

Метод блочного оптического распознавания инвентарных номеров железнодорожных подвижных единиц на основе комитетной нейроиммунной модели классификации

И.С. Артемьев, А.И. Лебедев, А.И. Долгий, А.Е. Хатламаджиян, В.Д. Меерович

Введение

Опыт внедрения системы автоматического распознавания номеров вагонов (АРНВ) [1] показывает необходимость максимального использования всей графической информации, нанесенной на железнодорожные подвижные единицы, в целях повышения качественных показателей системы.

Согласно утвержденному альбому «Знаки и надписи на вагонах грузового парка колеи 1520мм» № 632-2006 ПКБ ЦВ ОАО «РЖД» основные инвентарные номера, наносимые на борт вагона, дублируют в нижней части подвижной единицы: на кузове, раме, борте или нижней обвязке кузова, в зависимости от типа подвижной единицы. В связи с этим очевидна целесообразность применения в системе АРНВ нижних камер для распознавания номеров платформ и повышения устойчивости идентификации остальных подвижных единиц по дублирующим номерам.

Дублирующие номера вагонов имеют вдвое меньший размер по высоте и более широкий диапазон изменения масштаба, соответственно сильнее подвержены шуму, размытию, оптическим искажениям, что предъявляет повышенные требования к основным сложноорганизованным частям алгоритмов, таким как сегментация и распознавание, в существенной степени влияющим на итоговый результат [1].

Для решения задачи распознавания дублирующих номеров ж.-д. подвижных единиц к используемым алгоритмам предъявляются следующие требования:

- устойчивость к шуму;
- инвариантность к существующим изменениям масштаба номера;
- возможность сегментации размытых и слипшихся цифр номера, имеющих разные шрифты и начертания;
- минимальное время выполнения обработки номера;
- точность и производительность.

Для решения поставленной задачи необходимо разработать подход, объединяющий этапы сегментации и классификации цифр дублирующего номера вагона с учетом предъявленных требований. Одним из путей решения этой задачи являются подходы, основанные на блочном распознавании текста.

Блочное распознавание. На текущий момент, разработанные подходы блочного распознавания символов можно разделить по принципу формирования визуального стимула рецептивного поля искусственных нейронных сетей (ИНС)[2, 3]:

- 1) эвристические подходы [4];
- 2) подходы, основанные на использовании «скользящего окна» [5];
- 3) специализированные архитектуры ИНС [6].

К эвристическим, как наиболее творческим приемам в поиске решения задачи, относятся самые разнообразные варианты решения проблем традиционных методов сегментации.

Одни из них дополнительно ставят перед ИНС, классифицирующей символы, задачу определения класса «не символ», другие комбинируют результаты разных ИНС, распределяя задачу определения класса «символ – не символ» и определения класса символа.

Также стоит отметить эвристические подходы (рис.1) к решению распространенной проблемы сегментации традиционным методом – разделению слипшихся цифр, которые используют специализированные ИНС, обученные для классификации сдвоенных или строенных цифр, количество

которых на предоставляемом образе определяется эмпирически либо классифицируется отдельной ИНС.

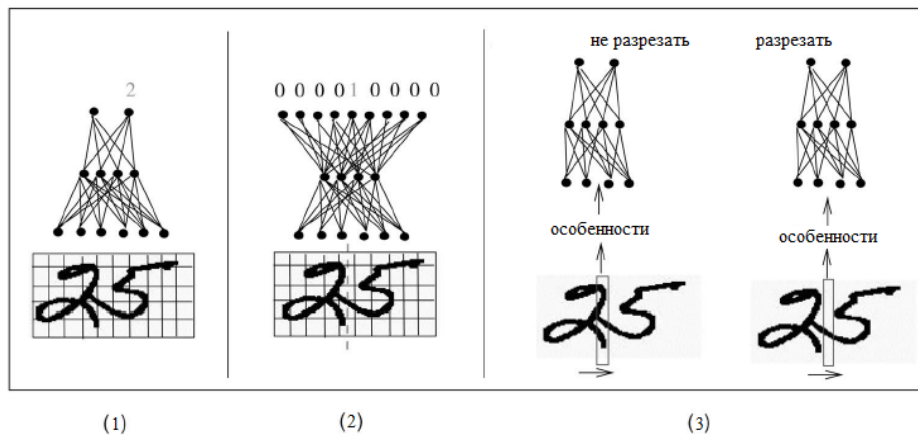


Рис. 1. – Эвристические методы блочного распознавания:

- 1) определение количества символов в блоке;
- 2) определение позиции разделения символов;
- 3) пошаговая классификация признаков связности символов

Что касается подходов, основанных на формировании входного изображения ИНС с помощью скользящего окна, их можно отнести к наиболее популярным и часто встречающимся методам сегментации. Задачей «сегментатора» является определение позиций окна, в которых символы попадают полностью. Такие образы называются «истинными объектами», а образы, полученные на пространствах между соседними символами, именуется «анти-объектами» (рис.2).



Рис. 2. – Блочное распознавание с помощью «скользящего окна»

Специализированные архитектуры ИНС, такие как SDNN - SpaceDisplacementNeuralNetwork (рис. 3), TDNN - TimeDelayNeuralNetwork, представляют собой сетевую архитектуру прямого распространения для блочного распознавания многосимвольных строк, основанную на репликации входного поля с помощью сверточных ИНС. Прежде всего, благодаря тому, что сверточные ИНС являются очень устойчивыми к сдвигам и масштабным вариациям входного изображения, а также шуму и посторонним меткам на входе. Входной слой сети соединен с модулем интерпретации результатов, работающим по алгоритму Витерби, который выбирает лучшую последовательность сегментированных и распознанных символов входной строки.

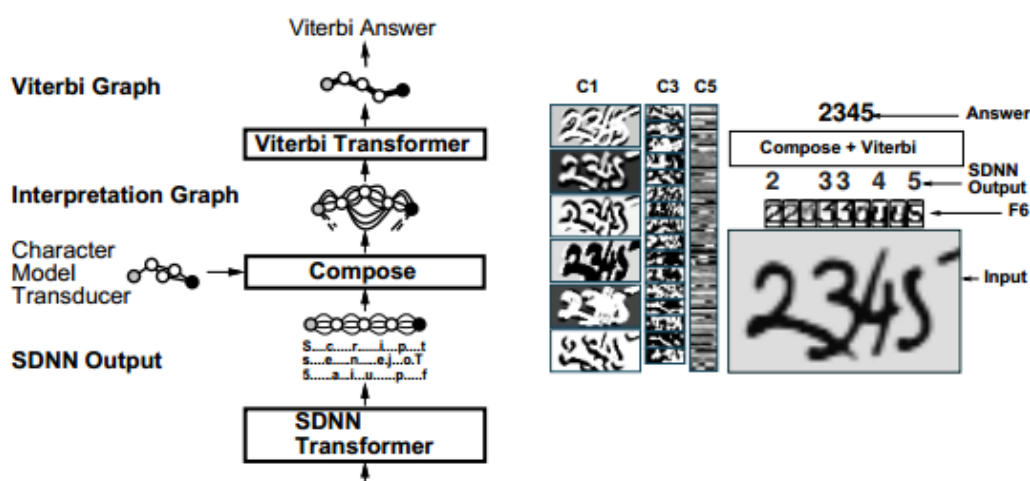


Рис. 3. – Блочное распознавание с помощью SDNN

Перечисленные выше подходы объединяет использование классификатора для разделения подаваемых образов на два класса «символ» и «не символ». Однако, формирование репрезентативной обучающей выборки для класса «не символ» зачастую является крайне нетривиальной задачей, т.к. элементами этого класса являются все оставшиеся мыслимые объекты, лежащие вне класса «символ».

В качестве решения этой задачи в настоящей статье предлагается метод на основе комитетной модели классификации.

Метод блочного распознавания инвентарных номеров железнодорожных подвижных единиц. В основе предлагаемого метода лежит предположение о том, что члены комитета неустойчивых классификаторов, обученные только на положительных примерах, будут чаще противоречить друг другу, чем когда им будет предъявлен образ из известного им класса. Однако при этом необходимо, чтобы классификаторы, участвующие в комитете имели достаточно высокое разнообразие. Для этого их необходимо обучать на различных выборках данных. Так как в нашем распоряжении имеется лишь одно множество примеров, то различающиеся подмножества с близким статистическим распределением можно получить путем применения бутстрепа [7] — случайной выборки с возвратом.

Подход, основанный на независимом обучении отдельных моделей на бутстреп-выборках из обучающего набора данных, был предложен в [8] и получил название «бэггинг».

Графическая интерпретация комитета неустойчивых классификаторов (нейронных сетей), разделивших при обучении на бутстреп-выборках множества классов гиперплоскостями, спроецированная на двумерную плоскость, демонстрирует суждение о результативности комбинирования неустойчивых моделей (рис.4).

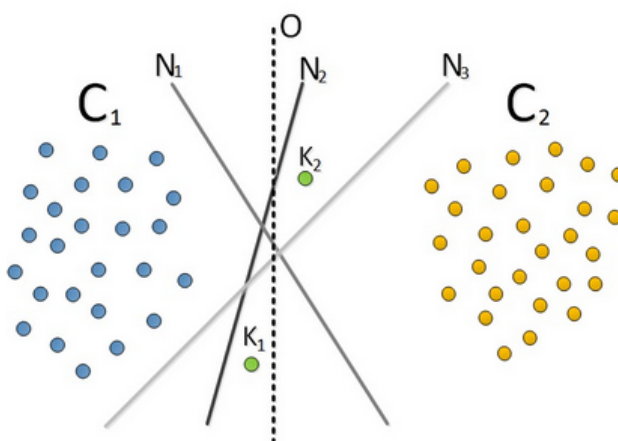


Рис. 4. – Графическая интерпретация комитета линейных пороговых классификаторов. C_1 , C_2 – множество объектов класса «1» и «2» соответственно, O – граница разделения классов, N_n – граница разделения классов по «мнению» n -й ИНС. K_1 , K_2 – новые объекты для классификации

Очевидно, что точность классификации комитетом повышается с увеличением количества нейросетей. Однако, при наращивании численности членов комитета, пропорционально увеличивается и время классификации, что не соответствует требованиям производительности алгоритмов систем, работающих в реальном времени. Соответственно, при формировании комитета, количество входящих в него классификаторов определяется некоторым оптимальным соотношением точности и времени классификации.

В качестве комитета неустойчивых классификаторов предлагается применение разработанной ранее нейроиммунной модели классификации (НИМК) [9] для блочного распознавания инвентарных номеров вагонов. НИМК обладает свойством редукции данных за счет применения механизма иммунной кластеризации и, в отличие от алгоритма-прототипа AIRS2 (Artificial Immune Recognition System v2) [10], имеет быстрый механизм классификации новых примеров без потери точности, основанный на применении комитета неустойчивых классификаторов – ИНС прямого распространения, обучаемых на бутстреп выборках.

Алгоритм НИМК состоит из следующих основных этапов:

1. Применяется метод главных компонент для сокращения размерности данных.
2. Формируется множество клеток памяти по методу, применяемому в алгоритме AIRS2.
3. Формирование обучающих множеств для ансамбля ИНС методом бэггинга на основе множества клеток памяти, полученных на предыдущем этапе.
4. Эвристическое формирование и обучение ансамбля ИНС.

5. Классификация: классифицируемые примеры подаются на вход ансамбля ИНС, полученного на предыдущем шаге.

При формировании визуального стимула рецептивного поля НИМК методом «скользящего окна», наблюдается подтверждение особенности комитета, предположенной нами выше – о противоречии большинства ИНС друг другу, при предъявлении образов неизвестного им класса. Графическая интерпретация выявленной особенности представлена на рис. 5.

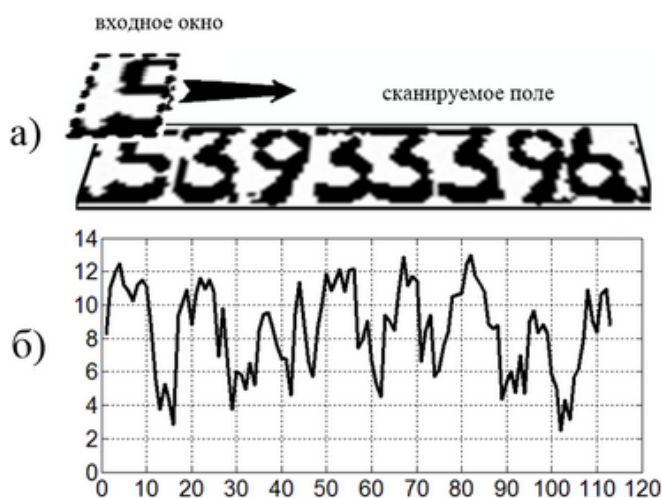


Рис. 5. – Графическое представление результатов классификации каждого шага скользящего окна с помощью НИМК

Изображение на рис. 5а) соответствует бинарному изображению области с инвентарным номером вагона, на рис.5б) (выходной сигнал) – результатам классификации комитетом каждого шага окна, где ось ординат – наибольшая величина суммы значений выходов комитета, ось абсцисс – координаты позиции окна.

Как видно по рис. 5б), существует необходимость подавления высокочастотной составляющей выходного сигнала для возможности корректной интерпретации полученных результатов. Для этого в разработанном методе используется КИХ-фильтр – линейный цифровой фильтр, характерной особенностью которого является ограниченность по времени его импульсной характеристики [11].

В задачах цифровой обработки сигналов и изображений КИХ-фильтр является эффективной унифицированной процедурой, осуществляющей базовую операцию – вычисление свертки входного изображения $x(n_1, n_2)$ и линейного фильтра $h(m_1, m_2)$, которая может быть представлена в виде (1).

$$y(n_1, n_2) = \sum_{(m_1, m_2) \in D} h(m_1, m_2)x(n_1 - m_1, n_2 - m_2), \quad (1)$$

где $y(n_1, n_2)$ – выходной ряд;

D - конечная область ненулевых отсчетов импульсной характеристики;

n_1, n_2 - отсчеты входного изображения;

m_1, m_2 - отсчеты коэффициентов фильтра.

Результат применения свертки выходного сигнала КИХ-фильтром изображен на рис. 6.

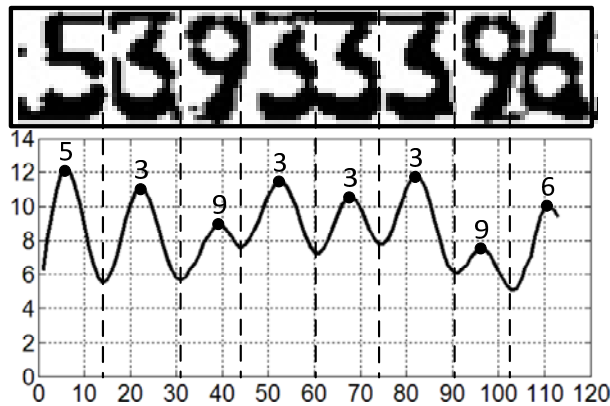


Рис. 6. – Результат блочного распознавания инвентарного номера

На полученном выходном ряде $y(n_1, n_2)$ локальные минимумы x_i^{\min} соответствуют вертикальным уровням разделения цифр, локальные максимумы x_i^{\max} – центральным позициям цифр распознаваемого номера, а также номеру класса, к которому они относятся. Обнаружение точек локальных экстремумов производится на этапе итеративной свертки КИХ-

фильтром с помощью вычисления определителя матрицы Гессе для точки x_i в апертуре $1 \times w$:

$$x_i \rightarrow x_i^{\max}, M_k > 0, k = 1, \dots, w,$$

$$x_i \rightarrow x_i^{\min}, (-1)^k M_k > 0, k = 1, \dots, w,$$

где M_k – диагональные миноры симметрической матрицы Гессе.

Таким образом, сформируем основные этапы предлагаемого нами метода блочного распознавания инвентарных номеров железнодорожных подвижных единиц, основанного на комитетной нейроиммунной модели классификации:

1. Формирование и обучение НИМК.
2. Классификация каждого шага входного изображения методом скользящего окна.
3. Сглаживание полученного ряда результатов классификации КИХ-фильтром.
4. Обнаружение локальных экстремумов за счет вычисления оператора матрицы Гессе для каждой точки ряда, полученного на предыдущем этапе.
5. Формирование результата распознавания по полученным на предыдущем этапе локальным максимумам.

Вычислительные эксперименты. Для проведения вычислительных экспериментов нами был реализован предлагаемый метод блочного распознавания инвентарных номеров железнодорожных подвижных единиц.

Обучение и тестирование НИМК производилось на множестве бинарных изображений цифр дублирующих номеров ж.-д. вагонов (рис.7), полученных на одном из объектов внедрения АРНВ.



Рис. 7. – Примеры бинарных изображений цифр дублирующих номеров ж.-д. вагонов

Обучающее множество содержало 10 тыс. уникальных примеров, а тестовое 1 тыс. размерностью 16 на 27 пикселей (432 признака).

Выборка для тестирования предлагаемого метода состояла из 1 тыс. примеров бинарных изображений дублирующих номеров размерностью 100 на 27 пикселей (рис. 8).

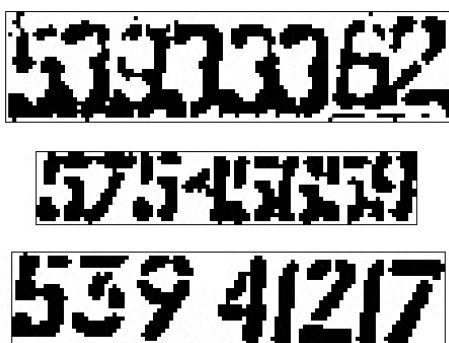


Рис. 8. – Примеры бинарных изображений дублирующих номеров

Эксперименты выполнялись на специализированном компьютере промышленного исполнения с установленным процессором IntelCore i7 частотой 3200 МГц.

В результате применения метода главных компонент размерность признакового пространства была снижена до 70.

Архитектура используемых в экспериментах ИНС – 70-35-20-10 нейронов.

Обучение всех ИНС производилось метом Левенберга-Марквардта.

Оптимальное количество ИНС в комитете подобрано эмпирически.

Количество правильно сегментированных и распознанных номеров составило 96,8% всей тестовой выборки. Среднее затраченное время на распознавание одного номера – 57 мс.

Заключение. Предложенный в настоящей статье метод блочного распознавания инвентарных номеров железнодорожных подвижных единиц основан на комитетной нейроиммунной модели классификации. Преимуществом использования комитета неустойчивых классификаторов является отсутствие необходимости формирования выборки отрицательных примеров.

Метод блочного распознавания инвентарных номеров подвижных единиц с использованием скользящего окна, объединяющий в себе этапы сегментации и классификации, позволяет достичь повышенной устойчивости к шуму, возможности сегментации размытых и слипшихся цифр номера, имеющих разные шрифты и начертания, а также инвариантности к существующим изменениям масштаба номера.

Кроме того, разработанный метод обладает важным свойством, унаследованными от модели НИМК – редукцией данных достигаемой за счет применения механизма иммунной кластеризации. Это свойство позволяет постоянно пополнять обучающую выборку комитета классификаторов новыми статистическими данными для последующего повышения точности классификации.

В настоящее время предложенный метод блочного распознавания инвентарных номеров железнодорожных подвижных единиц реализован в программном обеспечении системы АРНВ, находящейся в эксплуатации на сети дорог ОАО «РЖД».

Работа выполнена при финансовой поддержке РФФИ, проект № 13-07-00226 А, проект № 13-07-13109 офи_м_РЖД.

Литература:

1. Артемьев И.С., Лебедев А.И. Прогностическая модель сегментации трафаретных цифр в задаче оптической идентификации инвентарных номеров железнодорожных подвижных единиц // Труды

международной научно-практической конференции «Транспорт-2013». – Ростов н/Д., 2013. С. 7 – 9.

2. Романов Д.Е. Нейронные сети обратного распространения ошибки[Электронный ресурс] // «Инженерный вестник Дона», 2009, №3. – Режим доступа: <http://www.ivdon.ru/magazine/archive/n3y2009/143> (доступ свободный) – Загл. с экрана. – Яз.рус.

3. Лиля В.Б. Алгоритм и программная реализация адаптивного метода обучения искусственных нейронных сетей[Электронный ресурс] // «Инженерный вестник Дона», 2012, №1. – Режим доступа: <http://www.ivdon.ru/magazine/archive/n1y2012/626> (доступ свободный) – Загл. с экрана. – Яз.рус.

4. Marinai, S. Artificial neural networks for document analysis and recognition / S. Marinai, M. Gori, G. Soda // Pattern analysis and machine intelligence, pages 23-35, 2005.

5. Lee, SW. Integrated segmentation and recognition of handwritten numerals with cascade neural network / SW Lee, SY Kim // Systems, Man, and Cybernetics, Part C: Applications and Reviews, pages 285-290, 1999.

6. Y. LeCun, P. Haffner, L. Bottou and Y. Bengio. Object Recognition with Gradient-Based Learning. In Forsyth, D. (Eds), Feature Grouping, Springer, 1999– 28 p.

7. Hastie T., Tibshirani R., Friedman J., The elements of statistical learning: Data mining, inference, and prediction. Springer, 2001– 533 p.

8. Breiman, L. Bagging predictors. Univ. California Technical Report No. 421, September 1994 – 20 p.

9. Артемьев И.С., Долгий А.И., Суханов А.В., Хатламаджиян А.Е. Нейроиммунная модель классификации в задачах идентификации на транспорте // Сборник научных трудов VII-й Международной конференции «Интегрированные модели и мягкие вычисления в искусственном интеллекте». – М.: Физматлит, 2013. – С. 980-987.

10. Watkins, A. Exploiting Immunological Metaphors in the Development of Serial, Parallel, and Distributed Learning Algorithms. PhD Thesis, Mississippi State University, 2005–314 p.

11. Ifeachor, E.C. Digital Signal Processing: A Practical Approach / E.C. Ifeachor, B.W. Jervis // Prentice-Hall, Inc., Cliffs, 2002 – 933 p.