

УДК 004.942

Ильин Иван Александрович, аспирант,
Ульяновский государственный технический университет,
г. Ульяновск

Ташлинский Александр Григорьевич,
д. технических н., профессор,
Ульяновский государственный технический университет,
г. Ульяновск

Сучков Сергей Николаевич, аспирант,
Ульяновский государственный технический университет,
г. Ульяновск

АНАЛИЗ ЭФФЕКТИВНОСТИ ОПТИМИЗАЦИИ СТОХАСТИЧЕСКИХ ПРОЦЕДУР С ИСПОЛЬЗОВАНИЕМ ВЗАИМНОЙ ИНФОРМАЦИИ РЕНЬИ В ЗАДАЧАХ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ

Аннотация: В статье рассматривается применение методов оптимизации стохастических процедур, синтезированных по критерию максимума взаимной информации Реньи, в задачах распознавания изображений. Рассмотрены теоретические основы, обзор существующих методов, а также экспериментальное сравнение методов, включая сверточные нейронные сети (CNN) и машины опорных векторов (SVM).

Ключевые слова: взаимная информация Реньи, стохастические процедуры, оптимизация, распознавание изображений, нормализация.

Введение

Классические подходы к распознаванию изображений зачастую основываются на вероятностных или обучаемых моделях, таких как байесовские классификаторы, методы опорных векторов [11] или нейронные сети [7]. Однако, несмотря на высокую эффективность, эти методы требуют значительных объемов размеченных данных и подвержены переобучению в условиях ограниченной обучающей выборки [3].

Альтернативным подходом является синтез стохастических процедур, основанных на информационно-теоретических критериях, в частности, на взаимной информации (далее ВИ). В данной работе рассматривается частный случай этой меры – ВИ Реньи, которая является обобщением классической ВИ Шеннона [10] и позволяет гибко управлять чувствительностью к вероятностным различиям [9].

Целью настоящей статьи является практическое исследование применения критерия максимума ВИ Реньи для оптимизации стохастических процедур распознавания изображений. Основной акцент сделан на сравнительный анализ эффективности данного подхода, его применимость к задачам классификации изображений и сопоставление с другими существующими методами.

2. Обзор существующих методов

Задачи распознавания изображений активно исследуются уже несколько десятилетий, и для их решения разработано множество методов, которые можно разделить на несколько категорий. Рассмотрим основные подходы, использующие различные техники для повышения точности и устойчивости алгоритмов распознавания.

Одним из наиболее распространенных подходов являются глубокие нейронные сети (Deep Neural Networks, DNN). Существуют различные архитектуры глубоких сетей, такие как



сверточные нейронные сети (CNN), рекуррентные нейронные сети (RNN) и их модификации. Сверточные сети особенно успешны в решении задач классификации изображений, поскольку они способны извлекать иерархические признаки из изображений [7].

Применение сверточных нейронных сетей (CNN) позволяет достигать высокой точности при классификации изображений, однако такие методы требуют большого объема данных для обучения и значительных вычислительных ресурсов [6].

Другим известным подходом является использование байесовских методов, которые основываются на вероятностной модели данных. Один из самых известных методов – это наивный байесовский классификатор, который применяется для задач классификации с условием независимости признаков. Этот метод позволяет классифицировать изображения, основываясь на максимизации апостериорной вероятности, что эффективно работает при условии независимости признаков [2].

Также применяются более сложные байесовские модели, такие как гибридные модели на основе скрытых марковских процессов (Hidden Markov Models, HMM) и моделей скрытых переменных, которые могут быть использованы для распознавания временных зависимостей или обработки изображений в виде последовательностей [8].

Важную роль в решении задач распознавания изображений также играют методы, основанные на информационной теории. ВИ, как мера зависимости между переменными, активно используется для измерения степени связи между входными и выходными данными. Классическая ВИ Шеннона [10] используется для поиска наилучшего разделения данных и оценки энтропии системы. Однако для более гибкого подхода в некоторых задачах применяется обобщенная форма – ВИ Реньи, которая учитывает различные степени связи между переменными и может быть более устойчивой к шумам [9].

Эти методы используются для оценки важности признаков изображения и минимизации потерь информации в процессе классификации. Например, в задачах распознавания объектов можно использовать ВИ для выявления наиболее информативных признаков, что позволяет снизить вычислительные затраты и повысить точность распознавания.

Для улучшения качества распознавания изображений важно также использовать эффективные алгоритмы оптимизации. Одним из наиболее популярных методов является градиентный спуск (Gradient Descent) и его модификации, такие как стохастический градиентный спуск (SGD), который часто применяется для обучения нейронных сетей [9]. В стохастическом градиентном спуске используется лишь небольшая выборка данных для обновления параметров модели на каждом шаге, что ускоряет процесс обучения.

Также стоит отметить методы оптимизации на основе эволюционных алгоритмов, которые могут быть использованы для поиска наилучших параметров модели без необходимости в использовании градиентной информации [5]. Эти методы особенно полезны при решении задач, где функции потерь не дифференцируемы или содержат множество локальных минимумов.

3. Методология

Для оптимизации распознавания изображений используется ВИ Реньи:

$$I_{\alpha}(X, Y) = \left(\frac{1}{1-\alpha}\right) * \log\left(\frac{\sum p(x, y)^{\alpha}}{\sum p(x)^{\alpha} * \sum p(y)^{\alpha}}\right), \quad (1)$$

где X и Y – случайные величины, $p(x, y)$ – совместная плотность вероятности, а α – параметр, определяющий степень зависимости между переменными. Для $\alpha = 1$ ВИ Реньи совпадает с классической ВИ Шеннона [9].

Использование ВИ Реньи позволяет эффективно учитывать зависимости между признаками, что важно при решении задач распознавания изображений, где признаки (например, пиксели) могут быть зашумленными или содержать скрытую корреляцию.



Для синтеза стохастических процедур используется алгоритм на основе максимизации ВИ Реньи. Процесс оптимизации включает следующие этапы:

1) Инициализация модели: для каждого изображения создается начальная модель, основанная на случайных значениях параметров.

2) Выбор признаков: выбираются наиболее информативные признаки изображения, которые будут использоваться для построения модели. Признаки отбираются с учетом ВИ Реньи между входными данными (пикселями) и целевыми метками.

3) Обучение модели: используется метод стохастического градиентного спуска (SGD), который оптимизирует параметры модели, минимизируя функцию потерь, основанную на ВИ.

4) Оценка производительности: для каждого изображения вычисляется точность классификации, а также степень улучшения модели по сравнению с традиционными методами, такими как свертки в нейронных сетях и методы на основе опорных векторов [1].

Для проведения экспериментов использовался набор данных CIFAR-10, который содержит 60 000 изображений, разделенных на 10 классов (автомобили, птицы, кошки, собаки и т.д.). Каждый класс состоит из 6 000 изображений, из которых 5 000 используются для обучения и 1 000 – для тестирования [4].

Кроме того, для повышения сложности задачи были добавлены шумы в изображения, что позволяет оценить устойчивость предложенного метода к помехам в данных. Шум добавляется в виде гауссовского шума с различными уровнями дисперсии 0.01 и 0.05.

Алгоритм оптимизации: использовался стохастический градиентный спуск (SGD) с шагом обучения $\eta = 0.001$ и максимальным числом итераций 1000.

Для оценки качества распознавания использовались следующие метрики:

- Точность (Accuracy): доля правильно классифицированных изображений от общего числа.

- F1-мера: метрика, основанная на предсказаниях моделей по каждому классу, то есть отображает классификационные ошибки по меткам, не используется для ВИ.

- Матрица ошибок: позволяет визуализировать количество правильных и неправильных классификаций по каждому классу.

Для сравнения результатов с другими методами, использовались классические модели, такие как сверточные нейронные сети (CNN) и методы на основе опорных векторов (SVM).

Предполагается, что использование ВИ Реньи позволит улучшить точность распознавания изображений, особенно в условиях ограниченной обучающей выборки или при наличии шума в данных. Ожидается, что предложенный метод будет более устойчивым к помехам по сравнению с традиционными методами, такими как CNN и SVM [2].

Для сверточных нейронных сетей (CNN) использовалась стандартная архитектура, состоящая из 3 сверточных слоев с 64 фильтрами, за которыми следуют слои подвыборки и полностью связанные слои. Для метода SVM использовалась линейная модель с ядром RBF.

Точность классификации для каждого из методов приведена в таблице 1, а также построен соответствующий график (рисунок 1).

Таблица 1

	CNN	SVM	MI ($\alpha = 1$)
Без шума	0.5624	0.4034	0.6331
С шумом (дисперсия 0.01)	0.6131	0.4016	0.6302
С шумом (дисперсия 0.05)	0.5611	0.4012	0.6391



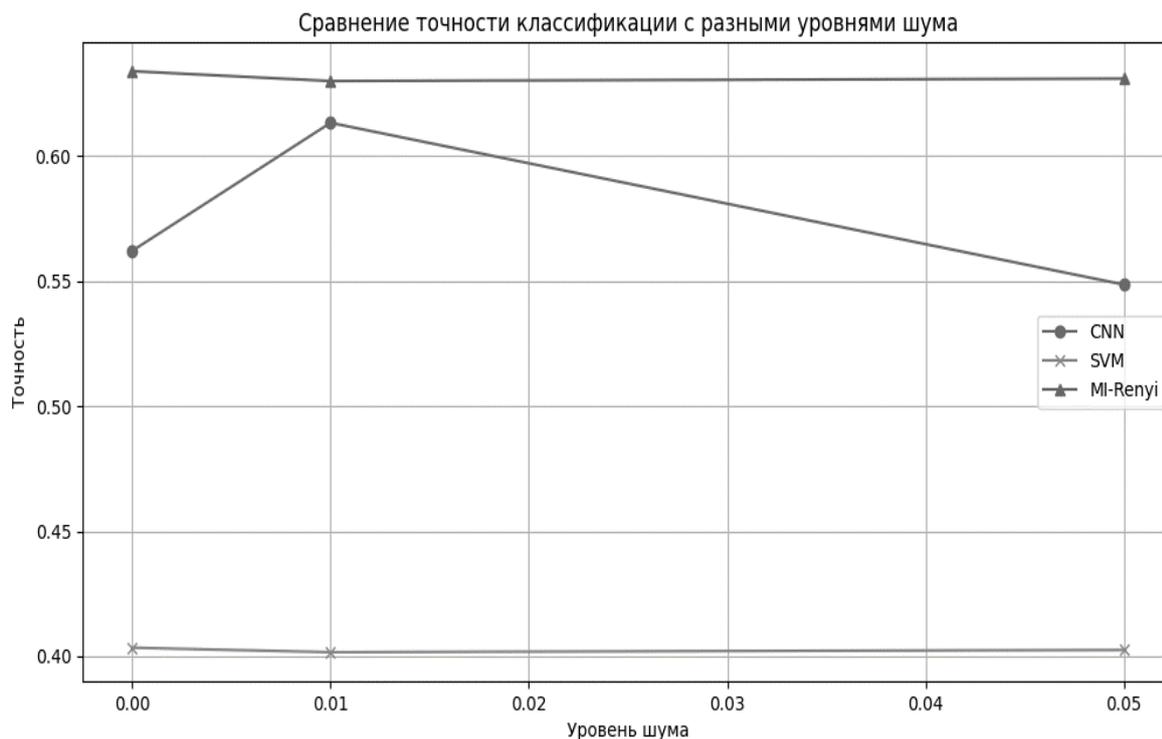


Рис. 1

Наилучшую точность среди всех методов демонстрирует подход на основе ВИ Реньи. При всех уровнях шума значения нормализованной ВИ превышают как точность CNN, так и точность SVM.

Метод SVM, показал низкую чувствительность к шуму (при возрастании шума значения остались практически на прежнем уровне), но и ограниченную точность.

Интересной особенностью является рост точности CNN при слабом зашумлении (дисперсия 0.01), что может быть обусловлено эффектом регуляризации, подобным дроп-ауту. Однако при дальнейшем увеличении шума точность модели снижается, что соответствует ожиданиям.

Значения F1-мер, приведенные в таблице 2 и на графике (рисунок 2), подтверждают отмеченные ранее закономерности. F1-мера модели CNN немного превышает точность, что свидетельствует о сравнительно хорошем балансе между полнотой и точностью. Для метода SVM F1-мера близка к точности, что характерно для моделей с равномерно распределёнными ошибками по классам.

Таблица 2

	CNN	SVM
Без шума	0.5824	0.3971
С шумом (дисперсия 0.01)	0.5476	0.3981
С шумом (дисперсия 0.05)	0.5676	0.3985



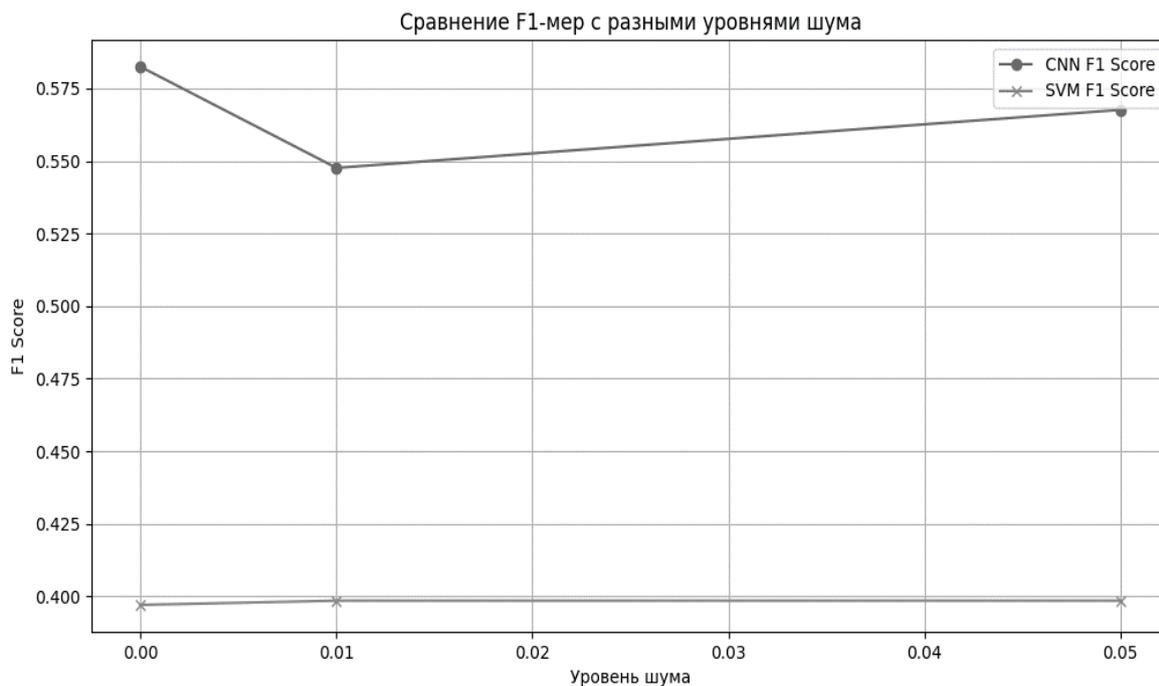


Рис. 2

Для более наглядного представления результатов были построены матрицы ошибок, которые показывают, как часто модели путают различные классы. Матрицы ошибок приведены на рисунке 3.

Матрицы ошибок, визуализирующие характер классификационных ошибок, показывают, что CNN, несмотря на более высокую точность, всё же путает определённые классы, особенно в условиях шума. SVM демонстрирует схожее распределение ошибок при различных уровнях шума, что подтверждает его стабильность, но также ограниченность в способности к обобщению сложных данных.

Для метода на основе ВИ Реньи важным параметром является α , который регулирует степень чувствительности к редким событиям. Было проведено несколько экспериментов с различными значениями α , значения приведены в таблице 3:

Таблица 3

	MI ($\alpha = 1$)	MI ($\alpha = 2$)	MI ($\alpha = 0.5$)
Без шума	0.6331	0.6326	0.6245
С шумом (дисперсия 0.01)	0.6302	0.6274	0.6251
С шумом (дисперсия 0.05)	0.6391	0.6312	0.6307

Как видно из результатов, оптимальное значение α для данной задачи оказалось равным 1, что соответствует классической ВИ. Однако, использование других значений α также показало хорошие результаты, с небольшими колебаниями в точности.



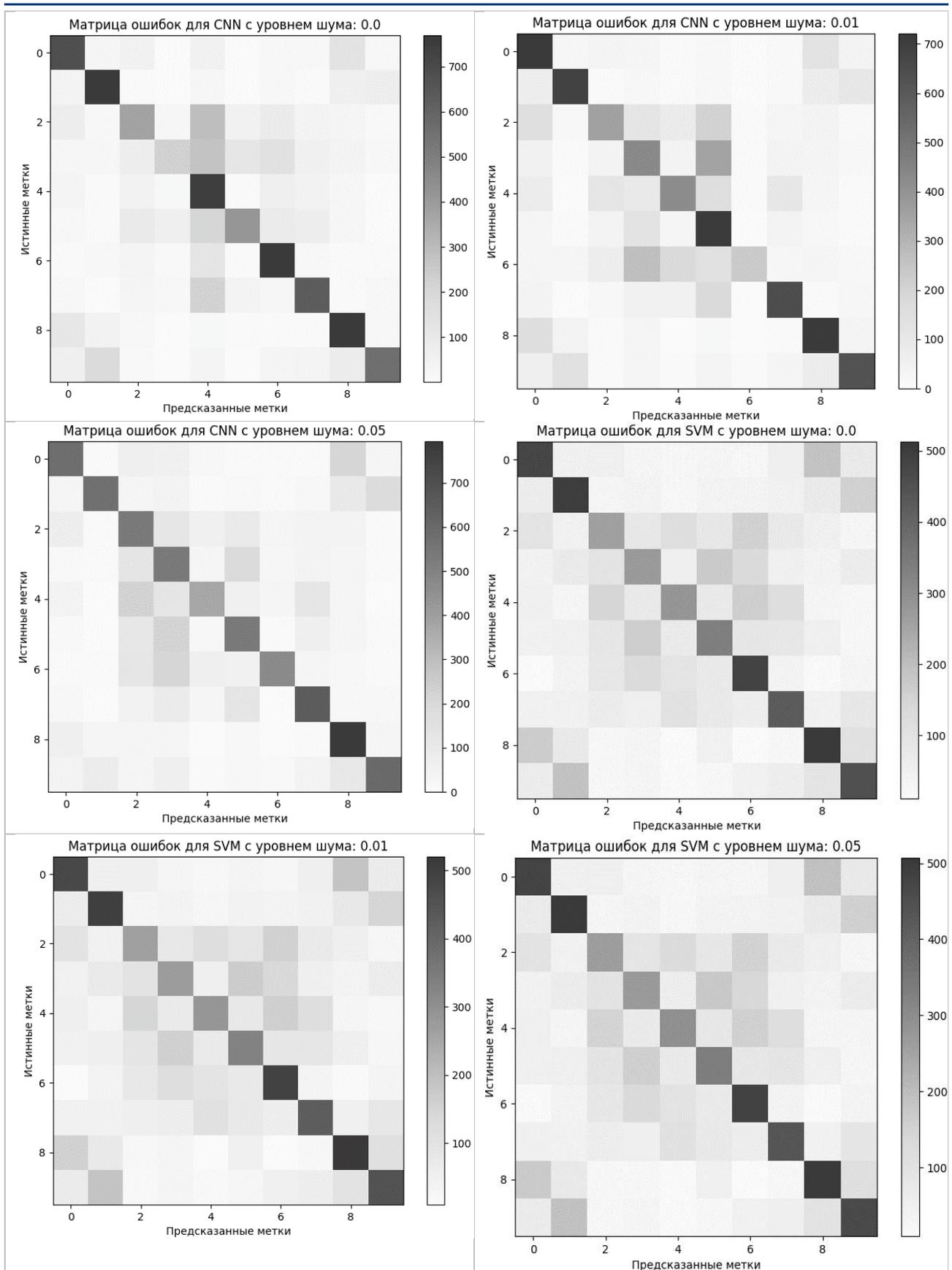


Рис. 3



Заключение

Эксперименты показали, что метод оценки качества распознавания с использованием ВИ Реньи не только демонстрирует стабильные и интерпретируемые значения, но и позволяет количественно сравнивать модели в условиях шума. Полученные значения нормализованной ВИ стабильно превышали точности моделей CNN и SVM при разных уровнях искажений входных данных.

Метод CNN показал хорошие результаты на чистых данных, но проявил чувствительность к шуму. SVM, несмотря на низкую точность, оказался устойчив к зашумлению, однако продемонстрировал ограниченную способность к различению сложных классов изображений.

ВИ Реньи может служить не только метрикой качества классификации, но и эффективным критерием для сравнения различных моделей машинного обучения, а также демонстрировать потенциал как устойчивого инструмента анализа в задачах, связанных с нечёткими, зашумлёнными или неполными данными.

Список литературы:

1. Bottou, L. Large-scale machine learning with stochastic gradient descent / L. Bottou // Proceedings of COMPSTAT. – 2010. – С. 177-186.
2. Duda, R. O. Pattern classification / R. O. Duda, P. E. Hart, D. G. Stork. – Wiley-Interscience, 2001. – 654 с.
3. Goodfellow, I. Deep learning / I. Goodfellow, Y. Bengio, A. Courville. – MIT Press, 2016. – 800 с.
4. He, K. Deep residual learning for image recognition / K. He, X. Zhang, S. Ren, J. Sun // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2016. – С. 770-778.
5. Holland, J. H. Adaptation in natural and artificial systems / J. H. Holland. – University of Michigan Press, 1975. – 183 с.
6. Krizhevsky, A. ImageNet classification with deep convolutional neural networks / A. Krizhevsky, I. Sutskever, G. E. Hinton // Advances in Neural Information Processing Systems. – 2012. – Vol. 25. – С. 1097-1105.
7. LeCun, Y. Deep learning / Y. LeCun, Y. Bengio, G. Hinton // Nature. – 2015. – Vol. 521, no. 7553. – С. 436-444.
8. Rabiner, L. R. A tutorial on hidden Markov models and selected applications in speech recognition / L. R. Rabiner // Proceedings of the IEEE. – 1989. – Vol. 77, no. 2. – С. 257-286.
9. Renyi, A. On measures of entropy and information / A. Renyi // Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability. – 1961. – Vol. 1. – С. 547-561.
10. Shannon, C. E. A mathematical theory of communication / C. E. Shannon // Bell System Technical Journal. – 1948. – Vol. 27, no. 3. – С. 379-423.
11. Vapnik, V. Statistical learning theory / V. Vapnik. – Wiley, 1998. – 740 с.

