

УДК 004.931

ПРИМЕНЕНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ВЫДЕЛЕНИЯ И РАСПОЗНАВАНИЯ АВТОМОБИЛЬНЫХ НОМЕРНЫХ ЗНАКОВ НА ИЗОБРАЖЕНИЯХ СО СЛОЖНЫМ ФОНОМ

Друки Алексей Алексеевич,аспирант кафедры вычислительной техники Института кибернетики ТПУ,
Россия, 634050, г. Томск, пр. Ленина, д. 30. E-mail: druki2008@yandex.ru

Актуальность исследования обусловлена необходимостью разработки методов, алгоритмов и программ, обеспечивающих повышение эффективности распознавания автомобильных номерных знаков на изображениях со сложным фоном.

Цель работы: повышение эффективности распознавания автомобильных номерных знаков на изображениях со сложным фоном за счёт создания методов, алгоритмов и программ, инвариантных к смещениям, поворотам, смене ракурса и прочим искажениям входных данных.

Методы исследования: для решения поставленных задач использованы методы искусственного интеллекта, идентификации и распознавания образов на изображениях, теории искусственных нейронных сетей, сверточные нейронные сети, эволюционные алгоритмы, математическое моделирование, теория вероятности и математическая статистика с использованием программного обеспечения Visual Studio и MatLab.

Результаты: разработана программная система, позволяющая распознавать автомобильные номерные знаки на изображениях со сложным фоном. Для выделения области расположения символов на изображении предложена сверточная нейронная сеть, состоящая из 7 слоев. Для выделения отдельных символов используется алгоритм, основанный на гистограммах средней интенсивности пикселей. Для распознавания символов реализована сверточная нейронная сеть, состоящая из 6 слоев. Представленная программная система позволяет распознавать автомобильные номерные знаки под большими углами наклона по вертикали, горизонтали и на плоскости при достаточно высоком быстродействии.

Ключевые слова:

Обработка изображений, искусственный интеллект, распознавание символов, нейронные сети, гистограммы средней интенсивности.

Введение

Одной из самых сложных задач в информационных технологиях является обработка и распознавание объектов на изображениях. О важности этой проблемы говорит тот факт, что исследования по распознаванию объектов, анализу изображений и речи включены в перечень приоритетных направлений развития науки и техники и критических технологий федерального уровня [1].

Современные методы распознавания символов на изображениях используются для решения широкого круга задач, таких как распознавание текста, изображений маркировки на поверхностях различных объектов и др.

Процесс распознавания символов на изображениях со сложным фоном можно разделить на 3 этапа:

1. Выделение области расположения символов на изображении.
2. Выделение отдельных символов.
3. Распознавание символов.

В настоящее время такие технологии реализуются тремя традиционными методами – структурным, признаковым и шаблонным. Каждый из этих методов ориентирован на свои условия применения, для которых они являются эффективными. Вместе с тем всем этим методам присущи недостатки. При регистрации изображений наибольшие искажения, влияющие на результат распознавания, вносят аффинные и проекционные искажения, возникающие в связи с изменением угла регистрации, изменением масштаба, погодными условиями. Также наличие посторонних объектов на изображениях со сложным фоном существенно

снижают надежность распознавания методами, используемыми в современных системах распознавания автомобильных номерных знаков на изображениях и видеопоследовательностях [2, 3].

Эти недостатки особенно ярко проявились при масштабной эксплуатации программно-технологических систем, использующих в своей основе эти методы. Практически у всех систем распознавания символов на изображениях точностные характеристики резко падают и становятся ниже технологически приемлемых при различных искажениях входного изображения. Вместе с тем технологические условия получения информации о маркировке не позволяют полностью устранить эти искажения [4, 5].

Анализ методов выделения и распознавания объектов на изображениях показал, что для решения данной задачи эффективно использовать искусственные нейронные сети в связи с тем, что они являются слабо чувствительными к искажениям входного сигнала, а также обеспечивают возможность получения классификатора, хорошо моделирующего сложную функцию распределения изображений символов, тем самым увеличивая скорость и точность решения задачи по сравнению с остальными методами.

Разработка сверточной нейронной сети для выделения области расположения символов на изображениях

Наиболее часто в задачах распознавания и идентификации изображений используются классические нейросетевые архитектуры (многослой-

ный перцептрон, сети с радиально-базисной функцией и др.), но из анализа данных 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 [6];
- большое количество параметров увеличивает вместимость системы и соответственно требует большей тренировочной выборки, увеличивает время и вычислительную сложность процесса обучения;
- для повышения эффективности работы системы желательно применять несколько нейронных сетей (обученные с различными начальными значениями синаптических коэффициентов и порядком предъявления образов), но это увеличивает вычислительную сложность решения задачи и время выполнения [7, 8];
- отсутствует инвариантность к изменениям масштаба изображения, ракурсов съёмки камеры и других геометрических искажений входного сигнала [1].

Поэтому для решения задачи выделения области расположения символов были выбраны сверточные нейронные сети, т. к. они обеспечивают частичную устойчивость к изменениям масштаба, смещениям, поворотам, смене ракурса и прочим искажениям.

Каждый слой сверточной нейронной сети представляет собой набор плоскостей, состоящих из нейронов. Нейроны одной плоскости имеют одинаковые синаптические коэффициенты, ведущие ко всем локальным участкам предыдущего слоя. Каждый нейрон слоя получает входы от некоторой области предыдущего слоя (локальное рецептивное поле), т. е. входное изображение предыдущего слоя сканируется небольшим окном и пропускается сквозь набор синаптических коэффициентов, а результат отображается на соответствующий нейрон текущего слоя. Таким образом, набор плоскостей представляет собой карты характеристик, и каждая плоскость находит «свои» участки изображения в любом месте предыдущего слоя. Размер локального рецептивного поля выбирается самостоятельно в процессе разработки нейронной сети [9].

Слои делятся на 2 типа: сверточные и подвыборочные. В сверточных слоях при сканировании рецептивные поля частично наслаиваются друг на друга по принципу черепицы, в подвыборочных слоях области соседних нейронов не перекрываются. Подвыборочный слой уменьшает масштаб плоскостей путём локального усреднения значений выходов нейронов, таким образом достигается иерархическая организация. Последующие слои привлекают более общие характеристики, меньше зависящие от искажений изображения. После прохождения нескольких слоев карта признаков выражается в вектор [10, 11].

Постепенно нейронная сеть обучается выделять ключевые характеристики автомобильных номерных знаков в поступающих на вход изображениях.

Для выделения области расположения символов на изображениях была разработана сверточная нейронная сеть, состоящая из 7 слоёв (рис. 1).

Входной слой размером 28×44 нейронов состоит из 1232 нейронов, не несет какой-либо функциональной нагрузки и служит лишь для подачи входного образа в нейронную сеть.

Следом за входным слоем находится первый скрытый слой, который является свёрточным. Этот слой состоит из 6 свёрточных плоскостей. Размер каждой плоскости этого слоя равен $24 \times 40 = 960$ нейронов.

Второй скрытый слой является подвыборочным, также состоит из 6 плоскостей, каждая из которых имеет синаптическую маску размером 2×2 . Размер каждой плоскости этого слоя $12 \times 20 = 240$ нейронов, что вдвое меньше, чем размер плоскости предыдущего слоя.

Третий скрытый слой является слоем свёртки. Он состоит из 18 плоскостей размером $16 \times 8 = 128$ нейронов.

Четвертый скрытый слой является подвыборочным и состоит из 18 плоскостей размером $4 \times 12 = 48$ нейронов.

Пятый скрытый слой состоит из 18 простых сигмоидальных нейронов, по одному на каждую плоскость предыдущего слоя. Роль этого слоя состоит в обеспечении классификации, после того, как выполнено извлечение особенностей и сокращение размерности входа. Каждый нейрон этого слоя полностью связан с каждым нейроном только одной плоскости предыдущего слоя.

Шестой слой является выходным слоем. Он состоит из одного нейрона, который полностью связан со всеми нейронами предыдущего слоя.

В соответствии с решаемой задачей в структуре нейронной сети достаточно одного выхода. Выходное значение нейронной сети находится в интервале $[-1; +1]$, что соответственно означает наличие или отсутствие автомобильного номерного знака на классифицируемом изображении [12, 13].

Таким образом, при сканировании входного изображения отклики нейронной сети образуют максимумы в местоположениях автомобильных номерных знаков. Отклики находятся в диапазоне $[-1; +1]$, в соответствии с выбранной активационной функцией [14, 15].

Размер синаптической маски в сверточных слоях равен 5×5 нейронов, в подвыборочных слоях – 2×2 нейрона.

Использование принципа объединения синаптических коэффициентов дает эффект уменьшения количества настраиваемых параметров нейронной сети.

Размер свёрточной плоскости определяется в соответствии со следующим выражением:

$$w_c = w_u - K + 1, \quad (1)$$

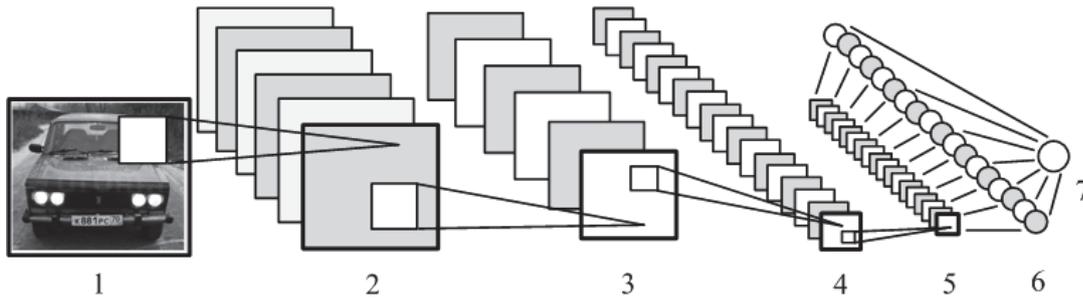


Рис. 1. Архитектура свёрточной нейронной сети для выделения области расположения символов: 1) вход; 2, 4) сверточные слои; 3, 5) подвыборочные слои; 6, 7) слои из обычных нейронов

$$h_c = h_u - K + 1, \quad (2)$$

где w_c, h_c – ширина и высота свёрточной плоскости соответственно; w_u, h_u – ширина и высота плоскости предыдущего слоя; K – ширина (высота) окна сканирования.

В качестве активационной функции был выбран гиперболический тангенс:

$$f(a) = A \tanh(Sa). \quad (3)$$

где $f(a)$ – искомое значение элемента, a – взвешенная сумма сигналов предыдущего слоя, A – амплитуда этой функции, S – определяет ее положение относительно начала отсчета.

Данная функция активации является нечётной, с горизонтальными асимптотами $+A$ и $-A$.

Эта функция имеет ряд преимуществ для решения задачи:

- симметричные активационные функции, типа гиперболического тангенса, обеспечивают более быструю сходимость, чем стандартная логистическая функция;
- функция имеет непрерывную первую производную;
- функция имеет простую производную, которая может быть вычислена через ее значение, что дает экономию вычислений.

Формула функционирования нейрона сверточного слоя:

$$y_k^{(i,j)} = b_k + \sum_{s=1}^K \sum_{t=1}^K w_{k,s,t} x^{((i-1)+s, (j-1)+t)}, \quad (4)$$

где $y_k^{(i,j)}$ – нейрон k -й плоскости свёрточного слоя; b_k – нейронной смещение k -й плоскости; K – размер рецептивной области нейрона; $w_{k,s,t}$ – матрица синаптических коэффициентов; x – выходы нейронов предыдущего слоя.

Формула функционирования нейрона подвыборочного слоя:

$$y_k^{(i,j)} = b_k + \frac{1}{4} w_k \sum_{s=1}^2 \sum_{t=1}^2 x^{((i,j)+s, (i,j))}. \quad (5)$$

Используется стандартный для нейронных сетей алгоритм обратного распространения ошибки. Для измерения качества распознавания использовалась функция среднеквадратической ошибки:

$$E_p = \frac{1}{2} \sum_j (t_{pj} - o_{pj})^2, \quad (6)$$

где E_p – величина функции ошибки для образа p ; t_{pj} – желаемый выход нейрона j для образа p ; o_{pj} – действительный выход нейрона j для образа p .

Окончательная коррекция синаптических коэффициентов происходит по формуле:

$$w_{ij}(t+1) = w_{ij}(t) + \eta \delta_{pj} o_{pj}, \quad (7)$$

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.

Когда ненастроенной сети предъявляется входной образ, она выдает некоторый случайный выход. 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.

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.

Для обучения сети была создана база данных из 1000 изображений автомобильных номерных знаков.

На рис. 2 представлены некоторые изображения автомобильных знаков из обучающей выборки, расположенные фронтально относительно регистрирующего устройства. Изображения получены при различных погодных условиях, времени суток и различном освещении и контрасте.

Для создания большего количества примеров и увеличения степени инвариантности нейронной сети к различным поворотам был создан набор изображений автомобильных номерных знаков, расположенных под различными углами по вертикали, горизонтали и на плоскости относительно регистрирующего устройства (рис. 3). Изображения получены при различных погодных условиях, времени суток, различном освещении и контрасте.



Рис. 2. Изображения автомобильных номерных знаков из обучающей выборки, расположенных фронтально относительно регистрирующего устройства



Рис. 3. Изображения автомобильных номерных знаков из обучающей выборки, расположенных под углами относительно регистрирующего устройства

Выделение отдельных символов с помощью гистограмм средней интенсивности

После выделения области расположения символов на изображении необходимо выделить отдельные символы, для дальнейшего их распознавания. Для этого предлагается использовать метод, основанный на построении гистограмм средней интенсивности.

Выделенная на предыдущем этапе область расположения символов сканируется попиксельно слева направо, сверху вниз, при этом вычисляется средняя интенсивность пикселей в каждом столбце. В тех местах, где нет символа, средняя интенсивность будет значительно отличаться от интенсивности тех мест, где символы присутствуют. Далее, выполняя ту же операцию по строкам, получается набор отдельных символов, которые уже можно анализировать [16, 17].

Для того чтобы отделить строку символов от всего изображения, для начала предлагается вычислить горизонтальные гистограммы. Так как на изображении самой яркой областью является фон номерной пластины, то два наибольших максимума будут соответствовать областям 1 и 2 (рис. 4).



Рис. 4. Построение горизонтальных гистограмм. Линиям 1 и 2 соответствуют два наибольших максимума. x – номер строки изображения, y – средняя интенсивность строки изображения

Далее строятся вертикальные гистограммы под углом взаимно перпендикулярным к n , и выделяется уже около 10 максимумов в промежутках между символами. Таким образом, выделяются области расположения отдельных символов на номерной пластине (рис. 5).

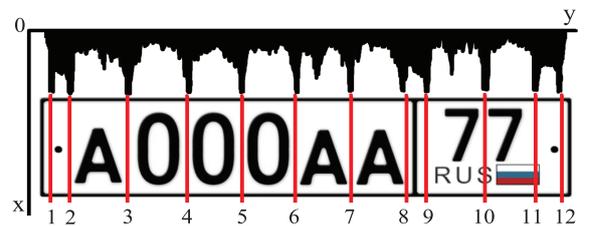


Рис. 5. Построение вертикальных гистограмм. Линиям от 1 до 12 соответствуют наибольшие максимумы. y – номер строки изображения, x – средняя интенсивность строки изображения

При регистрации изображение номерных пластин подвергается различным смещениям и искажениям, следовательно, линии, соответствующие областям 1 и 2, будут располагаться не по горизонтали, а под неизвестным углом. В связи с этим предлагается строить не одну, а n гистограмм средней интенсивности, каждая из которых строится не по горизонтали, а под заданным углом [18].

Необходимое количество гистограмм средней интенсивности определяется из технических условий регистрации изображений, по данным условиям угол разворота изображения не превышает 20° по горизонтали как в правую, так и в левую сторону, следовательно, $n=41$. Из n построенных гистограмм выбирается та, которая содержит наибольшее значение по y , т. к. наибольшее значение будет соответствовать области 1 или 2 (рис. 4) [19, 20].

Разработка сверточной нейронной сети для распознавания символов на изображениях

Для распознавания выделенных символов была разработана сверточная нейронная сеть с 4 скрытыми слоями (рис. 6).

Первый слой является входным и состоит из $28 \times 28 = 841$ нейрона.

Второй слой является свёрточным и состоит из шести плоскостей размером $24 \times 24 = 576$ нейронов.

Размер свёрточной плоскости определяется в соответствии с формулами (1, 2).

Третий слой является подвыборочным и также состоит из 5 плоскостей размером $12 \times 12 = 144$ нейронов.

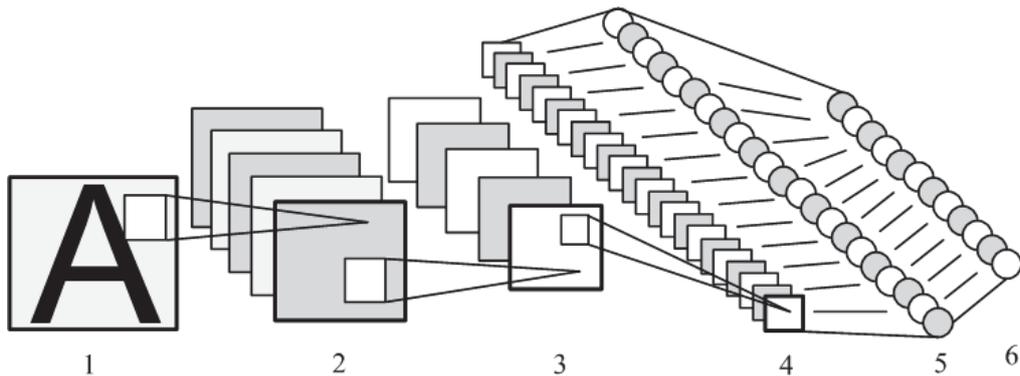


Рис. 6. Архитектура свёрточной нейронной сети для распознавания символов: 1) вход; 2, 4) свёрточные слои; 3) подвыборочный слой; 5, 6) слои из обычных нейронов

Четвёртый слой является свёрточным и состоит из 50 плоскостей размером $8 \times 8 = 64$ нейронов.

Пятый слой состоит из 126 простых сигмоидальных нейронов. Роль этого слоя – обеспечение классификации, после того, как выполнено извлечение особенностей и сокращение размерности входных данных.

Последний, шестой, слой является выходным слоем и состоит из 21 нейрона.

Согласно ГОСТ Р 50577–93 регистрационные знаки транспортных средств могут содержать следующие символы: А, В, Е, К, М, Н, О, Р, С, Т, Х, У и все цифры от 0 до 9. Поэтому выходной слой состоит из 21 нейрона, так как распознается 21 символ.

В качестве активационной функции был выбран гиперболический тангенс (3), формула функционирования нейрона свёрточного слоя (4), формула функционирования нейрона подвыборочного слоя (5). Для обучения сети используется алгоритм обратного распространения ошибки (6). Коррекция синаптических коэффициентов происходит по формуле (7).

Для обучения сети была использована база из 60000 изображений рукописных цифр (MNIST) и создана база из 20000 изображений букв. Размер тестирующей выборки – 10000 символов.

В табл. 1 отображены значения вероятности верного распознавания символов в зависимости от изменения углов регистрации.

На основе представленных алгоритмов разработана программная система, которая обеспечивает вероятность распознавания автомобильных номерных знаков на изображениях со сложным фоном не менее 98 % при следующих условиях регистрации:

- время обработки: 35 миллисекунд;
- высота символов не менее 12 пикселей;
- освещенность пластины номерного знака в диапазоне от 50 до 1000 лк;
- угол отклонения пластины номерного знака по горизонтали относительно регистрирующего устройства до $\pm 60^\circ$;
- угол отклонения пластины номерного знака по вертикали относительно регистрирующего устройства до $\pm 65^\circ$;

- угол поворота пластины номерного знака на плоскости до $\pm 20^\circ$.

Таблица 1. Вероятность верного распознавания символов в зависимости от изменения углов регистрации

Угол отклонения	Вероятность распознавания, %		
	по вертикали	по горизонтали	на плоскости
-75	87	85	–
-70	98	93	–
-65	100	99	–
-60	100	100	–
-55	100	100	–
-50	100	100	–
-45	100	100	–
-40	100	100	–
-35	100	100	–
-30	100	100	–
-25	100	100	–
-20	100	100	100
-15	100	100	100
-10	100	100	100
-5	100	100	100
0	100	100	100
5	100	100	100
10	100	100	100
15	100	100	100
20	100	100	100
25	100	100	–
30	100	100	–
35	100	100	–
40	100	100	–
45	100	100	–
50	100	100	–
55	100	100	–
60	100	100	–
65	100	98	–
70	99	92	–
75	88	86	–

Сравнение технических характеристик систем распознавания автомобильных номерных знаков

В табл. 2, 3 приведено сравнение технических характеристик разработанной программной системы с существующими на рынке системами распознавания автомобильных номеров.

Таблица 2. Сравнение технических характеристик систем распознавания автомобильных номеров

Название системы	Вероятность распознавания, %	Время распознавания	Освещенность, лк	Минимальная высота номера в кадре
«Авто-инспектор»	95	не указано	не менее 50	не указано
«Авто-интеллект»	90	не указано		не указано
«SL-Traffic»	90	не указано		25 пикселей
«Дигнум- авто»	90	не указано		не указано
«CarFlow II»	93...98	60 мс		не указано
Разработанное ПО	98	35 мс	от 50 до 1000	12 пикселей

Таблица 3. Возможности распознавания систем при различных углах отклонения пластины автомобильного номерного знака

Название системы	Угол отклонения, град.		
	Вертикальный	Горизонтальный	На плоскости
«Авто-инспектор»	±40	±30	±15
«Авто-интеллект»	±30	±20	±10
«SL-Traffic»	±60	±30	±10
«Дигнум- авто»	±40	±40	не указано
«CarFlow II»	±45	±45	±15
Разработанное ПО	±65	±60	±20

СПИСОК ЛИТЕРАТУРЫ

- 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.
- Елизаров А.И., Афонасенко А.В. Методика построения систем распознавания автомобильного номера // Известия Томского политехнического университета. – 2006. – Т. 309. – № 8. – С. 118–121.
- 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.
- Болотова, Ю.А., Спицын В.Г. Сравнение способов обучения модели НТМ для задачи распознавания цифр // Молодежь и современные информационные технологии: сборник трудов IX Всероссийской научно-практической конференции студентов, аспирантов и молодых ученых. – Томск: Изд-во СПБ Графикс, 2011. – Т. 1. – С. 252–253.
- Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072 с.
- Болотова, Ю.А., Спицын В.Г., Кермани А.К. Распознавание символов на цветном фоне на основе иерархической временной модели с предобработкой фильтрами Габора // Электромагнитные волны и электронные системы. – 2012. – Т. 16. – № 1. – С. 14–19.
- Bundzel M., Hashimoto S. Object identification in dynamic images based on the memory-prediction theory of brain function // Journal of Intelligent Learning Systems and Applications. – 2010. – V. 2. – № 4. – P. 212–220.
- Буй Тхи Тху Чанг, Фан Нгюк Хоанг, Спицын В.Г. Алгоритмическое и программное обеспечение для классификации цифровых изображений с помощью вейвлет-преобразования Хаара и нейронных сетей // Известия Томского политехнического университета. – 2011. – Т. 319. – № 5. – С. 103–106.

Как видно из табл. 2, разработанная программная система по всем параметрам не уступает существующим на рынке системам, а по некоторым параметрам превосходит.

Как видно из табл. 3, разработанная программная система способна распознавать автомобильные номерные знаки при самых больших углах отклонения в сравнении с другими системами.

Заключение

Для существенного повышения эффективности выделения и распознавания автомобильных номерных знаков на изображениях со сложным фоном предложено использовать сверточные нейронные сети, т. к. они являются устойчивыми к смещениям, поворотам, смене ракурса и прочим искажениям входных данных.

Разработана сверточная нейронная сеть, обеспечивающая нахождение и выделение области расположения символов на изображениях со сложным фоном. Для выделения отдельных символов предложено использовать алгоритм, основанный на построении гистограмм средней интенсивности пикселей. Для распознавания отдельных символов разработана сверточная нейронная сеть, работающая как классификатор символов. Проведена оптимизация структуры разработанных нейронных сетей, что позволило повысить эффективность их работы.

- Болотова Ю.А., Спицын В.Г., Фомин А.Э. Применение модели иерархической временной памяти в распознавании изображений // Известия Томского политехнического университета. – 2011. – Т. 318. – № 5. – С. 60–63.
- Кермани К.А., Спицын В.Г., Хамкер Ф. Нахождение параметров и удаление постоянной составляющей фильтра Габора для обработки изображений // Известия Томского политехнического университета. – 2011. – Т. 318. – № 5. – С. 57–59.
- Hansen D.W., Hansen J.P., Nielsen M. Eye typing using Markov and active appearance models // Applications of computer vision. – 2002. – V. 12. – P. 132–136.
- Rowley H.A., Baluja S., Kanade T. Neural network-based face detection // Pattern anal. mach. intell. – 1998. – V. 20. – P. 23–38.
- 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.
- George D., Hawkins J. A hierarchical bayesian model of invariant pattern recognition in the visual cortex // Proceedings. 2005 IEEE International Joint Conference on Neural Networks. – Montreal, Canada: IEEE Computer Society, 2005. – V. 3. – P. 1812–1817.
- Chapelle O., Haffner P., Vapnik V. Support vector machines for histogram-based image classification // IEEE Trans. on Neural Networks & Learning Systems – 1999. – V. 10. – № 5. – P. 1055–1064.
- 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. – 2001. – V. 23. – № 23. – P. 42–53.
- Karnowski T., Arel I., Rose D. Deep Spatiotemporal Feature Learning with Application to Image Classification // The 9th International conference on machine learning and applications (ICMLA'10). – 2010. – P. 883–888.

18. Кульчин Ю.Н., Ноткин Б.С., Седов В.А. Нейро-итерационный алгоритм томографической реконструкции распределенных физических полей в волоконно-оптических измерительных системах // Компьютерная оптика. – 2009. – Т. 33. – № 4. – С. 446–455.
19. Lai Z., Hongbin G., Ben N. Visual Hand Pose Estimation Based on Hierarchical Temporal Memory in Virtual Reality Cockpit Simulator // Information Technology Journal. – 2011. – V. 10. – № 9. – P. 1809–1816.
20. 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.

Поступила 05.12.2013 г.

UDC 004.931

APPLICATION OF CONVOLUTIONAL NEURAL NETWORKS FOR EXTRACTION AND RECOGNITION OF CAR NUMBER PLATES ON IMAGES WITH COMPLEX BACKGROUND

Aleksey A. Druki,

Tomsk Polytechnic University, 30, Lenin avenue, Tomsk, 634050, Russia.

E-mail: druki2008@yandex.ru

The urgency of the discussed issue is caused by the need to develop methods, algorithms and programs to ensure efficiency of car number plate recognition on images with a complex background.

The main aim of the study: *to increase the efficiency of character recognition on images with a complex background by developing methods, algorithms and programs invariant to affine and projective transformations of the input data.*

The methods used in the study: *To solve the task the author has used the methods of the artificial Intelligence, identification and pattern recognition in images, theory of artificial neural networks, convolutional neural networks, evolutionary algorithms, mathematical modeling, probability theory and mathematical statistics with the help of software Visual Studio and MatLab.*

The results: *The author developed the software allowing the recognition of car number plates on images with a complex background. The seven-layer convolutional neural network for character area selection on images is proposed. The algorithm based on the average pixel intensity histograms for individual characters selection is used. The six-layer convolutional neural network for character recognition on images is implemented. The represented software system can recognize license plates with deviation horizontally, vertically and in a plane angles. The system has high speed work.*

Key words:

Image processing, artificial intelligence, character recognition, neural networks, histogram of average intensity.

REFERENCES

1. Makarenko A.A., Kalayda V.T. Metodika lokalizatsii izobrazhenia litsa dlya sistem videokontrolya na osnove neyronnoy seti [The locating face method for video surveillance systems based on neural network]. *Bulletin of Tomsk Polytechnic University*, 2006, vol. 309, no. 8, pp. 113–118.
2. Elizarov A.I., Afonassenko A.V. Metodika postroeniya sistem raspoznavaniya avtomobilnogo nomera [The constructing method of license plate recognition systems]. *Bulletin of Tomsk Polytechnic University*, 2006, vol. 309, no. 8, pp. 118–121.
3. Le Cun Y., Bengio Y. Convolutional networks for images, speech and time series. *The handbook of brain theory and neural networks*, 1998, vol. 7, no. 1, pp. 255–258.
4. Bolotova Yu.A., Spitzyn V.G. Sravnenie sposobov obucheniya modeli HTM dlya zadachi raspoznavaniya tsifr [Comparing ways of HTM model learning for digits recognition]. *Sbornik trudov IX Vserossiiskoy nauchno-prakticheskoy konferentsii studentov, aspirantov i molodykh uchenykh «Molodezh i sovremennye informatsionnye tekhnologii»* [Proceedings of the IX Russian scientific and practical conference of students, graduate students and young scientists. Youth and modern information technology]. Tomsk, Publ. SPB Grafiks, 2011, vol. 1, pp. 252–253.
5. Gonzalez R., Woods R. Digital image processing. Moscow, Technosphere, 2005. 1072 p.
6. Bolotova Yu.A., Spitzyn V.G., Kermani K.A. Raspoznavaniye simbolov na tsvetnom fone na osnove ierarkhicheskoy vremennoy modeli s predobrabotkoy filtrami Gabora [Character recognition on a colored background based on HTM model pretreatment with Gabor filters]. *Electromagnetic waves and electronic systems*, 2012, vol. 16, no. 1, pp. 14–19.
7. Bundzel M., Hashimoto S. Object identification in dynamic images based on the memory-prediction theory of brain function. *Journal of Intelligent Learning Systems and Applications*, 2010, vol. 2, no. 4, pp. 212–220.
8. Buy Thi Thu Chang, Fan Ngok Hoang, Spitzyn V.G. Algoritmicheskoe i programmnoe obespechenie dlya klassifikatsii tsifrovyykh izobrazheniy s pomoshchyu Veyvlet preobrazovaniya Haara i neyronnykh setey [Algorithms and software for digital images classification with the help the Haar wavelet transform and neural networks]. *Bulletin of Tomsk Polytechnic University*, 2011, vol. 319, no 5, pp. 103–106.
9. Bolotova Yu.A., Spitzyn V.G., Fomin A.E. Primeneniye modeli ierarkhicheskoy vremennoy pamyati dlya raspoznavaniya izobrazheniy [The hierarchical temporal memory model application for image recognition]. *Bulletin of Tomsk Polytechnic University*, 2011, vol. 318, no. 5, pp. 60–63.
10. Kermani K.A., Spitzyn V.G., Hamker F. Nakhozhdeniye parametrov i udalenie postoyannoy sostavlyayushchey filtra Gabora dlya obrabotki izobrazheniy [Finding parameters and removing the constant component of Gabor filter for image processing]. *Bulletin of Tomsk Polytechnic University*, 2011, vol. 318, no. 5, pp. 57–59.

11. Hansen D.W., Hansen J.P., Nielsen M. Eye typing using Markov and active appearance models. *Workshop on applications of computer vision*. Orlando, Florida, USA, IEEE Computer Society, 2002, vol. 12, pp. 132–136.
12. Rowley H.A., Baluja S., Kanade T. Neural network-based face detection. *Pattern anal. mach. intell.*, 2000, vol. 5, pp. 23–38.
13. LeCun 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, pp. 97–104.
14. George D., Hawkins J. A hierarchical bayesian model of invariant pattern recognition in the visual cortex. *Proceedings. 2005 IEEE International Joint Conference on Neural Networks*. Montreal, Canada, IEEE Computer Society, 2005, vol. 3, pp. 1812–1817.
15. Chapelle O., Haffner P., Vapnik V. SVMs for histogram-based image classification. *IEEE Trans. Neural Networks*. Washington, DC, USA, IEEE Computer Society, 1999, vol. 10, no. 5, pp. 1055–1064.
16. 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*, 2002, vol. 3, no. 23, pp. 42–53.
17. Karnowski T., Arel I., Rose D. Deep Spatiotemporal Feature Learning with Application to Image Classification. *The 9th International conference on machine learning and applications (ICMLA'10)*. Washington, DC, USA, IEEE Computer Society, 2010, pp. 883–888.
18. Yu N., Notkin B.S., Sedov V.A. Neuro-iterative algorithm of tomographic reconstruction of the distributed physical fields in the fiberoptic measuring systems. *Computer optics*, 2009, vol. 33, no. 4, pp. 446–455.
19. Lai Z., Hongbin G., Ben N. Visual Hand Pose Estimation Based on Hierarchical Temporal Memory in Virtual Reality Cockpit Simulator. *Information Technology Journal*, 2011, vol. 10, no. 9, pp. 1809–1816.
20. 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.

УДК 004.932

ПРИМЕНЕНИЕ МНОГОШАГОВОЙ СЕГМЕНТАЦИИ ДЛЯ РАСПОЗНАВАНИЯ НЕЧЕТКИХ ДУБЛИКАТОВ ИЗОБРАЖЕНИЙ

Немировский Виктор Борисович,

канд. физ.-мат. наук, доцент каф. информатики и проектирования систем
Института кибернетики Томского политехнического университета,
Россия, 634050, г. Томск, пр. Ленина, д. 30. E-mail: vbn@tpu.ru

Стоянов Александр Кириллович,

канд. техн. наук, доцент каф. информатики и проектирования систем
Института кибернетики Томского политехнического университета,
Россия, 634050, г. Томск, пр. Ленина, д. 30. E-mail: sto@tpu.ru

Актуальность работы обусловлена необходимостью распознавания нечётких дубликатов изображений в системах технического зрения, в работе с большими цифровыми архивами, а также при поиске изображений в сети Интернет.

Цель работы: исследование возможности применения многошаговой сегментации для распознавания нечётких дубликатов изображений.

Методы исследования: в выполненных исследованиях сегментация реализуется за счёт кластеризации яркостей пикселей изображения. Для кластеризации используется рекуррентная нейронная сеть, моделируемая одномерными точечными отображениями. Для оценки близости изображений применено косинусное расстояние между ранговыми распределениями мощностей кластеров яркости.

Результаты: Предложен поисковый образ изображения, основанный на ранговом распределении мощностей кластеров яркостей, выделенных на изображении. Приводятся экспериментальные результаты по распознаванию дубликатов изображений, основанному на применении предложенного образа. Показано, что использование многошаговой сегментации и рангового распределения мощности кластеров яркости позволяет надёжно определять нечёткие дубликаты оригинала изображения с большой степенью искажения на них, вплоть до радиуса гауссовых искажений, равного 8 пикселям. Применение такого подхода позволяет надёжно решать и обратную задачу обнаружения оригинала изображения даже по его пятикратно уменьшенной копии с радиусом гауссовых искажений на ней до 8 пикселей.

Ключевые слова:

Изображение, пиксель, точечные отображения, рекуррентная нейронная сеть, кластеризация, сегментация, распознавание изображений, ранговое распределение.

Введение

Задача распознавания изображений актуальна для систем технического зрения в робототехнике и иных технических системах; для поиска изобра-

жений по образцу в Интернете, цифровых архивах и библиотеках; при анализе содержания космических и аэрофотоснимков и т. д. [1, 2]. В основном это связано с отсутствием универсальных методов,