

УДК 004.021  
DOI: 10.15827/0236-235X.133.151-156

Дата подачи статьи: 30.07.20  
2021. Т. 34. № 1. С. 151–156

## **Адаптация модели нейронной сети LSTM для решения комплексной задачи распознавания образов**

В.С. Тормозов<sup>1</sup>, старший преподаватель, 007465@pnu.edu.ru

<sup>1</sup> Тихоокеанский государственный университет, г. Хабаровск, 680035, Россия

В статье рассматривается адаптация модели искусственных нейронных сетей прямого распространения с блоками долгой краткосрочной памяти (LSTM) для комплексной задачи распознавания образов. Для искусственных нейронных сетей контекст может быть извлечен из вектора поступающих на вход сигналов и из значений весовых коэффициентов обученной сети. Однако при учете контекста значительного объема растут количество нейронных связей и сложность процедур обучения и работы сети. Контекст может временно храниться в специальном буфере памяти, откуда в дальнейшем извлекаться и использоваться в качестве сигнала при работе искусственных нейронных сетей (LSTM). Преимущество сетей такого типа в том, что они используют блоки памяти, ассоциированные с каждым нейроном скрытого слоя, что позволяет сохранять данные, относящиеся к контексту, при формировании шаблонов распознавания.

Предлагается метод линейного переключения блоков LSTM в зависимости от значения проходящего сигнала. Проведен вычислительный эксперимент, направленный на исследование эффективности предложенного метода и ранее разработанной нейронной сети прямого распространения аналогичной структуры. Выполнено машинное обучение для каждого вида искусственных нейронных сетей на одной и той же последовательности обучающих примеров. Проведено сравнение результатов экспериментов для искусственной нейронной сети прямого распространения, рекуррентной нейронной сети аналогичной архитектуры с таким же количеством нейронов на каждом слое и сетью нейромодулирующего взаимодействия с одной задержкой по обратной связи. Критерием оптимизации в данном случае является погрешность работы нейронной сети на обучающей выборке, состоящей из примеров, не представленных в тестирующей выборке. Эффективность решения классификационной задачи оценивается по двум критериям: погрешность обучения на обучающей выборке и погрешность тестирования на тестирующей выборке.

**Ключевые слова:** блок долгой краткосрочной памяти, распознавание образов, машинное обучение, искусственный интеллект, искусственная нейронная сеть.

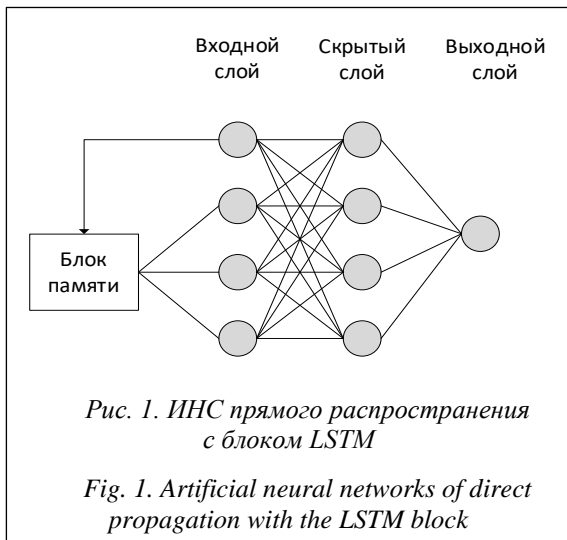
При разработке автоматизированных интеллектуальных систем классификаций в условиях неопределенности немаловажную роль играет смысловой контекст решаемой задачи. К такому контексту относятся и информация о прошлых состояниях системы. Сегодня высокую популярность для задач классификации входных образов приобретают нейросетевые математические модели. Для *искусственных нейронных сетей* (ИНС) контекст может быть извлечен из вектора поступающих на вход сигналов и из значений весовых коэффициентов обученной сети [1]. Немаловажным преимуществом нейросетевого подхода является возможность его параллельной работы на современных микропроцессорах и GPU. Однако при учете контекста значительного объема растут количество нейронных связей и сложность процедур обучения и работы сети. Контекст может временно храниться в специальном буфере памяти, откуда в дальнейшем извлекаться и использоваться в качестве сигнала при работе

ИНС. Такой вид памяти носит название *долгой краткосрочной памяти* (ДКП) (Long Short-Term Memory, LSTM), в ней сохраняются предыдущие значения входного вектора.

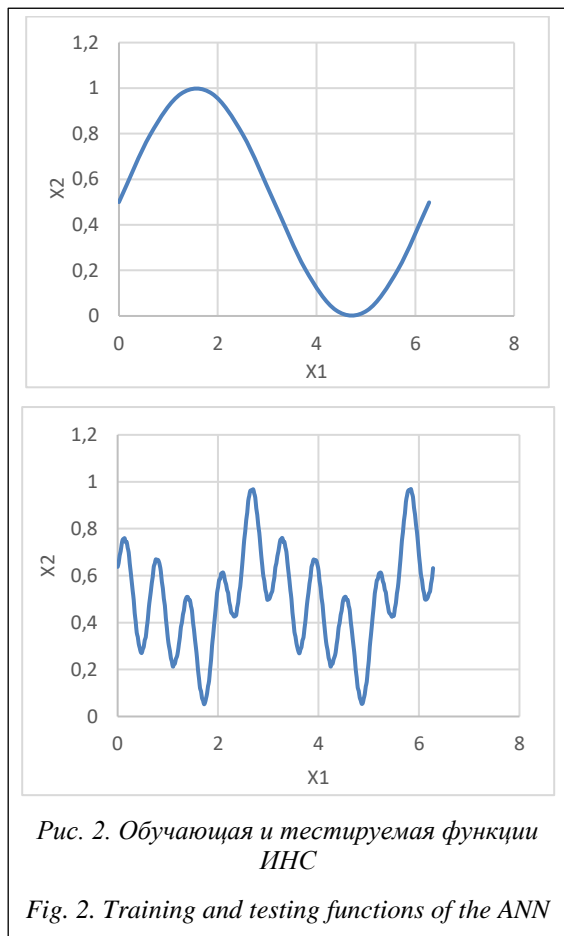
### **Постановка комплексной задачи распознавания**

Информационно-технические методы решения прикладных задач находят все более широкую область применения [2]. Для предварительной обработки и фильтрации зашумленных аналоговых сигналов можно использовать ИНС, представленную на рисунке 1 [3, 4]. Нейроны первого слоя принимают сигналы от блока LSTM и обрабатывают их так же, как и нейроны скрытого слоя. В качестве нелинейной дифференцируемой функции активизации было выбрано нормальное распределение для нейронов всей сети.

Такая ИНС может быть применена для предсказания будущей последовательности



значений синусоидального сигнала с зашумлением. В блоке памяти сохраняются значения десяти предыдущих входных сигналов. Учебная выборка состоит из значений одной синусоиды  $\frac{1}{2} + \frac{1}{2}\sin(x_1)$  в диапазоне  $[0, 2\pi]$ . Тестовая выборка состоит из значений суммы трех синусоид переменных частот в том же диапазоне (рис. 2). Погрешность обучения достигла



8,5 % после  $7 \cdot 10^5$  итераций обучения, в то время как погрешность тестирования составила 16,7 %.

Распознавание образов – область науки, посвященная разработке принципов, построению программных и (или) аппаратных систем, предназначением которых является определение принадлежности рассматриваемого объекта к какому-либо из заранее заданных классов. Объектом могут быть любой сигнал, реально существующий предмет или абстракция, явление и другое.

Объект определяется конечной последовательностью признаков или свойств. Признаки – это данные, полученные в результате измерения объекта для того, чтобы использовать их в процессе классификации. Множество признаков одного объекта образует так называемый вектор признаков, который может быть отмечен как точка в гиперпространстве признаков. Каждому образу сопоставлен только один вектор признаков, как и вектору признаков только один образ. Отнесение объекта к одному из классов называется классификацией. Классификация выполняется с применением эталонных объектов, класс которых уже известен.

Образ – описание объекта, используемое для решения задачи распознавания образов, отнесения объекта к одному из классов. Для решения задачи распознавания образов как задачи дискриминантного анализа необходимо, чтобы все возможные образы объектов разбивались на конечное множество классов. Также должно быть задано множество эталонных объектов. Каждый эталонный объект принадлежит одному из классов. Существует и объект, принадлежность которого какому-либо классу неизвестна. Задача распознавания образов заключается в том, чтобы установить, к какому классу принадлежит этот объект, либо разработать алгоритм, выполняющий это для любого рассматриваемого объекта. Чем чаще система верно определяет принадлежность объекта классу, тем точнее система распознавания образов.

Качество работы системы распознавания может быть оценено с точки зрения точности назначения класса предъявляемым объектам. Различают распознавание с обучением и без обучения. Если распознавание выполняется с применением имеющегося множества эталонных объектов, то задачу распознавания называют классификацией с обучением.

Нейросетевой подход с сохранением предыдущих значений в блоках LSTM – очень эффек-

тивный инструмент для задач как распознавания, так и прогнозирования временных рядов. Тем не менее, у такого подхода имеются ограничения по функциональности, связанные с предопределенностью блока LSTM. Развитием нейросетевой архитектуры с LSTM является ИНС, сохраняющая промежуточные значения проходящих сигналов во внутренних слоях. Такое усовершенствование позволяет сети производить самонастройку с помощью блоков памяти, осуществляющих чтение и запись значений сигнала в долговременную память.

### Обзор существующих исследований

Метод распознавания образов, базирующийся на обучении модели ИНС с ДКП непосредственно из последовательностей элементов-представителей высокого уровня с сохранением порядка следования этих представительства, подробно рассмотрен в работе [5]. Эти элементы представляют собой выборку из всего множества обрабатываемых элементов. Таким образом, было предложено объединение архитектур ИНС и блоков ДКП для моделирования предложений естественного языка. Чтобы воспользоваться преимуществами как одной, так и другой архитектуры, была спроектирована простая сквозная унифицированная архитектура, соединяющая выход однослойной СНС с ДКП. Были проведены два эксперимента – для классификации предложений и классификации типа вопросов. Получившаяся гибридная система для задачи классификации предложений показала 94,6 % точности на тестовых данных.

Подход, основанный на модели нейронной машины Тьюринга (NTM, НТМ), был реализован в работе [6]. НТМ объединяет возможности ИНС по нечеткому сопоставлению с образцом и возможности программируемых алгоритмических компьютерных систем. НТМ имеет контроллер нейронной сети, связанный с внешними ресурсами памяти, с которыми он взаимодействует посредством механизмов внимания. Взаимодействия с памятью различаются между собой, что позволяет оптимизировать их с помощью метода градиентного спуска. НТМ сетевой контроллер с ДКП позволяет создавать многие алгоритмы, такие как копирование, сортировка и ассоциативное обращение только путем обучения на примерах. В архитектуру НТМ встроены структуры памяти, которым доступны команды чтения, записи и удаления данных. Сеть была обучена последовательно-

стью из двадцати 8-битных чисел, а тестирование осуществлялось последовательностью из 120 тестов с очень хорошими результатами. При таком подходе увеличение памяти может быть достигнуто не только корректировкой значений весовых коэффициентов, но и сохранением обучаемой информации в структурах памяти.

### Методы решения комплексной задачи

Архитектура исследуемой нейронной сети включает в себя ИНС прямого распространения с блоком LSTM, связанным с каждым нейроном в каждом из скрытых слоев. Входной и выходной слои составлены только из стандартных нейронов, в то время как скрытые слои из нейронов с подключенными LSTM. Выход каждого нейрона формируется сигмоидальной функцией комбинации его входных сигналов. Выход нейрона работает как линейный переключатель активации подключенного к нему блока LSTM. Нейронная сеть такой архитектуры в процессе обучения приобретает способность извлекать записанный в блоки памяти контекст по мере необходимости [7].

ИНС с блоками LSTM имеет входной и выходной слои, состоящие из стандартных нейронов и скрытых слоев, которые, в свою очередь, состоят из стандартных нейронов с подключенными к ним блоками LSTM. Нейроны скрытых слоев имеют сигмоидальную функцию активации. Каждый нейрон связан с блоком памяти, и каждый блок памяти состоит из множества буферов памяти. Выходы нейронов используются в качестве линейных переключателей, чтобы при необходимости активизировать определенный буфер памяти.

Процедуры обучения и тестирования сети состоят из трех базовых этапов: прямое распространение сигналов по сети через каждый слой, обучение сети с помощью алгоритма обратного распространения ошибки, сохранение значений сигнала памяти в каждом скрытом слое как ортогональный набор в многомерном пространстве.

Вектор состояния вывода для стандартных нейронов каждого слоя увеличивается с каждым проходом сигнала через скрытые слои, содержащие LSTM [8]. Выходной сигнал каждого блока памяти – это содержание того буфера памяти, к которому отнесен выходной сигнал подключенного к нему нейрона. Выполняется линейное переключение блока памяти в зависимости от того, к какому отрезку допу-

стимых значений сигнала относится проходящий через нейрон сигнал. При таком подходе нейрон выступает в роли линейного переключателя, включающего конкретный буфер памяти в зависимости от значения выхода данного нейрона. Нейроны скрытого слоя получают входной сигнал от предыдущего слоя, в то время как блоки памяти воспринимают входной сигнал только от нейрона, к которому они подключены [9]. Обучение с использованием алгоритма обратного распространения ошибки для ИНС с блоками памяти аналогично обучению ИНС прямого распространения с использованием этого алгоритма [10].

Каждый нейрон скрытого слоя содержит блок LSTM. Таким образом, в сети сохраняются воспоминания о предыдущих прохождениях сигнала. Каждое уникальное воспоминание в нейроне ортогонально всем другим воспоминаниям, представленным в этом блоке. Последовательность воспоминаний представлена в виде вектора в многомерном пространстве и называется составной памятью. Направление вектора воспоминаний указывает на то, какие воспоминания присутствуют, а значения величины вектора при проекции на оси каждого воспоминания – на порядок, в котором они произошли [11]. Вновь занесенные в блок памяти значения задаются ортогональным отображением выходного сигнала подключенного и соседнего нейронов. Ближайшие соседние нейроны могут быть сформулированы в пространственном смысле – как нейроны того же скрытого слоя с близкими значениями индексов. Ближайшие соседние нейроны несут информацию о состоянии предыдущего слоя.

### Вычислительный эксперимент

Набор данных состоит из выборки восьми случайно сгенерированных многомерных нормальных распределений. Целью работы сети является классификация, определяющая, из какого распределения была взята заданная точка. Помимо этого, были протестированы *ИНС прямого распространения* (ИНС ПР) и *рекуррент-*

*ная нейронная сеть* (РНС) аналогичной архитектуры (с таким же количеством нейронов на каждом слое), *сеть нейромодулирующего взаимодействия* (СНВ) [12] с одной задержкой по обратной связи. Минимальная классификационная погрешность обучения исследуемой РНС составила 1,9 %. Результаты проведенного исследования показаны на рисунке 3.

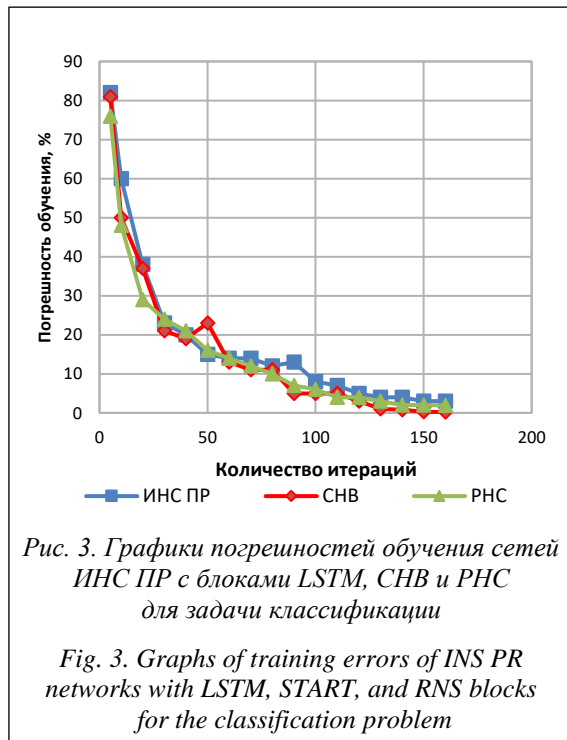


Рис. 3. Графики погрешностей обучения сетей ИНС ПР с блоками LSTM, СНВ и РНС для задачи классификации

Fig. 3. Graphs of training errors of INS PR networks with LSTM, START, and RNS blocks for the classification problem

### Заключение

Предложенный метод линейных переключателей блока LSTM на скрытых слоях ИНС целесообразно применять, когда результат текущей обработки коррелирует с прошедшими значениями сигналов сети. Такая ИНС показывает хорошие результаты для задач классификации и распознавания образов в сравнении с РНС. Преимущество сетей данного типа в том, что они используют блоки памяти, ассоциированные с каждым нейроном скрытого слоя, что позволяет сохранять контекст при формировании шаблонов распознавания.

Работа выполнена при поддержке Минобрнауки РФ, дополнительное соглашение № 075-02-2020-1529/1 от 21.04.2020.

### Литература

1. Hopfield J.J. Neural networks and physical systems with emergent collective computational abilities. PNAS, 1982, vol. 79, no. 8, pp. 2554–2558. DOI: 10.1073/pnas.79.8.2554.

2. Золкин А.Л. Разработка информационно-управляющей системы для контроля износа коллекторов тяговых электродвигателей // Вестн. ДААТ. 2019. № 2. С. 65–74.
3. Тормозов В.С. Метод детектирования и классификации транспортных средств на спутниковых снимках сверхвысокого разрешения // Промышленные АСУ и контроллеры. 2019. № 6. С. 18–24.
4. Тормозов В.С. Автоматическое детектирование дорожного покрытия на космических снимках сверхвысокого разрешения // Молодые ученые – Хабаровскому краю: матер. XIX краевого конкурса молодых ученых и аспирантов. 2017. С. 131–135.
5. Zhou C., Sun C., Liu Z., Lau F.C.M. C-LSTM neural network for text classification. Arxiv, 2015, art. 08630. URL: <https://arxiv.org/abs/1511.08630> (дата обращения: 21.10.2020).
6. Graves A., Wayne G., Danihelka I. Neural turing machines. Arxiv, 2014, art. 5401. URL: <https://arxiv.org/abs/1410.5401> (дата обращения: 21.10.2020).
7. Sainath T.N., Vinyals O., Senior A., Sak H. Convolutional, long short-term memory, fully connected deep neural networks. Proc. ICASSP, IEEE, 2015, pp. 4580–4584. DOI: 10.1109/ICASSP.2015.7178838.
8. Weston J., Chopra S., Bordes A. Memory networks. Arxiv, 2014, art. 3916. URL: <https://arxiv.org/abs/1410.3916> (дата обращения: 21.10.2020).
9. Hochreiter S., Bengio Y., Frasconi P., Schmidhuber J. Gradient flow in recurrent nets: the difficulty of learning long-term dependencies. In: A Field Guide to Dynamical Recurrent Neural Networks, 2001, pp. 114–132.
10. Estebon M.D. Perceptrons: An Associative Learning Network. 1997. URL: <https://ei.cs.vt.edu/~history/Perceptrons.Estebon.html> (дата обращения: 21.10.2020).
11. Schmidhuber J., Wierstra D., Gagliolo M., Gomez F. Training recurrent networks by Evolino. Neural Computation, 2007, vol. 19, pp. 757–779. DOI: 10.1162/neco.2007.19.3.757.
12. Vecoven N., Ernst D., Wehenkel A., Drion G. Introducing neuromodulation in deep neural networks to learn adaptive behaviours. PloS ONE, 2020, vol. 15, no. 1, art. e0227922. DOI: 10.1371/journal.pone.0227922.

Software & Systems  
DOI: 10.15827/0236-235X.133.151-156

Received 30.07.20  
2021, vol. 34, no. 1, pp. 151–156

### The adaptation of the LSTM neural network model to solve the pattern recognition complex problem

V.S. Tormozov<sup>1</sup>, Senior Lecturer, 007465@pnu.edu.ru

<sup>1</sup> Pacific National University, Khabarovsk, 680035, Russian Federation

**Abstract.** The paper examines the adaptation of the model of artificial neural networks of direct distribution with blocks of long short-term memory (LSTM) for the complex problem of pattern recognition. For artificial neural networks (ANN), the context can be extracted from the input signal vector and from the weight values of the trained network. However, considering the context of a significant volume, the number of neural connections and the complexity of training procedures and network operation increase. Instead of receiving context from input values, the context can also be temporarily stored in a special memory buffer, from where it can later be extracted and used as a signal in the ANN's operation. This type of memory is called LSTM. The advantage of networks of this type is that they use memory blocks associated with each neuron of the latent layer, which allows context-related data to be stored when forming recognition patterns.

There is the method of linear switching of LSTM units depending on the value of the transmitted signal in the paper. A computational experiment was conducted aimed at investigating the effectiveness of the proposed method and the previously developed neural network of direct distribution of a similar structure. Machine learning was performed for each type of ANN on the same sequence of training examples. The test results were compared for: an ANN of direct propagation, a recurring neural network (RNS) of a similar architecture: with the same number of neurons on each layer, and a network of neuromodulating interaction with one feedback delay. The optimization criterion, in this case, is the error of the neural network on the training sample, consisting of examples not presented in the test. The efficiency of solving the classification problem is evaluated according to two criteria: learning error on the training sample and testing error on the testing sample.

**Keywords:** long short-term memory unit, image recognition, machine learning, artificial intelligence, artificial neural network.

**Acknowledgements.** This paper was financially supported by Ministry of Science and Higher Education of the Russian Federation, supplementary agreement no. 075-02-2020-1529/1 of 21.04.2020.

### References

1. Hopfield J.J. Neural networks and physical systems with emergent collective computational abilities. *PNAS*, 1982, vol. 79, no. 8, pp. 2554–2558. DOI: 10.1073/pnas.79.8.2554.
2. Zolkin A.L. Development of the management information system for control of wear of collectors of traction electric motors. *Vestn. DAAT*, 2019, no. 2, pp. 65–74 (in Russ.).
3. Tormozov V.S. Method of detection and classification of vehicles on ultra-high resolution satellite images. *Industrial Automatic Control Systems and Controllers*, 2019, no. 6, pp. 18–24 (in Russ.).
4. Tormozov V.S. Automatic detection of road surface on space images of ultra-high resolution. *Young Scientists to the Khabarovsk Territory: Proc. 19th Regional Competition of Young Scientists and Graduates*, 2017, pp. 131–135 (in Russ.).
5. Zhou C., Sun C., Liu Z., Lau F.C.M. C-LSTM neural network for text classification. *Arxiv*, 2015, art. 08630. Available at: <https://arxiv.org/abs/1511.08630> (accessed October 21, 2020).
6. Graves A., Wayne G., Danihelka I. Neural Turing machines. *Arxiv*, 2014, art. 5401. Available at: <https://arxiv.org/abs/1410.5401> (accessed October 21, 2020).
7. Sainath T.N., Vinyals O., Senior A., Sak H. Convolutional, long short-term memory, fully connected deep neural networks. *Proc. ICASSP, IEEE*, 2015, pp. 4580–4584. DOI: 10.1109/ICASSP.2015.7178838.
8. Weston J., Chopra S., Bordes A. Memory networks. *Arxiv*, 2014, art. 3916. Available at: <https://arxiv.org/abs/1410.3916> (accessed October 21, 2020).
9. Hochreiter S., Bengio Y., Frasconi P., Schmidhuber J. Gradient flow in recurrent nets: the difficulty of learning long-term dependencies. In: *A Field Guide to Dynamical Recurrent Neural Networks*, 2001, pp. 114–132.
10. Esteban M.D. *Perceptrons: An Associative Learning Network*. 1997. Available at: <https://ei.cs.vt.edu/~history/Perceptrons.Esteban.html> (accessed October 21, 2020).
11. Schmidhuber J., Wierstra D., Gagliolo M., Gomez F. Training recurrent networks by Evolino. *Neural Computation*, 2007, vol. 19, pp. 757–779. DOI: 10.1162/neco.2007.19.3.757
12. Vecoven N., Ernst D., Wehenkel A., Drion G. Introducing neuromodulation in deep neural networks to learn adaptive behaviours. *PLoS ONE*, 2020, vol. 15, no. 1, art. e0227922. DOI: 10.1371/journal.pone.0227922.

### Для цитирования

Тормозов В.С. Адаптация модели нейронной сети LSTM для решения комплексной задачи распознавания образов // Программные продукты и системы. 2021. Т. 34. № 1. С. 151–156. DOI: 10.15827/0236-235X.133.151-156.

### For citation

Tormozov V.S. The adaptation of the LSTM neural network model to solve the pattern recognition complex problem. *Software & Systems*, 2021, vol. 34, no. 1, pp. 151–156 (in Russ.). DOI: 10.15827/0236-235X.133.151-156.