

УДК 004.[932.75'1+89]

Я. А. БУРЫЙ, Д. И. САМАЛЬ

ПРИМЕНЕНИЕ ЭВОЛЮЦИОННОЙ ПАРАДИГМЫ К ПРОЕКТИРОВАНИЮ АРХИТЕКТУРЫ НЕЙРОННОЙ СЕТИ ДЛЯ РАСПОЗНАВАНИЯ ИСКАЖЁННОГО ТЕКСТА

Белорусский государственный университет информатики и радиоэлектроники

В статье представлена попытка объединения эволюционных алгоритмов и нейронных сетей при проектировании и обучении системы распознавания искажённых изображений текста.

В течение последних десятилетий искусственные нейронные сети хорошо себя зарекомендовали во многих областях искусственного интеллекта, таких, например, как прогнозирование, оптимизация, анализ данных, распознавание образов и принятие решений. Тем не менее, традиционные эвристические подходы к разработке топологии многослойных нейронных сетей основываются на рекомбинации уже существующих нейросетевых архитектур. Такой подход позволяет решать широкий спектр задач, но подразумевает соблюдение специфических условий для качественной работы алгоритмов.

Естественные аналоги подобных интеллектуальных систем в живой природе, однако, достаточно универсальны, чтобы адаптироваться практически к любой среде обитания.

Несмотря на их чрезвычайную сложность и ограниченные возможности к исследованию их структур, известно, что эти конструкции были сформированы в результате эволюционного процесса. И если на сегодняшний день невозможно определить точную архитектуру связей в биологических нейросистемах, то, по крайней мере, можно попытаться воспроизвести сам процесс их формирования с целью получения более универсального алгоритма, чем те, что разработаны к настоящему моменту.

В рассматриваемой работе окончательная структура ядра системы классификации образуется в результате эволюционного процесса, с учетом известных сегодня знаний об особенностях развития и строения нервной системы позвоночных.

Использование описываемого подхода позволяет абстрагироваться от ограничений существующих нейросетевых алгоритмов, обусловленных сферой применения конкретных типов их структур.

Ключевые слова: нейронные сети, эволюционные алгоритмы, генетический алгоритм, распознавание текста.

Введение

Эволюционное моделирование, в силу близости соответствующих алгоритмов к их естественным аналогам, позволяет воспроизвести происходящие в окружающей природе процессы, целью которых является адаптация заданной стохастической модели к некоторой среде её функционирования.

В качестве прототипа такой модели целесообразно выбрать уже существующий в природе аналог разрабатываемой системы и применить к его конструированию эволюционный подход, наподобие того, как это происходило исторически, согласно известным на сегодняшний день палеонтологическим и антропологическим сведениям.

Распознавание искажённых изображений текста (далее просто – «искажённого текста»)

представляет собой одну из многих задач в направлении распознавания образов и обработки изображений. Её можно определить как нахождение некоторой последовательности символов заданного алфавита, которая максимально соответствует цифровому изображению, поданному на вход системы распознавания.

В случае распознавания искажённого текста всё изображение или его части могут быть деформированы, но не более той степени, которая делала бы невозможным распознавание начертанного текста человеком, обученным чтению.

Если абстрагироваться от семантики заданного текста, то любая текстовая конструкция укладывается в рамки стохастической модели, данной как результат взаимодействия системы с окружающей средой.

Эволюционное моделирование базируется на замене самого процесса разработки информационной системы имитацией её эволюционного развития. Наиболее развитым направлением эволюционного моделирования в настоящее время являются, так называемые, генетические алгоритмы.

Нейронные сети являются одним из подходов к построению решений задач многомерной оптимизации и имитируют мыслительную деятельность нервной системы позвоночных. Они широко используются в качестве классификаторов в системах распознавания образов. Однако существующие в настоящее время системы распознавания текста, в том числе на базе нейронных сетей, не обладают достаточной инвариантностью к существенным изменениям размеров, а также к различным неафинным и/или нелинейным преобразованиям (например, волнообразным, сгибающим, растягивающим, сдвигающим и т. п. искажениям) изображений распознаваемых символов.

Постановка задачи

Разработанные к настоящему моменту методы и подходы к классификации и распознаванию образов основываются на математических преобразованиях в многомерных пространствах признаков. Несмотря на разнообразие используемых метрик и то, что в качестве признаков могут выступать достаточно сложные характеристики входных образов, эти признаковые гиперпространства обладают свойствами конечномерных евклидовых пространств.

Этого может быть вполне достаточно для описания исходных объектов распознавания, однако, в случае распознавания искажённых изображений (рис. 1), пространство признаков, в котором задано описание объекта, анизотропно. В таком случае конкретное распознаваемое изображение есть представление объекта в одном из тензорных полей на исходном пространстве признаков, а все возможные вариации допустимых искажений есть не что иное, как тензорное расслоение на этом пространстве.

Теоретически возможно восстановить исходное описание объекта в изначальном изо-

тропном признаковом пространстве и для дальнейшей классификации использовать существующий математический аппарат, но для этого необходимо точно установить характеристики тензорного поля, в котором дано искажённое описание образа. Это значит, что в заданном тензорном расслоении необходимо отыскать такое сечение, которое в точности отображает исходный эталонный неискажённый образ одного из символов алфавита на распознаваемое изображение. При этом этот найденный эталонный символ и будет соответствовать решению задачи распознавания.

Проблема обладает свойствами каскадной задачи оптимизации, где необходимо решить следующие подзадачи:

- построение такого тензорного расслоения, которое реализует в каком-то приближении все возможные допустимые искажения исходного пространства признаков;
- устранение искажений из изображения, т. е. нахождение вышеуказанного отображающего сечения в тензорном расслоении;
- распознавание восстановленного неискажённого образа в исходном изотропном признаковом пространстве.

Решению данной проблемы прямыми численными методами, например, полным перебором (при поддержке уже разработанных механизмов с доказанной эффективностью), препятствует экспоненциальный рост и без того существенных вычислительных затрат, выраженных количеством операций с плавающей запятой. Это обусловлено тем, что размерности векторных расслоений на топологических многообразиях могут на порядки превышать размерности самих исходных многообразий.

Однако в природе, как известно, имеется более элегантное и менее требовательное к вычислительным ресурсам решение – система восприятия человека.

Если рассматривать это решение в историческом контексте с точки зрения имеющейся к настоящему моменту информации о его строении и развитии, можно определить в нём некоторые однозначные характеристики, такие как очевидный нейросетевой характер архитектуры, а также эволюционную природу процесса формирования структуры связей между нейронами.

Многослойные нейронные сети способны с различной точностью аппроксимировать лю-



Рис. 1. Искажённое изображение текста

бую из существующих логических функций [1]. Следовательно, стохастическое описание эволюционно сформированной зрительной системы человека может быть дано с точки зрения нейросетевого моделирования.

Существуют различные способы использования системы распознавания. С одной стороны, это может подразумевать распознавание длинных кусков текста, с другой – система может использоваться для анализа коротких текстов ограниченного алфавита с хорошо различаемыми символами, таких как паспортные данные, адреса домов на улице и т. п. При этом с эволюционной точки зрения вторая будет являться прародителем первой, т. к. функционально значительно проще и предоставляет разработчику широкие возможности для её конструирования.

Таким образом, можно сформулировать следующую *гипотезу*: с помощью эволюционного моделирования за конечное число итераций соответствующего алгоритма можно построить такую архитектуру связей нейронной сети, которая будет способна распознавать с заданной точностью искажённые изображения коротких участков текста, написанного ограниченным алфавитом хорошо различающихся между собой символов.

Задачей данного исследования является проверка указанной гипотезы.

Обычно приспособленность каждой особи (генотипа) оценивается значением среднеквадратической погрешности, рассчитанной по соответствующей этой особи нейронной сети (фенотипу) [2]. В данной работе под «особью» будет подразумеваться конкретный вариант структуры нейронной сети, включающий в себя матрицы связей и весовых коэффициентов.

Таким образом, в экспериментальной проверке гипотезы необходимо исследовать следующие основные и вспомогательные выходные характеристики системы распознавания, полученные в результате функционирования соответствующего эволюционного алгоритма и имеющие значение с точки зрения оценки качества полученного результата:

- среднеквадратическая ошибка распознавания символов текста

$$\varepsilon = \sqrt{\frac{\sum_{i=1}^N \sum_{j=1}^M (e_{i,j})^2}{NM}},$$

где $e_{i,j}$ – ошибка распознавания j -го символа i -го текста из тестирующей выборки;

- абсолютная точность распознавания. Этот параметр является основным для оценки качества полученной системы распознавания с точки зрения её применения, т. к. означает, какая именно часть изображений текста распознана успешно. Рассчитывается по формуле:

$$\tau = 1 - \frac{\sum_{i=1}^N \sum_{j=1}^M E_{i,j}}{NM},$$

где

$$E_{i,j} = \begin{cases} 1, & \text{если } i, j\text{-й символ распознан} \\ & \text{с ошибкой,} \\ 0, & \text{если } i, j\text{-й символ распознан} \\ & \text{без ошибки.} \end{cases}$$

Кроме того, для оценки эффективности эволюционного алгоритма целесообразно ориентировочно оценить следующие характеристики: скорость схождения эволюционного алгоритма подбора параметров системы распознавания; вычислительная стоимость схождения эволюционного алгоритма; итоговая сложность полученной архитектуры нейронной сети.

Результаты экспериментов

Для проверки вышесформулированной гипотезы была проведена серия более чем из 50 экспериментов для различных комбинаций характеристик эволюционного процесса и проектируемой нейронной сети. Для их проведения использовалась база изображений «The Street View House Numbers (SVHN) Dataset» [3]. На рис. 2 представлены некоторые изображения из обучающей и тестирующей выборок этой базы.

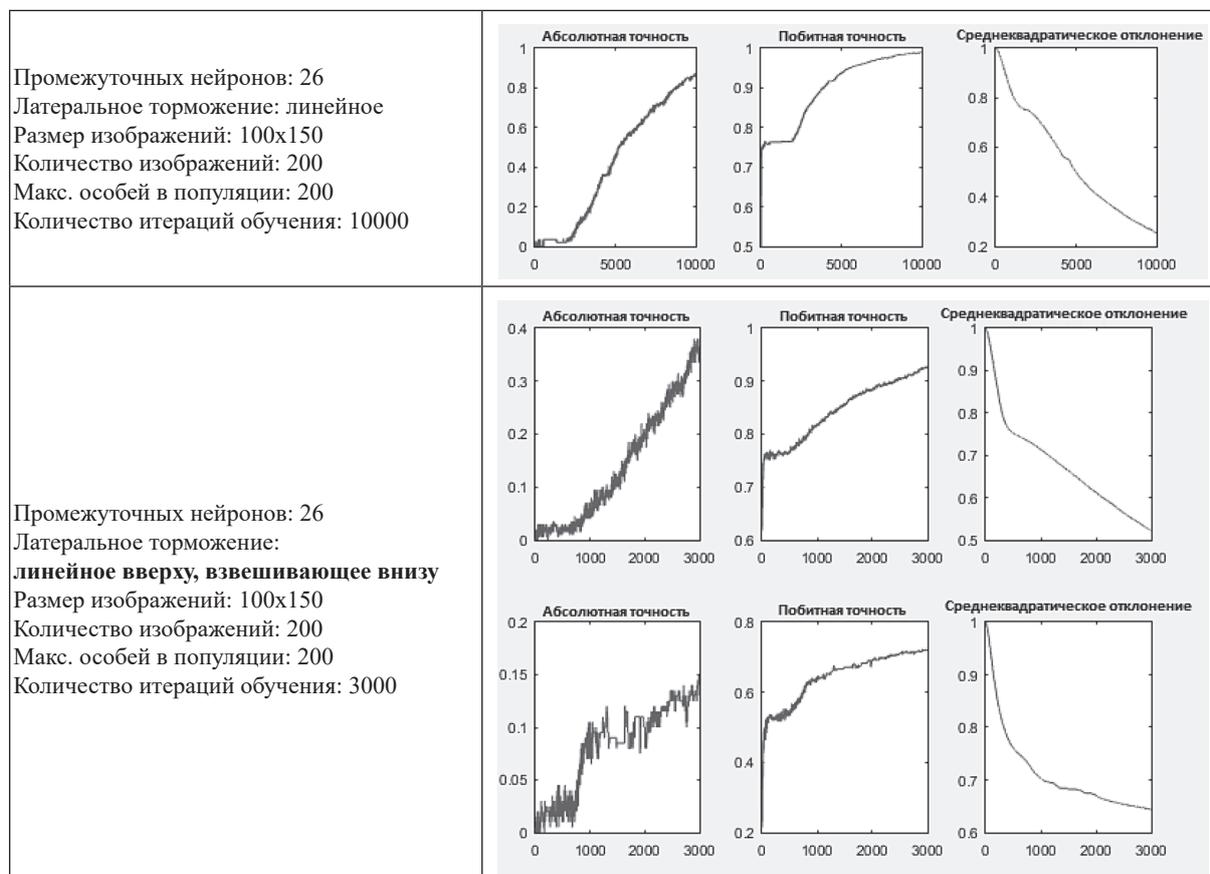
В таблице представлены отдельные экспериментальные данные, характерные для всего эксперимента в целом.

В ней приведены графики следующих выходных характеристик полученного в результате эволюционного моделирования алгоритма распознавания: абсолютная точность распознавания – отношение успешно распознанных изображений к общему количеству изображений в тестовой выборке, побитная точность – отношение успешно распознанных бит в коде выходного сигнала к общему количеству бит сигнала по всей выборке, среднеква-



Рис. 2. Примеры изображений из базы SVHN

Результаты работы генетического алгоритма настройки системы распознавания



дратическое отклонение ошибки выходного сигнала. В данной серии экспериментов под «битами сигнала» подразумевается битовое представление кода распознанного числа.

Из графиков видно, как в результате работы генетического алгоритма настройки нейронной сети происходит устойчивое последовательное улучшение целевых функций, что говорит о принципиальной возможности нахождения архитектуры связей нейронной сети для распознавания искажённого текста с помощью эволюционного моделирования, а также

о жизнеспособности конкретного выбранного эволюционного решения.

Несмотря на то, что в качестве основной взвешивающей функции алгоритма оптимизации используется среднеквадратическое отклонение, на графиках хорошо видно, что вместе с падением ошибки абсолютная и побитная точность устойчиво растут к единице. При различных числовых и структурных характеристиках эволюционного алгоритма и проектируемой нейронной сети наблюдалась большая относительная разница в скорости сжог-

дения алгоритма. Для разных комбинаций указанных параметров достигнутое за конечное количество шагов значение среднеквадратической ошибки варьировалось в пределах интервала $\varepsilon \in [0.1, 0.6]$, при пороговых значениях выхода нейрона $u \in [-1, 1]$, и в среднем составило $\varepsilon = 0.2$. Примерно в 86% всех случаев наблюдалось устойчивое улучшение целевых функций по траектории, близкой к логарифмическому закону. Оставшиеся 14% случаев схождения целевой функции к локальному минимуму градиента имели место при экстремальных значениях характеристик алгоритма.

Также заметно, что применение линейного порогового элемента в качестве функции латерального торможения показывает заметно лучший результат, т. к. в этом случае используется только одна матрица весовых коэффициентов,

что значительно уменьшает сложность эволюционного алгоритма, а также требования системы к вычислительным ресурсам.

Кроме того, в ходе экспериментов были обнаружены технические ограничения, указывающие на необходимость ряда улучшений алгоритма с целью уменьшения требований к памяти и процессорному времени.

Заключение

На основе анализа экспериментальных данных можно сделать вывод о том, что изначальная гипотеза может быть принята с вероятностью 86%. В то же время, для достижения более значимых результатов, на последующих этапах работы системы целесообразно применять так же семантический анализ распознанных вариантов текста.

Литература

1. Хижняков, Ю. Н. Алгоритмы нечеткого, нейронного и нейро-нечеткого управления в системах реального времени / Ю. Н. Хижняков. – Пермь: ПНИПУ, 2013. – 160 с.
2. Рутковская, Д. Нейронные сети, генетические алгоритмы и нечеткие системы / Д. Рутковская, М. Пилиньский, Л. Рутковский – М.: Горячая линия – Телеком, 2013. – 384 с.
3. База изображений The Street View House Numbers (SVHN) Dataset – [Электронный ресурс]. – Режим доступа: <http://ufldl.stanford.edu/housenumbers>. – Дата доступа: 01.10.2017.
4. Плотников, А. Д. Математическое программирование / А. Д. Плотников – Минск: Новое знание, 2007. – 171 с.
5. Головкин, В. А. Нейронные сети: обучение, организация и применение. Кн. 4 / В. А. Головкин, Общая ред. А. И. Галущкин – М.: ИПРЖР, 2001. – 256 с.
6. Монтгомери, Д. К. Планирование эксперимента и анализ данных / Д. К. Монтгомери, Л.: Судостроение, 1980. – 384 с.
7. Хайкин, С. Нейронные сети. Полный курс / С. Хайкин. – М., СПб., Киев: Вильямс, 2006. – 1104 с.

References

1. Hizhnjakov, Ju. N. Algorithms of fuzzy, neural and neural-fuzzy control in real-time systems / Ju. N. Hizhnjakov. – Perm: PNIPU, 2013. – 160 s.
2. Rutkovskaja, D. Neural networks, genetic algorithms and fuzzy systems / D. Rutkovskaja, M. Pilińskij, L. Rutkovskij – M.: Gorjachaja linija – Telekom, 2013. – 384 s.
3. The Street View House Numbers (SVHN) Dataset – [Electronic resource]. – Access mode: <http://ufldl.stanford.edu/housenumbers>. – Access date: 01.10.2017.
4. Plotnikov, A. D. Mathematical programming / A. D. Plotnikov – Minsk: Novoe znanie, 2007. – 171 s.
5. Golovko, V. A. Neural networks: training, organization and application. Kn. 4 / V. A. Golovko, A. I. Galushkin – M.: IPRZhR, 2001. – 256 s.
6. Montgomery, D. K. Experiment planning and data analysis / D. K. Montgomery, L.: Sudostroenie, 1980. – 384 s.
7. Hajkin, S. Neural networks. Full course / S. Hajkin. – M., SPb., Kiev: Vil'jams, 2006. – 1104 s.

Поступила После доработки
17.10.2017

Принята к печати
15.11.2017

15.12.2017

Bury Y. A., Samal D. I.

APPLICATION OF THE EVOLUTIONARY PARADIGM TO DESIGNING ARCHITECTURE OF A NEURAL NETWORK FOR RECOGNIZING THE DISTORTED TEXT

Belarusian State University of Informatics and Radioelectronics

The paper presents an attempt to apply of evolutionary methods to the design and training of a system for recognizing distorted text.

Over the past decades, artificial neural networks are widely used in many areas of artificial intelligence, such as forecasting, optimization, data analysis, pattern recognition and decision making. Nevertheless, the traditional heuristic approaches to design of multi-layer neural networks are based on the recombination of already existing neural network architectures.

This approach allows us to solve a wide range of problems, but implies compliance with specific conditions for the quality work of algorithms.

The natural analogues of such intelligent systems in living nature, however, are universal enough to adapt to virtually any habitat.

Despite their extreme complexity and limited ability to study their structures, it is known that these structures were formed as a result of the evolutionary process. And if today it is impossible to determine the exact architecture of the links in biological neural systems, then at least one can try to reproduce the very process of their formation in order to obtain a more universal algorithm than those developed to the present moment.

In opposite to them we form the final structure of the core of the classification system by evolutionary process, taking into account the knowledge about the features of the development and construction of the nervous system of vertebrates.

Applying of the approach makes it possible to abstract from the limitations of existing neural network algorithms, caused by the scope of application of specific types of their structures.

Keywords: Neural networks, evolutionary algorithms, genetic algorithm, text recognition.



Бурый Ярослав Анатольевич Ассистент кафедры электронных вычислительных машин БГУИР, аспирант кафедры электронных вычислительных машин БГУИР.

Bury Y. A. Assistant of the Department of Electronic Computing Machines BSUIR, Post-graduate student of the department of electronic computers BSUIR.

Email: edidici@tut.by



Самаль Дмитрий Иванович Зав. кафедрой электронных вычислительных машин БГУИР, кандидат технических наук, доцент.

Samal D. I. Head of the Department of Electronic Computing Machines BSUIR, Ph. D., Associate Professor.

Email: samal@bsuir.by