

ПРИМЕНЕНИЕ НЕЙРОСЕТЕВЫХ АЛГОРИТМОВ ДЛЯ РАСПОЗНАВАНИЯ РУКОПИСНЫХ ПОДПИСЕЙ НА ИЗОБРАЖЕНИЯХ

Друки А.А., Милешин М.А.

Томский политехнический университет, 634050, Россия, г. Томск, пр. Ленина, 30

E-mail: druki2008@yandex.ru

Введение

В настоящее время многие научные, технические и производственные направления в значительной степени ориентируются на развитие систем обработки и распознавания изображений. Однако, до сих пор при решении задач в этой области возникает ряд сложных научных, технических и технологических проблем.

Распознавание изображений находит широкое применение в различных областях – распознавание лиц, объектов, символов, автомобильных номеров, контроль доступа к информации по идентификации личности, оперативный поиск в картотеке изображений, дактилоскопия и др. [1].

В этот круг задач так же можно включить распознавание рукописных подписей с целью идентификации их владельца. В настоящее время не существует программного обеспечения, которое решает данную задачу, хотя эта задача является достаточно актуальной.

Целью работы является разработка алгоритмов, позволяющих распознавать рукописные подписи и идентифицировать их владельца.

Анализ существующих методов распознавания объектов на изображениях показал, что данную задачу сложно решить с помощью классических математических методов, т.к. это требует больших вычислительных затрат. Поэтому для решения поставленной задачи было предпринято использовать искусственные нейронные сети, в связи с тем, что они обеспечивают возможность получения классификатора, хорошо моделирующего сложную функцию распределения рукописных подписей на изображениях, тем самым, увеличивая точность решения по сравнению с остальными методами.

Так же преимущество использования нейронных сетей для распознавания объектов заключается в обучаемости системы для выделения ключевых характеристик объектов на изображениях из учебных наборов [2, 3].

Разработка сверточной нейронной сети для распознавания рукописных подписей

Наиболее часто в задачах распознавания и идентификации изображений используются классические нейросетевые архитектуры (многослойный персептрон, сети с радиально-базисной функцией и др.), но из анализа многих работ и экспериментальных исследований следует, что применение классических нейросетевых архитектур в

данной задаче неэффективно по следующим причинам:

- изображения имеют большую размерность, соответственно возрастает размер нейронной сети;
- большое количество параметров увеличивает время и вычислительную сложность процесса обучения [4].

Поэтому для решения поставленной задачи были выбраны сверточные нейронные сети, т. к. они обеспечивают частичную устойчивость к вышеперечисленным недостаткам [5].

Разработана сверточная нейронная сеть, состоящая из 5 слоёв (рис. 1).

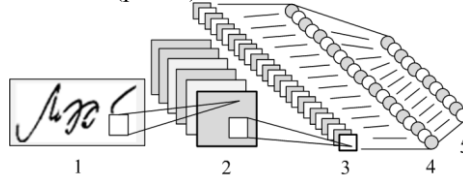


Рис. 1. Архитектура свёрточной нейронной сети для распознавания рукописных подписей: 1) вход; 2, 3) сверточные слои; 4, 5) слои из обычных нейронов

Размер входного слоя 40 на 80 нейронов, что соответствует размеру распознаваемых подписей.

Следом за входным слоем располагается сверточный слой, который состоит из 6 плоскостей размером 20 на 40 нейронов.

Третий слой так же является сверточным и состоит из 50 плоскостей размером 6 на 12 нейронов.

Четвёртый слой состоит из 100 обычных сигмоидальных нейронов.

Пятый слой является выходным и состоит из 20 сигмоидальных нейронов, соответствующих 20 типам рукописных подписей.

В ходе экспериментальных исследований было установлено, что в данной задаче эффективно использовать локальное рецептивное поле равное четырем пикселям.

Выходные нейроны принимают значения в интервале $[-1; +1]$ в соответствии с выбранной активационной функцией (1).

Активационная функция гиперболический тангенс:

$$f(a) = \text{Atanh}(Sa).$$

Формула нейрона сверточного слоя:

$$y_k^{(i,j)} = b_k + \sum_{s=1}^K \sum_{t=1}^K w_{k,s,t} x^{((i-1)+s, (j-1)+t)}.$$

Формула нейрона подвыборочного слоя:

$$y_k^{(i,j)} = b_k + \frac{1}{4} w_k \sum_{s=1}^2 \sum_{t=1}^2 x^{((i,j)+s,(i,j))}.$$

Для обучения сети используется алгоритм обратного распространения ошибки. Для измерения погрешности обучения используется средняя квадратичная ошибка:

$$E = \frac{1}{2} \sum_j \sum_s (y_j^s - d_j^s)^2.$$

Окончательная коррекция синаптических коэффициентов происходит по формуле:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_{pj} o_{pi},$$

где η – коэффициент пропорциональности, влияющий на скорость обучения.

На каждой итерации алгоритм обратного распространения ошибки рассчитывается для всего обучающего набора данных, чтобы вычислить средний или истинный градиент.

В качестве весовых коэффициентов устанавливаются случайные значения [6].

Функция ошибки представляет собой разность между текущим выходом сети и идеальным выходом, который необходимо получить. Для успешного обучения сети требуется приблизить выход сети к желаемому выходу, т. е. последовательно уменьшать величину функции ошибки. Это достигается настройкой межнейронных связей. Каждый нейрон в сети имеет свои веса, которые настраиваются, чтобы уменьшить величину функции ошибки [7].

Значения весовых коэффициентов были выбраны случайным образом из нормального распределения с нулевым средним и стандартным отклонением:

$$\delta_w = \sqrt{m},$$

где m – число связей входящих в нейрон.

Предварительная обработка данных

Для обучения сети была создана база данных из 1000 изображений рукописных подписей, принадлежащих разным людям (рис. 2).



Рис. 2. Образцы рукописных подписей из базы данных

Данные, которые подаются на вход нейронной сети, должны быть единообразны с точки зрения формата, размера и графического содержания. Для этого изображение подписи должно быть обработано согласно следующему алгоритму:

- 1) Применение сглаживающего фильтра, чтобы очистить изображение от шумовых помех.
- 2) Перевод изображения в черно-белый цвет.
- 3) Центрирование рукописной подписи.
- 4) Изменение размера полученного изображения.

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

Заключение

На основе представленных алгоритмов была разработана программная система, которая обеспечивает вероятность точного распознавания рукописных подписей на изображениях не менее 94%. Время распознавания составляет 35 миллисекунд.

Список литературы

1. Kozin, N.E. Gradual learning the radial neural networks // Computer Optics. – 2004. – № 26. – P. 138-141.
2. Le Cun Y., Bengio Y. Convolutional networks for images, speech and time series// The handbook of brain theory and neural networks. – 1998. – V. 7. – № 1. – P. 255–258.
3. Bolotova J.A., Spitzyn V.G., Fomin A.E. The hierarchical temporal memory model application for image recognition // News of Tomsk Polytechnic University. – 2011. – V. 318. – № 5. – P. 60–63.
4. LeCun, Y. Gradient Based Learning Applied to Document Recognition / IEEE Press. – 1998. – P. 46.
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. Kermani Kolankeh A., Spitzyn V.G., Hamker F. The finding parameters and remove the constant component of Gabor filter for image processing // News of Tomsk Polytechnic University. – 2011. – V. 318. – № 5. – P. 57–59.
7. Haykin, S. Neural Networks – a comprehensive foundation – Moscow: Williams, 2006. – 1104 p.