождественность $\det(\boldsymbol{G})$ нулю. Следовательно, $\det(\boldsymbol{G}(\boldsymbol{\phi})) \neq 0$ для почти всех $\boldsymbol{\phi} \in [0, 2\pi)^P$, что и означает линейную независимость столбцов $\boldsymbol{R}$ с вероятностью 1.

Замечание 1. Лемма 1 утверждает, что случайная фазовая маска с вероятностью 1 обеспечивает максимально возможный ранг матрицы состояний (при условии $N \geq M$). Это свойство не зависит от конкретных значений входов, если только они различны. На практике N может быть очень большим (миллионы пикселей), и

условие $N \geq M$ выполняется почти всегда для разумных размеров обучающей выборки.

Из леммы 1 и стандартной теории гребневой регрессии немедленно вытекает, что для любого набора меток Y можно достичь сколь угодно малой ошибки на обучающих данных, выбрав параметр регуляризации β достаточно малым (см., например, [9]). Этот вывод не зависит от природы матрицы $\boldsymbol{R}$ при условии её полного столбцового ранга.

Численная проверка теоретических результатов

Для численной проверки леммы 1 проведено моделирование на базе лиц ORL[1]: 20 классов, $M = 140$ обучающих образцов (по 7 на класс), 60 тестовых. Размер резервуара $N = 1024$ (32 × 32), длина волны $\lambda = 532$ нм, расстояние до детектора $L = 0{,}05$ м; дифракционная свёртка вычислялась через алгоритм быстрого преобразования Фурье (БПФ). Программная реализация выполнена на языке MATLAB[2]. Проверка леммы 1 подтвердила $\mathrm{rank}(\boldsymbol{R}) = M = 140$. Все сингулярные числа матрицы состояний превышают порог машинной точности на 11 порядков.

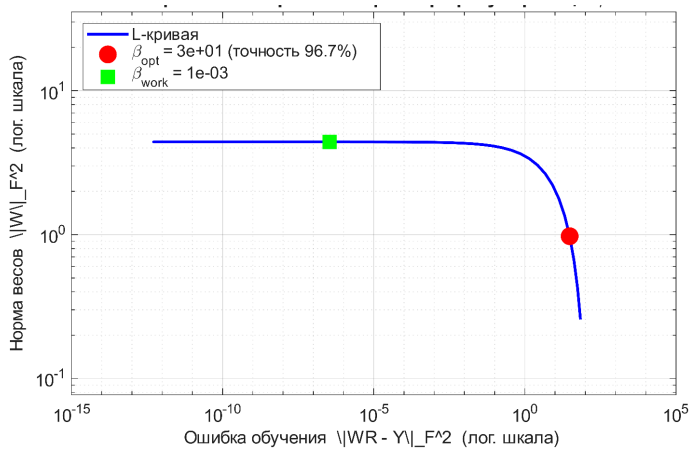


Рис. 3. L-кривая: зависимость нормы весов $\|W_\beta\|_F^2$ от ошибки обучения $\|W_\beta R - Y\|_F^2$ в двойной логарифмической шкале при $\beta \in [10^{-8}, 10^2]$. Красная точка — оптимальное значение $\beta^* = 30$ (точность на тесте 96,7 %); зелёный квадрат — рабочее значение $\beta = 10^{-3}$ (95 %), расположенное на горизонтальном плато

[1] База лиц ORL (AT&T Laboratories Cambridge, 1994) содержит 400 изображений 40 человек. В эксперименте использовано 200 изображений: 140 для обучения, 60 для теста. Доступна по адресу: https://www.kaggle.com/datasets/tavarez/the-orl-database-for-training-and-testing

[2] Исходный код доступен в репозитории: https://gitlab.com/reservoir-computer-for-face-recognition-based-on-fourier-optics-and-speckle-patterns/Convergence-of-training.

Параметр регуляризации β задаёт компромисс между ошибкой обучения и нормой весов. Для иллюстрации этой зависимости β варьировался в диапазоне $[10^{-8}, 10^{2}]$. Результат представлен в виде L-кривой (рис. 3).

L-кривая воспроизводит теоретический компромисс между ошибкой обучения и нормой весов. Уменьшение β снижает ошибку обучения (горизонтальное плато), тогда как норма весов монотонно возрастает (вертикальный спуск). Значение $\beta = 10^{-3}$ даёт точность 95 %. Оптимальное значение $\beta^* = 30$ находится вблизи изгиба кривой и даёт 96,7 %, что соответствует классическому критерию выбора по L-кривой. Небольшой прирост точности при увеличении β обусловлен подавлением переобучения при $M = 140 << N = 1024$. Численный эксперимент подтверждает теоретические выводы.

Заключение

В предложенном обосновании ключевую роль играет линейный оператор L (свертка с ядром Френеля (2)), который обладает свойством нелокальности: каждая точка выходной плоскости зависит от всех точек входного изображения. Это обеспечивает смешивание информации, необходимое для чувствительности отображения T к глобальной структуре образа. При тождественном отображении ($L = I$) вектор состояния $\boldsymbol{r}$ был бы покомпонентным квадратом модуля входного поля, и различные входы могли бы давать линейно зависимые образы. Строгое доказательство показывает, что оптические резервуары на шероховатых поверхностях обладают универсальной аппроксимирующей способностью. В отличие от стандартных резервуарных вычислений, где требуются динамические системы с эхо-состоянием, в нашем случае статическая нелинейность (квадрат модуля) и случайное проектирование обеспечивают аналогичные свойства. Это позволяет интерпретировать ОРК как частный случай метода «случайных признаков» (random features) для ядра, порождаемого генератором T. Из доказанной леммы 1 следует, что для достижения высокой точности классификации достаточно обеспечить $N \geq M$ и использовать гребневую регрессию (8). Параметр β балансирует между ошибкой обучения (малое β) и устойчивостью к шумам (большое β). Параметр β можно выбирать с помощью L-кривой или кросс-валидации.

Основной результат работы — доказательство того, что с вероятностью 1 матрица состояний резервуара имеет полный ранг при $N \geq M$, что гарантирует сходимость обучения линейного классификатора. Результаты создают теоретический фундамент для развития гибридных оптико-электронных систем машинного обучения.

Благодарности

Авторы благодарят А.А. Сироту за поддержку и интерес, проявленный к работе.

Список использованной литературы

1. Nakajima M. Scalable reservoir computing on coherent linear photonic processor / M. Nakajima, K. Tanaka, T. Hashimoto // Communications Physics. — 2021. — Vol. 4. — P. 20.

2. Experimental demonstration of reservoir computing on a silicon photonics chip / K. Vandoorne, P. Mechet, Van Vaerenbergh T. [et al.] // Nature Communications. — 2014. — Vol. 5. — P. 3541.

3. Human action recognition with a large-scale brain-inspired photonic computer / P. Antonik, N. Marsal, D. Brunner, D. Rontani // Nature Machine Intelligence. — 2019. — Vol. 1. — P. 530–537.

4. Jaeger H. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication / H. Jaeger, H. Haas // Science. — 2004. — Vol. 304. — P. 78–80.

5. Головинский П.А. Дорожная карта реализации оптических машин экстремального обучения / П.А. Головинский. — DOI 10.17150/2713-1734.2023.5(3).239-250. — EDN MDTJBC // System Analysis and Mathematical Modeling. — 2023. — Т. 5, № 3. — С. 239–250.

6. Goodman J.W. Introduction to Fourier Optics / J.W. Goodman. — 3rd ed. — Englewood : Roberts & Company Publishers, 2005. — 509 p.

7. Rahimi A. Random features for large-scale kernel machines / A. Rahimi, B. Recht // Advances in Neural Information Processing Systems. — 2007. — P. 1177–1184.

8. Cover T.M. Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition / T.M. Cover // IEEE Transactions on Electronic Computers. — 1965. — Vol. EC-14. — P. 326–334.

9. Hastie T. The Elements of Statistical Learning / T. Hastie, R. Tibshirani, J. Friedman. — 2nd ed. — Springer, 2009. — 764 p.

10. Łojasiewicz S. Ensembles semi-analytiques / S. Łojasiewicz. — Bures-sur-Yvette, 1964. — 153 p.

11. Krantz S.G. A Primer of Real Analytic Functions / S.G. Krantz, H.R. Parks. — 2nd ed. — Birkhäuser, 2002. — 205 p.

References

1. Nakajima M., Tanaka K., Hashimoto T. Scalable Reservoir Computing on Coherent Linear Photonic Processor. *Communications Physics*, 2021, vol. 4, pp. 20.

2. Vandoorne K., Mechet P., Van Vaerenbergh T., M. Fiers, G. Morthier [et al.]. Experimental Demonstration of Reservoir Computing on a Silicon Photonics Chip. *Nature Communications,* 2014, vol. 5, pp. 3541.

3. Antonik P., Marsal N., Brunner D., Rontani D. Human Action Recognition with a Large-scale Brain-inspired Photonic Computer. *Nature Machine Intelligence*, 2019, vol. 1, pp. 530–537.

4. Jaeger H., Haas H. Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication. *Science,* 2004, vol. 304, pp. 78–80.

5. Golovinsky P.A. Roadmap for Implementation Optical Extreme Learning Machines. *System Analysis and Mathematical Modeling*, 2023, vol. 5, no. 3, pp. 239–250. (In Russian). EDN: MDTJBC. DOI: 10.17150/2713-1734.2023.5(3).239-250.

6. Goodman J.W. *Introduction to Fourier Optics*. 3rd ed. Englewood, Roberts & Company Publishers, 2005. 509 p.

7. Rahimi A., Recht B. Random Features for Large-Scale Kernel Machines. *Advances in Neural Information Processing Systems*, 2007, pp. 1177–1184.

8. Cover T.M. Geometrical and Statistical Properties of Systems of Linear Inequalities with Applications in Pattern Recognition. *IEEE Transactions on Electronic Computers,* 1965, vol. EC-14, pp. 326–334.

9. Hastie T., Tibshirani R., Friedman J. *The Elements of Statistical Learning*. 2nd ed. Springer, 2009. 764 p.

10. Łojasiewicz S. *Ensembles Semi-Analytiques*. Bures-sur-Yvette, 1964. 153 p.

11. Krantz S.G., Parks H.R. A *Primer of Real Analytic Functions*. 2nd ed. Birkhäuser, 2002. 205 p.

Информация об авторах

Абалмасов Александр Михайлович — магистр по специальности «Электроника и наноэлектроника», аспирант по специальности «Искусственный интеллект и машинное обучение», Воронежский государственный университет, 394018, Российская Федерация, г. Воронеж, Университетская пл., 1, e-mail: alexander.abalmasov@gmail.com.

Головинский Павел Абрамович — доктор физико-математических наук, профессор, кафедра теории обработки и защиты информации, кафедра оптики и спектроскопии, ведущий научный сотрудник, Воронежский государственный университет, 394018, Российская Федерация, г. Воронеж, Университетская пл., 1, e-mail: golovinski@bk.ru.

Information about the Authors

Alexander M. Abalmasov — M.Sc. in Electronics and Nanoelectronics, PhD Student in Artificial Intelligence and Machine Learning, Voronezh State University, 1 Universitetskaya Square, Voronezh, 394018, Russian Federation, e-mail: alexander.abalmasov@gmail.com.

Pavel A. Golovinsky — D.Sc. in Physics and Mathematics, Professor, Department of Information Processing and Security Technologies, Department of Optics and Spectroscopy, Leading Researcher, Voronezh State University, 1 Universitetskaya Square, Voronezh, 394018, Russian Federation, e-mail: golovinski@bk.ru.

Вклад авторов

Все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Contribution of the Authors

The authors contributed equally to this article. The authors declare no conflicts of interests.