

АЛГОРИТМЫ РАСПОЗНАВАНИЯ РУКОПИСНЫХ ПОДПИСЕЙ НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ

Друки А.А., Милешин М.А.

*Национальный исследовательский Томский политехнический университет,
Институт кибернетики, Томск, e-mail: druki2008@yandex.ru*

Данная статья посвящена разработке методов, алгоритмов и программ, обеспечивающих распознавание рукописных подписей на изображениях. В процессе исследований было определено, что для решения поставленной задачи эффективнее использовать нейронные сети. Разработаны две нейронные сети разных видов: сверточная нейронная сеть и классическая нейронная сеть. Было проведено исследование различных конфигураций разработанных нейронных сетей по следующим параметрам: число слоёв, число нейронов в слоях, число связей между нейронами, тип активационной функции, тип алгоритма обучения. Для обучения и тестирования нейронных сетей была создана база данных изображений рукописных подписей. Наилучшие результаты при обучении и тестировании показала сверточная нейронная сеть, состоящая из шести слоёв, при использовании активационной функции гиперболический тангенс и алгоритма обучения Левенберга – Марквардта. Данная сверточная нейронная сеть легла в основу программного обеспечения для распознавания рукописных подписей на изображениях. Разработанное программное обеспечение показывает высокую вероятность точного распознавания при достаточно высокой скорости работы.

Ключевые слова: обработка изображений, искусственный интеллект, распознавание объектов, рукописные подписи, нейронные сети

HANDWRITTEN SIGNATURE RECOGNITION ALGORITHMS BASED ON THE NEURAL NETWORKS

Druki A.A., Mileschin M.A.

*National Research Tomsk Polytechnic University, Institute of Cybernetics,
Tomsk, e-mail: druki2008@yandex.ru*

This article focuses on the development of methods, algorithms and programs for recognition of handwritten signatures on images. During the research, that for solve of this task effectively use neural networks it was determined. Two different types of neural networks were developed: a convolution neural network and a classical neural network. A study of different configurations of neural networks on the following parameters was conducted: the number of layers, number of neurons in layers, the number of connections between neurons, the type of activation function, the type of learning algorithm. The database of handwritten signatures images for training and testing of neural networks was created. The best results of training and testing were obtained by using the convolution neural network consisting of six layers, using a hyperbolic tangent activation function and the learning algorithm of Levenberg-Marquardt. This convolution neural network provided the basis for the software to recognize handwritten signatures on images. The developed software shows a high probability of accurate recognition of handwritten signatures in images at a sufficiently high speed.

Keywords: image processing, artificial intelligence, pattern recognition, handwritten signatures, neural networks

Распознавание изображений находит широкое применение в различных областях: это может быть классификация объектов, распознавание лиц, символов, маркировки объектов, автомобильных номеров знаков, отпечатков пальцев и т.д. На сегодняшний день данные технологии применяются в сотовых телефонах, планшетных компьютерах, системах контроля доступа к секретным объектам, оперативном поиске в картотеке и т.д. Однако до сих пор при решении задач в этой области возникает ряд сложных научных и технических проблем.

В связи с этим можно предположить, что в ближайшем будущем будут актуальны задачи идентификации и распознавания рукописных подписей, т.к. подпись каждого человека может являться важным идентификатором личности. В настоящее время

не существует программного обеспечения, которое решает данную задачу.

Рукописные подписи имеют достаточно сложную структуру и мелкую детализацию, всё это представляет большую сложность для решения данной задачи математическими методами и требует больших вычислительных затрат.

При анализе методов распознавания образов на изображениях было решено использовать искусственные нейронные сети, т.к. они могут выполнять роль классификатора, хорошо моделирующего сложную функцию распределения рукописных подписей на изображениях, тем самым увеличивая точность распознавания по сравнению с остальными методами [1, 2].

Целью работы является разработка алгоритма распознавания рукописных подписей на изображениях, обеспечивающего высокий процент точного распознавания.

Разработка сверточной нейронной сети

В процессе анализа эффективности работы нейронных сетей было предпринято реализовать сверточную нейронную сеть и классическую нейронную сеть, чтобы экспериментально сравнить результаты их работы. Разработанные нейронные сети обучались и тестировались на одних и тех же базах данных изображений.

Архитектура сверточных нейронных сетей адаптирована для работы с двумерными поверхностями и обеспечивает высокую степень инвариантности к пространственным искажениям входных данных, таких как изменение масштаба, смещение, поворот, смена ракурса и т.д. [3].

Входной слой представляет из себя плоскость, размеры которой соответствуют размеру входного изображения. Скрытые слои состоят из набора плоскостей и предназначены для выделения характерных признаков из входного изображения. Каждый нейрон слоя связан с областью нейронов предыдущего слоя (локальные рецепторные поля), причём все нейроны в плоскости имеют одинаковые веса. Это обеспечивает нахождение характерных

признаков в любом месте изображения и уменьшает общее число весовых коэффициентов [4].

С каждым следующим слоем количество плоскостей в слоях увеличивается, а их размеры уменьшаются. На каждом следующем уровне выделяются все более абстрактные признаки классифицируемых объектов [5]. Последние слои состоят из простых сигмоидальных нейронов, которые проводят финальную классификацию признаков, выделенных в предыдущем слое.

Проведено исследование различных конфигураций сверточной нейронной сети по следующим параметрам: число слоёв, число плоскостей в слоях, число связей между слоями, функции активации нейронов. В ходе экспериментальных исследований было установлено, что для решения поставленной задачи эффективнее использовать сверточную нейронную сеть без подвыборочных слоёв, т.к. подвыборочные слои могут привести к потере мелких характерных признаков рукописных подписей. Наилучшее качество классификации было достигнуто при использовании сети с конфигурацией, представленной на рис. 1.

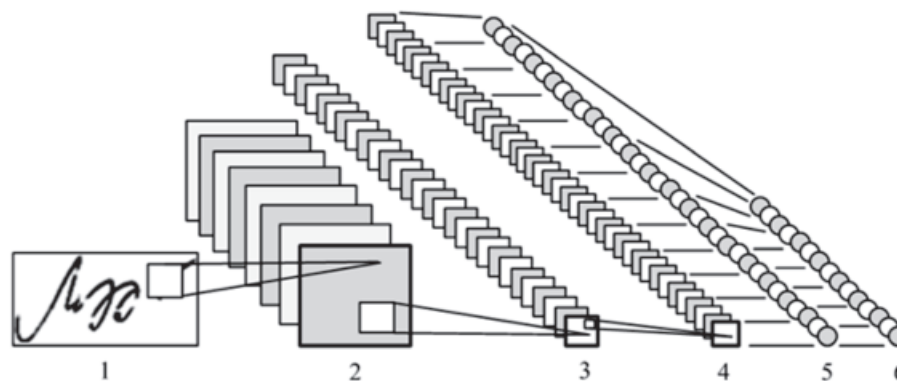


Рис. 1. Архитектура свёрточной нейронной сети для распознавания рукописных подписей:
1 – входной слой; 2, 3 – сверточные слои; 4, 5 – слои из обычных нейронов

Размер входного слоя – 40 на 80 нейронов, что соответствует размеру распознаваемых подписей. Второй слой является свёрточным и состоит из 8 плоскостей размером 20 на 40 нейронов. Третий слой также является свёрточным и состоит из 16 плоскостей размером 12 на 22 нейронов. Четвертый слой также является свёрточным и состоит из 48 плоскостей

размером 6 на 10 нейронов. Пятый слой состоит из 100 обычных сигмоидальных нейронов. Шестой слой является выходным и состоит из 15 нейронов, соответствующих 15 типам рукописных подписей. Размер локального рецептивного поля равен четырем пикселям.

В табл. 1 представлена конфигурация сверточной нейронной сети.

Таблица 1

Конфигурация сверточной нейронной сети

Слой	1	2	3	4	5	6
Число плоскостей	1	8	16	48	100	15
Размер плоскостей	40×80	20×40	12×22	6×10	1	1
Размер рецептивной области	4×4	4×4	4×4	4×4	1	100

Сверточные слои осуществляют фильтрацию изображений прямоугольной маской из общих синаптических коэффициентов нейронов. Результат фильтрации после нелинейного преобразования функцией активации отображается на выходы нейронов [6].

Процесс функционирования нейрона сверточного слоя задается выражением

$$y_k^{(i,j)} = f \left(b_k + \sum_{s=0}^{N-1} \sum_{t=0}^{M-1} w_{k,s,t} x^{((i-1)+s, (j-1)+t)} \right),$$

где $y_k^{(i,j)}$ – нейрон k -й плоскости сверточного слоя; b_k – нейронное смещение k -й плоскости; N, M – размер рецептивной области нейрона; $w_{k,s,t}$ – элемент матрицы, синаптических коэффициентов; x – выходы нейронов предыдущего слоя; f – функция активации нейрона.

Постепенно нейронная сеть обучается выделять ключевые характеристики рукописных подписей в поступающих на вход изображениях [7].

Была протестирована работа двух видов активационных функций: лог-сигмоидная и гиперболический тангенс. Наилучшие результаты показала функция гиперболический тангенс:

$$f(x) = \tanh(x).$$

Эта функция имеет ряд преимуществ для решения задачи:

- симметричные функции обеспечивают более быструю сходимость;
- функция имеет непрерывную первую производную;
- функция имеет простую производную, которая может быть вычислена через ее значение, что дает экономии вычислений [8].

Данная функция имеет интервал $[-1; +1]$, соответственно выходные нейроны сети принимают значения в данном интервале.

Разработка классической нейронной сети

Для сравнения экспериментальных результатов работы сверточной нейронной сети была разработана классическая нейронная сеть: многослойный персептрон. В процессе разработки классической нейронной сети было проведено исследование различных конфигураций сети по следующим параметрам: число слоёв, число нейронов в слоях, число связей между нейронами. Наилучшее качество классификации было достигнуто при использовании сети с конфигурацией, представленной на рис. 2.

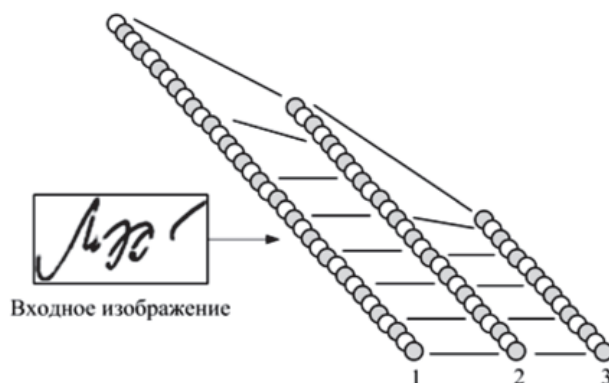


Рис. 2. Архитектура классической нейронной сети для распознавания рукописных подписей: 1 – входной слой; 2 – скрытый слой; 3 – выходной слой

Так как размер изображений рукописных подписей 40 на 80 пикселей, то соответственно входной слой состоит из 3200 (40×80) нейронов. Второй слой состоит из 30 нейронов. Выходной слой состоит из 15 нейронов, соответствующих 15 типам рукописных подписей. Слои являются полностью связанными.

В табл. 2 представлена конфигурация классической нейронной сети.

В качестве активационной функции использовалась функция гиперболический тангенс, т.к. она показала наилучшие результаты в сравнении с лог-сигмоидной функцией.

Таблица 2
Конфигурация классической нейронной сети

Слой	1	2	3
Число нейронов	3200	30	15
Число связей	96000	96450	450

Обучение и тестирование нейронных сетей

Для обучения нейронных сетей была создана база данных из 2000 изображений рукописных подписей, принадлежащих разным людям (рис. 3).



Рис. 3. Образцы рукописных подписей из базы данных

Для тестирования сети также была создана база данных из 2000 изображений рукописных подписей. Базы данных не содержат повторяющихся изображений подписей, каждая подпись является уникальной.

Данные, которые подаются на вход нейронных сетей, должны быть единообразны с точки зрения формата и размера. Для входных изображений производится процедура предварительной обработки согласно следующему алгоритму:

1. Применение сглаживающего фильтра с единичной матрицей 3 на 3 пикселя, чтобы очистить изображение от шумовых помех.
2. Перевод изображения в черно-белый формат с указанием минимального порога яркости.
3. Центрирование рукописной подписи.
4. Изменение размера полученного изображения. Размер входных изображений приводится к фиксированному значению 40 на 80 пикселей.

Обучение нейронных сетей производилось тремя различными алгоритмами: Backpropagation, Quickprop, Левенберга-Марквардта.

Для сверточной нейронной сети лучшее качество обучения было достигнуто при использовании алгоритма Левенберга – Марквардта (табл. 3). К недостаткам данного алгоритма можно отнести долгое время обучения, в несколько раз превышающее остальные алгоритмы. Число неверно классифицированных изображений тестовой выборки равно 32. Время распознавания рукописных подписей сверточной нейронной сетью составляет 21 мс.

В классической нейронной сети алгоритм Левенберга – Марквардта также показал наилучшие результаты в сравнении с другими методами, при наибольшем времени обучения (табл. 4). Число неверно классифицированных изображений тестовой выборки равно 83 изображением. Время распознавания рукописных подписей классической нейронной сетью составляет 40 мс.

Таблица 3

Сравнение алгоритмов обучения сверточной нейронной сети

Алгоритм обучения	Погрешность при обучении	Погрешность при тестировании	Число эпох	Время обучения (часы)
Backpropagation	26,2 %	31,8 %	49	5,1
Quickprop	3,4 %	3,2 %	15	4,8
Левенберга-Марквардта	0,8 %	1,2 %	24	17,7

Таблица 4

Сравнение алгоритмов обучения классической нейронной сети

Алгоритм обучения	Погрешность при обучении	Погрешность при тестировании	Число эпох	Время обучения (часы)
Backpropagation	38,27 %	40,14 %	61	4,5
Quickprop	8,58 %	10,2 %	19	4,1
Левенберга – Марквардта	6,2 %	6,8 %	27	12,3

Из табл. 3 и 4 видно, что при одинаковых алгоритмах обучения и одинаковых обучающих выборках сверточная нейронная сеть показывает лучшие результаты, нежели классическая нейронная сеть. Поэтому можно сделать вывод, что применение классических нейросетевых архитектур в данной задаче неэффективно по следующим причинам:

- изображения имеют большую размерность, вследствие чего увеличивается размер нейронной сети, время и вычислительная сложность процесса обучения;

- игнорируется топология входных данных. Компоненты входного слоя могут быть представлены в любом порядке, без учета цели обучения. Однако изображения имеют строгую двумерную структуру, в которой существует зависимость между пространственно соседними пикселями.

- рукописные подписи каждого человека могут быть написаны с искажениями, а классические нейронные сети достаточно чувствительны к искажениям входного сигнала.

Разработанная сверточная нейронная сеть легла в основу программного обеспечения для распознавания рукописных подписей.

При распознавании выделенной рукописной подписи программная система выводит информацию о её владельце из базы данных: «ФИО», «Должность», «Телефон», «E-mail», «Фото», «Подпись».

Программная система разрабатывалась и тестировалась на процессоре Intel(R) Core(TM)2 Quad CPU 3.00 GHz. Процесс реализации программного обеспечения проходил в среде разработки Visual Studio 2010 на языке C++.

Заключение

В процессе разработки системы распознавания рукописных подписей было реализовано две нейронные сети: сверточная нейронная сеть и классическая нейронная сеть. Было проведено исследование различных конфигураций нейронных сетей по следующим параметрам: число слоёв, число нейронов в слоях, число связей между нейронами, тип активационной функции, тип алгоритма обучения.

Наилучшее качество распознавания показала сверточная нейронная сеть с применением активационной функции гиперболический тангенс и алгоритма обучения Левенберга – Марквардта.

Разработанное программное обеспечение показывает вероятность точного распознавания рукописных подписей на изображениях не менее 98 %. Время распознавания составляет 21 мс.

Список литературы

1. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с.

2. Макаренко А.А., Калайда В.Т. Методика локализации изображения лица для систем видеоконтроля на основе нейронной сети // Известия Томского политехнического университета. – 2006. – Т. 309. – № 8. – С. 113–118.

3. Солдатова О.П., Гаршин А.А. Применение сверточной нейронной сети для распознавания рукописных цифр // Компьютерная оптика. – 2010. – Т. 34. – № 2. – С. 252–259.

4. Калиновский И.А., Спицын В.Г. Алгоритм обнаружения лиц на основе сверточной нейронной сети // Нейрокомпьютеры: разработка, применение. – 2013. – № 10. – С. 48–53.

5. Le Cun Y., Huang F., Bottou L. Learning Methods for Generic Object Recognition with Invariance to Pose and Lighting // Proceedings of CVPR'04. – Washington, DC, USA: IEEE Computer Society, 2004. – P. 97–104.

6. Rowley H.A., Baluja S., Kanade T. Neural network-based face detection // Pattern anal. mach. intell. – Washington, DC, USA: IEEE Computer Society, 1998. – Vol. 20. – P. 23–38.

7. Feraud R., Bernier O., Viallet J., Collobert M. A fast and accurate face detector based on neural networks // IEEE Transactions on pattern analysis and machine intelligence. – Washington, DC, USA: IEEE Computer Society, 2001. – Vol. 23. – № 23. – P. 42–53.

8. Le Cun Y., Kavukcuoglu K., Farabet C. Convolutional Networks and Applications in Vision // International Symposium on Circuits and Systems (ISCAS'10). – Paris, France: IEEE Computer Society, 2010. – P. 253–256.

References

1. Gonzalez R., Woods R. Digital image processing. M.: Technosphere, 2005, 1072 p.

2. Makarenko A.A., Kalayda V.T. Metodika lokalizatsii izobrazheniya litsa dlya sistem videokontrolya na osnove neironnoy seti [The locating face method for video surveillance systems based on neural network]. *Bulletin of the Tomsk Polytechnic University*, 2006, vol. 309, no. 8, pp. 113–118.

3. Soldatova O.P., Garshin A.A. Primenenie svertochnoy neironnoy seti dlya raspoznavaniya rukopisnykh cifr [The application of convolutional neural network for recognition of handwritten digits]. *Computer optics*, 2010, vol. 34, no. 2, pp. 252–259.

4. Kalinovskiy I.A., Spitsyn V.G. Algoritim obnaruzheniya lic na osnove svertochnoy neironnoy seti [Algorithm of face detection based on the convolution neural network]. *Neurocomputers: development and application*, 2013, no. 10, pp. 48–53.

5. LeCun Y., Huang F., Bottou L. Learning Methods for Generic Object Recognition with Invariance to Pose and Lighting. *Proceedings of CVPR'04*. IEEE Press, Washington, DC, USA: IEEE Computer Society, 2004, pp. 97–104.

6. Rowley H.A., Baluja S., Kanade T. Neural network-based face detection. *Pattern anal. mach. intell.* Washington, DC, USA: IEEE Computer Society, 2000, vol. 5, pp. 23–38.

7. Feraud R., Bernier O., Viallet J., Collobert M. A fast and accurate face detector based on neural networks. *Transactions on pattern analysis and machine intelligence*. Washington, DC, USA: IEEE Computer Society, 2002, vol. 3, no. 23, pp. 42–53.

8. LeCun Y., Kavukcuoglu K., Farabet C. Convolutional Networks and Applications in Vision. *International Symposium on Circuits and Systems (ISCAS'10)*. Paris: IEEE, 2010, pp. 253–256.

Рецензенты:

Кориков А.М., д.т.н., профессор, зав. кафедрой автоматизированных систем управления (АСУ), факультет систем управления (ФСО), Томский государственный университет систем управления и радиоэлектроники, г. Томск;

Тузовский А.Ф., д.т.н., профессор, кафедра ОСУ, институт кибернетики, Национальный исследовательский, Томский политехнический университет, г. Томск.

Работа поступила в редакцию 18.11.2014.