

УДК 004.93'1

*М.Д. Поляк*  
*Санкт-Петербург, Россия*

## **МОДЕЛЬ СТОХАСТИЧЕСКОЙ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ В ЗАДАЧАХ РАСПОЗНАВАНИЯ ОБРАЗОВ**

*M.D. Polyak*  
*St.-Petersburg, Russia*

### **A STOCHASTIC ARTIFICIAL NEURAL NETWORK MODEL FOR PATTERN RECOGNITION**

---

Рассмотрена стохастическая искусственная нейронная сеть на основе машины Больцмана и особенности ее применения к задачам обработки данных, в частности, к задаче распознавания рукописных чисел. Предложен алгоритм оценки внутреннего состояния такой сети с помощью карт признаков.

СТОХАСТИЧЕСКАЯ ИСКУССТВЕННАЯ НЕЙРОННАЯ СЕТЬ. МАШИНА БОЛЬЦМАНА. РАСПОЗНАВАНИЕ ОБРАЗОВ.

A stochastic artificial neural network based on the Boltzmann machine is examined and some aspects of applying stochastic networks to information processing tasks such as handwritten digits recognition are considered. An algorithm is proposed for evaluating internal state of the network with feature maps.

STOCHASTIC ARTIFICIAL NEURAL NETWORK. BOLTZMANN MACHINE. PATTERN RECOGNITION.

---

Традиционные алгоритмы компьютерного зрения и машинного обучения не могут сравниться с производительностью человеческого мозга при решении таких задач, как распознавание рукописного текста, выделение объектов на изображении, распознавание лиц и т. п. Искусственные нейронные сети позволяют алгоритмам распознавания образов приблизиться к производительности человеческого мозга. С помощью нейросетевого подхода становится возможным автоматизировать процесс выделения информативных признаков из изображения или сложного по своей структуре сигнала.

Искусственную нейронную сеть принято называть *стохастической*, если ее весовые коэффициенты, характеризующие связи между нейронами, принимают случайные значения, либо если передаточная функция нейрона является случай-

ной функцией. В последнем случае нейронную сеть также называют *машиной Больцмана* [1, 2].

Стохастические нейронные сети позволяют строить вероятностные модели сложных процессов. В ходе обучения на основе имеющихся экспериментальных данных (обучающей выборки) такая сеть формирует статистическую модель, способную самостоятельно воспроизводить данные в соответствии с их эмпирическим законом распределения [3]. Подобная модель может использоваться для генерации достоверных данных в случае, если экспериментальных данных недостаточно. Детальный анализ нейросетевой модели может помочь пролить свет на структуру исходных процессов, оказавших влияние на формирование экспериментальных данных. Такой подход может применяться в интеллектуальном анализе данных (data mining), при решении задач

классификации, моделирования, прогнозирования и др.

Несомненное достоинство рассматриваемого класса нейросетевых алгоритмов – возможность визуального представления результатов анализа данных, визуализация модели.

В настоящей статье рассматривается стохастическая искусственная нейронная сеть на основе машины Больцмана и возможности ее применения к задачам обработки данных, в частности, для выделения признаков, классификации, распознавания. Приводится анализ нейросетевого подхода к задаче распознавания рукописных чисел. При обработке изображений нейронная сеть может быть представлена как адаптивная система с множеством входов, а при анализе временных рядов нейронная сеть представляется системой с одним входом [4], что позволяет использовать одни и те же нейросетевые алгоритмы для решения задач разного типа.

При решении задачи распознавания образов можно выделить следующие два характерных этапа: выделение информативных признаков из исходного изображения или сигнала и классификация объекта по набору описывающих его признаков. Стохастическая нейронная сеть может использоваться на каждом из этапов по отдельности или задача распознавания образов может целиком решаться с помощью одной многослойной сети. Цель данной статьи – рассмотрение особенностей использования многослойной стохастической нейронной сети для решения задачи распознавания образов.

**Стохастическая нейронная сеть.** Классическая машина Больцмана является обобщением нейросетевой модели Хопфилда [2, 5]. Сеть Хопфилда представляет собой рекуррентную нейронную сеть, состоящую из единственного слоя двоичных нейронов, функционирующих

в соответствии с моделью искусственного нейрона МакКаллока–Питтса и имеющих пороговую функцию активации [6]. Модель Хопфилда представляет собой ассоциативную память, она позволяет восстанавливать зашумленные и неполные образы. Ее основные недостатки – малый объем памяти, избыточное количество нейронов, недостоверность результатов (ложные образы) [2].

Искусственная нейронная сеть на основе машины Больцмана имеет два существенных отличия от сети Хопфилда. Во-первых, в машине Больцмана решение о том, что нейрон должен перейти из состояния  $-1$  в состояние  $+1$  (или наоборот) является вероятностным:

$$s_i = \begin{cases} +1, & \text{с вероятностью } p(+1) \\ -1, & \text{с вероятностью } 1 - p(+1) \end{cases}, \quad (1)$$

где  $p(+1)$  – вероятность того, что нейрон  $s_i$  перейдет из состояния  $-1$  («off») в состояние  $+1$  («on»), а функция, определяющая вероятность активации нейрона, является сигмоидальной. Выбор множества  $\{0;1\}$  возможных состояний двоичного стохастического нейрона вместо принятого в модели Хопфилда множества  $\{-1;+1\}$  не влияет на структуру модели.

Во-вторых, все нейроны в машине Больцмана делятся на два класса: «видимые» и «скрытые». Видимые нейроны в свою очередь могут подразделяться на «входы» и «выходы» сети. Сеть Хопфилда целиком состоит из видимых нейронов и не имеет скрытых. Скрытые нейроны или, как их еще называют, скрытые переменные, позволяют машине Больцмана успешно находить существенные корреляционные связи между признаками, поступающими на вход сети.

Машина Больцмана представляет собой полностью связанный граф, представленный на рис. 1 а, в котором могут быть выделены два слоя: слой видимых и слой скрытых вершин. Алгоритмы обу-

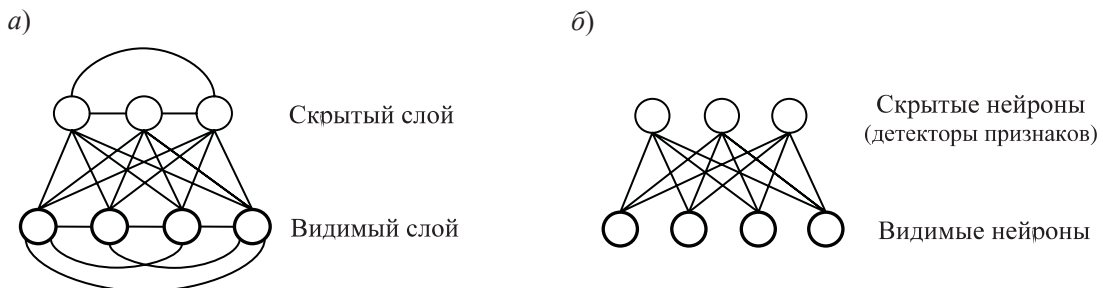


Рис. 1. Графическая модель машины Больцмана с четырьмя видимыми нейронами и тремя скрытыми: а – машина Больцмана; б – машина Больцмана с ограничениями

чения и функционирования машины Больцмана подробно рассмотрены в литературе [1, 7, 8]. Из-за большого количества связей в сети алгоритм обучения машины Больцмана труднореализуем, поэтому на практике вводят ряд ограничений.

Машина Больцмана с ограничениями (МБО) – неориентированная графическая модель с двухслойной архитектурой. В отличие от рассмотренной выше общей машины Больцмана, в МБО запрещены все связи видимый-видимый и скрытый-скрытый между нейронами одного слоя. В такой сети присутствуют только связи видимый-скрытый, как показано на рис. 1 б. МБО является частным случаем общей машины Больцмана, в которой веса всех связей между нейронами одного слоя равны нулю.

МБО часто рассматривают в качестве мультипликативной экспертной модели [7], поскольку она совмещает в себе множество отдельных составных моделей (экспертов), перемножая их результаты. Такими экспертами являются двоичные скрытые нейроны, которые можно рассматривать как детекторы признаков (рис. 3 а). Нейрон-эксперт (детектор признака) выдает единицу на своем выходе с вероятностью

$$P(h_j = 1 | \mathbf{v}) = \frac{1}{1 + \exp\left(-\sum_i w_{ij} v_i - a_j\right)}, \quad (2)$$

где  $h_j = s_j$  – состояние скрытого нейрона  $j$ ,  $h_j \in \{0, 1\}$ ;  $\mathbf{v}$  – двоичный вектор, описывающий состояние всех видимых нейронов;  $v_i = s_i$  – состояние видимого нейрона  $i$ , т. е.  $i$ -я координата вектора  $\mathbf{v}$ ,  $v_i \in \{0, 1\}$ ;  $w_{ij} = w_{ji}$  – вес связи между  $i$ -м видимым и  $j$ -м скрытым нейронами;  $a_j$  – смещение скрытого нейрона  $j$ .

Таким образом, вероятность  $P(h_j = 1 | \mathbf{v})$  в выражении (2) характеризует вероятность того, что образ  $\mathbf{v}$ , поданный на вход сети, содержит признак  $j$ . Из (2) также следует, что условное распределение для вектора состояния скрытых нейронов может быть задано следующим образом:

$$p(\mathbf{h} | \mathbf{v}; W) = \prod_j P(h_j | \mathbf{v}), \quad (3)$$

где  $\mathbf{h}$  – двоичный вектор, описывающий состояние всех скрытых нейронов, т. е. вектор признаков;  $W$  – симметричная матрица весов связей  $w_{ij}$  между всеми нейронами сети. Выражение (3) описывает мультипликативную экспертную модель.

Существенное ограничение МБО – использование двоичных нейронов, поскольку многие



Рис. 2. Примеры изображений рукописных чисел

реальные данные являются вещественными. Эта проблема решается введением модели стохастического нейрона с вещественным выходом [9]. Функция активации такого нейрона имеет следующий вид:

$$s_i = \frac{\theta_L + (\theta_H - \theta_L)}{1 + \exp\left(-a_i \left(\sum_j w_{ij} s_j + N_i(0, \sigma^2)\right)\right)}, \quad (4)$$

где  $\theta_L$  и  $\theta_H$  – нижняя и верхняя границы принимаемых нейроном вещественных значений;  $s_j$ ,  $s_j$  – видимый или скрытый нейрон;  $a_i$  – параметр;  $N_i(0, \sigma^2)$  – гауссов белый шум с нулевым математическим ожиданием и среднеквадратическим отклонением  $\sigma$ .

Нейрон с функцией активации (4) фактически представляет собой модель зашумленного сигнала и в процессе обучения нейронная сеть учится игнорировать шумы (нормальные отклонения) в экспериментальной выборке. Результатом обучения такой сети является робастная нейросетевая модель.

**Распознавание рукописного текста.** Для исследования процесса выбора информативных признаков стохастической нейронной сетью при решении задачи распознавания образов на основе МБО в данной работе построена нейросетевая модель, осуществляющая распознавание изображений рукописных чисел из базы данных MNIST [10]. Пример распознаваемых изображений из этой базы приведен на рис. 2. Числа на изображениях представлены оттенками серого, фон имеет черный цвет. Обучающая выборка содержит 60 000 изображений размерами  $28 \times 28$  пикселей, разбитых на 10 классов. Тестовая выборка вклю-

чает в себя дополнительные 10 000 изображений.

Исследуемая модель состоит из трех слоев. Первый слой видимых нейронов включает в себя 784 элемента, которые соответствуют изображению с размерами  $28 \times 28 = 784$  пикселей. Изображение подается на вход сети путем задания в качестве вероятности активации каждого видимого нейрона нормированной от 0 до 1 яркости соответствующего пикселя. Вторым слоем является слой скрытых нейронов, количество которых варьировалось в процессе эксперимента. В третьем слое содержатся метки, соответствующие цифрам от 0 до 9.

В данной модели слой скрытых нейронов осуществляет выделение информативных признаков из изображения, а слой меток выступает в роли классификатора. Качество распознавания можно существенно улучшить, если использовать несколько слоев скрытых нейронов, т. е. выделять информативные признаки из информативных признаков. Такие «признаки признаков» сложно интерпретировать, поэтому в данной работе детально исследовался только первый слой скрытых нейронов. Процесс выделения признаков дополнительными слоями скрытых нейронов в многослойной стохастической нейронной сети

полностью аналогичен процессу выделения первичных признаков из исходного изображения.

Исследовалось влияние числа нейронов в скрытом слое на качество распознавания изображений. Наименьшая ошибка классификации, не превышавшая 1,5 %, получена для нейронной сети с 479 нейронами в первом скрытом слое и 300 – во-втором. Помимо этого сеть включала в себя 784 видимых входных нейрона, информация с которых поступала на первый скрытый слой, и 10 видимых нейронов-меток, считывавших информацию со второго скрытого слоя. Выбор числа скрытых нейронов в первом и втором слое осуществлялся с использованием графического представления весов обученной сети.

Для графической интерпретации результата обучения сети построены карты признаков (рис. 3 *в, з*). Карта признаков состоит из областей чувствительности всех скрытых нейронов сети. Область чувствительности нейрона-признака может быть представлена графически следующим образом. Веса  $w_{ij}$  всех связей данного скрытого нейрона с входными нейронами (рис. 3 *а*) нормируются и интерпретируются как яркость соответствующих пикселей области чувствительности

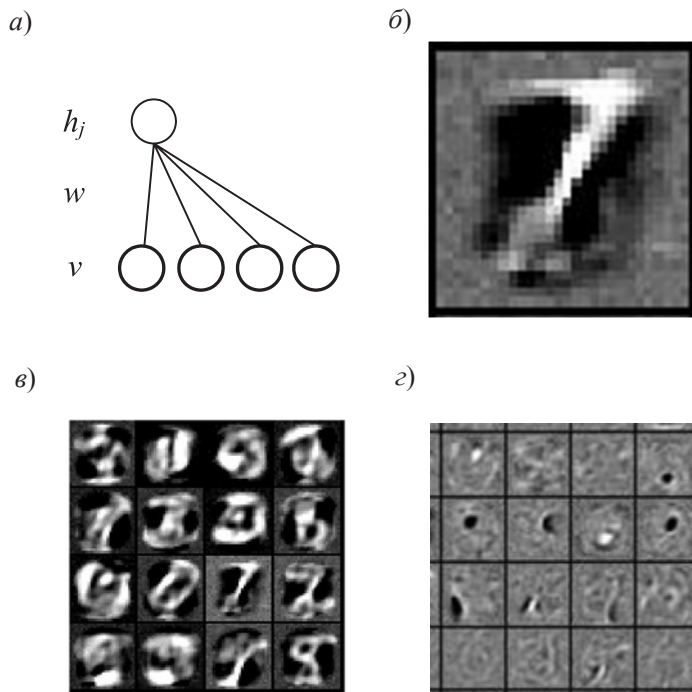


Рис. 3. Графическое представление области чувствительности нейронов:

*а* – скрытый нейрон (эксперт), определяющий наличие во входных данных некоторого признака;  
*б* – область чувствительности скрытого нейрона; *в* – карта признаков нейронной сети с 16 скрытыми нейронами; *з* – фрагмент карты признаков нейронной сети с 400 скрытыми нейронами



(рис. 3 б). Светлые и темные пиксели на рис. 3 б соответствуют областям исходного изображения, которые должны обладать соответственно высокой или низкой интенсивностью для того, чтобы нейрон принял положительное решение о наличии в исходном изображении некоторого признака. Серые пиксели вокруг центральной области на рис. 3 б соответствуют области исходного изображения, которая не несет никакой полезной информации для принятия решения о наличии в изображении признака, за который отвечает данный нейрон. Фактически, область чувствительности нейрона представляет собой маску, которая накладывается на изображение. Чем лучше маска подходит к поданному на вход сети изображению, тем выше вероятность активации скрытого нейрона. В соответствии со своей областью чувствительности, нейрон, представленный на рис. 3 б, будет иметь самую высокую вероятность активации, когда на вход сети подается изображение чисел 2 и 7 или похожее на них.

Для нейронной сети с 16 скрытыми нейронами карту признаков можно представить в виде матрицы размерностью  $4 \times 4$ , состоящей из изображений областей чувствительности всех скрытых нейронов, как показано на рис. 3 в. Каждый элемент матрицы соответствует одному скрытому нейрону-признаку и описывает его область чувствительности, т. е. степень влияния каждого пикселя исходного изображения на принятие решения экспертом о наличии в изображении данного признака.

При увеличении числа скрытых нейронов происходит «размывание» карты признаков. Для сети с 400 скрытыми нейронами область чувствительности любого из них представляет собой серый (нечувствительный) квадрат с небольшим белым или черным пятном (рис. 3 г). Такой нейрон будет активен, если соответствующая пятну область на поданном на вход сети изображении будет закрашена (для нейронов с белым пятном на области чувствительности) или будет являться незакрашенным фоном (для нейронов с черным пятном). Остальные пиксели изображения не

вносят существенного вклада в процесс принятия решения данным нейроном. Нейроны, область чувствительности которых представляет собой сплошной серый квадрат без белых или черных пятен, являются избыточными, поскольку на их активацию не влияет исходное изображение. После обучения первого слоя скрытых нейронов, но перед началом обучения второго, избыточные нейроны могут быть удалены, т. к. обучение многослойной стохастической нейронной сети осуществляется послойно. Это позволит ускорить процесс дальнейшего обучения, повысить информативность внутреннего представления модели и увеличить точность классификации.

Визуальное представление матрицы весов  $W$  в виде карты признаков может использоваться для оценки качества обучения стохастической нейронной сети как после завершения обучения для сравнения между собой нескольких обученных сетей, так и в реальном времени во время обучения. Визуализация области чувствительности скрытых нейронов позволяет лучше понять и интерпретировать признаки изображения, за которые отвечают эти нейроны.

В ходе данного исследования разработан алгоритм визуализации внутреннего состояния стохастической нейронной сети на карте признаков, состоящей из графического представления областей чувствительности нейронов. Данный алгоритм может использоваться для уточнения числа скрытых нейронов в сети, оценки качества ее обучения, что было показано на примере распознавания рукописных чисел.

Использование нескольких скрытых слоев нейронов в машине Больцмана усложняет процесс обучения, но позволяет существенно повысить количество правильно распознанных образов. Обучение многослойной стохастической сети осуществляется послойно, поэтому перед началом обучения нового слоя на основе оценки внутреннего состояния сети целесообразно осуществлять селекцию нейронов-признаков с целью уменьшения их количества.

#### СПИСОК ЛИТЕРАТУРЫ

1. Polyak, M. Data Processing Using Boltzmann Machines [Text] / M. Polyak // XIII International Forum Modern information society formation – problems, perspectives, innovation approaches: Proc. of the Forum. St. Petersburg, 5-10 Sept.: SUAI. –SPb., 2012. –P. 44–50.  
2. Хайкин, С. Нейронные сети: полный курс [Текст]

/ С. Хайкин. –М.: ИД «Вильямс», 2006. –1104 с.

3. Montufar, G. Refinements of Universal Approximation Results for DBNs and RBMs [Text] / G. Montufar, N. Ay // Neural Computation. –2011. –Vol. 23 (5). –P. 1306–1319

4. Уидроу, Б. Адаптивная обработка сигналов

[Текст] / Б. Уидроу, С. Стирнз. –М.: Радио и связь, 1989. –440 с.

5. **Hopfield, J.J.** Neural networks and physical systems with emergent collective computational properties [Text] / J.J. Hopfield // Proc. of the National Academy of Sciences of the USA. –1982. –Vol. 79 (8). –P. 2554–2558.

6. **McCullough, W.S.** A logical calculus of the ideas immanent in nervous activity [Text] / W.S. McCullough, W.H. Pitts // Bulletin of Mathematical Biophysics. –1943. –Vol. 5. –P. 115–133.

7. **Hinton, G.E.** Training Products of Experts by Minimizing Contrastive Divergence [Text] / G.E. Hinton // Neural Computation. –2002. –Vol. 14 (8). –P. 1771–1800.

8. **Woodford, O.** Notes on Contrastive Divergence [Электронный ресурс] / **O. Woodford.** –Режим доступа: <http://www.robots.ox.ac.uk/~ojw/files/NotesOnCD.pdf> (Дата обращения 11.01.2013)

9. **Chen, H.** A Continuous Restricted Boltzmann Machine with an Implementable Training Algorithm [Text] / H. Chen, A.F. Murray // IEE Proc. of Vision, Image and Signal Processing. –2003. –Vol. 150 (3). –P. 153–158.

10. The MNIST Database of handwritten digits by Yann LeCun [Электронный ресурс] / Режим доступа: <http://yann.lecun.com/exdb/mnist/> (Дата обращения 11.01.2013)