

**Программные системы:**  
**теория и приложения**

УДК 004.932.75'1+004.89

doi 10.25209/2079-3316-2023-14-3-3-36



## Распознавание цифровых последовательностей с использованием свёрточных нейронных сетей

Игорь Викторович **Винокуров**

Финансовый Университет при Правительстве Российской Федерации, Москва, Россия

✉ [igvvinokurov@fa.ru](mailto:igvvinokurov@fa.ru)

**Аннотация.** В статье показана актуальность задачи преобразования в текстовый формат элементов изображений, содержащих последовательности машинописных цифр. На примере распознавания табличной информации из отсканированных документов ШПК «Роскадастр» предложено возможное решение этой задачи с использованием свёрточных нейронных сетей (CNN). Описаны принципы формирования наборов данных и моделей CNN для распознавания последовательностей из двух, трёх и четырёх цифр. Приведены результаты экспериментального исследования этих моделей и показана их эффективность. Описана интеграция моделей CNN в разрабатываемую в настоящее время информационную систему (ИС), предназначенную для автоматизированного перевода отсканированных документов в их текстовые аналоги.

**Ключевые слова и фразы:** Распознавание цифр, свёрточные нейронные сети, CNN, Keras, Python

Для цитирования: Винокуров И. В. *Распознавание цифровых последовательностей с использованием свёрточных нейронных сетей* // Программные системы: теория и приложения. 2023. Т. 14. № 3(58). С. 3–36.  
[https://psta.psisras.ru/read/psta2023\\_3\\_3-36.pdf](https://psta.psisras.ru/read/psta2023_3_3-36.pdf)

## Введение

Процесс перехода на электронный документооборот, реализуемый в настоящее время многим организациями, предполагает перевод электронных копий документов в их тестовые аналоги. В ППК «Роскадастр» основными документами, подлежащими распознаванию и переводу в текстовые форматы (txt, json и другие), являются изображения таблиц с кадастровыми координатами объектов. Детектирование таблиц на изображениях документов достаточно эффективно реализуется с использованием сетей глубокого обучения YOLO [1]. Не менее эффективно с использованием CNN, могут быть распознаны и одиночные символы машинописного текста [2]. Однако при плохом качестве отсканированного документа его сегментация с использованием и стандартных (например, *findContours()* библиотеки OpenCV [3]) и собственных методов их обнаружения может привести к выявлению регионов (ROI, *Region Of Interest, англ.*), содержащих более одного символа. В этих случаях предложенный в [2] подход неприменим: требуется распознать на фрагментах изображений непрерывные последовательности из нескольких цифр.

Пример сегментов (фрагментов) изображения, содержащих от одной до трёх цифр, приведён на рисунке 1. Выявление регионов, содержащих более одной цифры, как это показано на рисунке 1б, объясняется наличием градаций серого цвета, которые пороговая подсистема трансформации цветов в ИС [2] воспринимает как элементы изображения и, как следствие, относит их к текущему ROI.



(а) Координаты объекта

(б) Формирование ROI

РИСУНОК 1. Сегментация координат объекта

Решение задачи распознавания цифровых последовательностей, при массовом автоматизированном преобразовании отсканированных изображений в их текстовые аналоги, является в настоящее время актуальной и значимой для ППК «Роскадастр».

Ниже, в разделе 1 приведены существующие в настоящее время подходы к распознаванию последовательностей цифр. Раздел 2 посвящён формированию и исследованию моделей CNN для распознавания последовательностей из двух, трёх и четырёх цифр. В разделе 3 описаны особенности выбора сформированных моделей CNN для распознавания документов с кадастровыми координатами объектов.

## 1. Анализ возможных подходов к решению задачи

В настоящее время существует достаточно большое количество работ, посвящённых задаче распознавания буквенно-цифровых последовательностей. Интерес к этой задаче объясняется тем, что результаты её решения могут быть использованы в нетривиальном компьютерном зрении с целью распознавания содержимого электронных копий документов, обработки форм с результатами опросов, идентификации почтовых индексов и адресов и т.п.

Определённый интерес с точки зрения поиска возможного подхода к решению задачи представляют методы исследовательских групп из разных стран по распознаванию рукописных буквенно-цифровых последовательностей, представленные на Международной конференции по распознаванию рукописного текста (ICFHR) 2014 [4]. Все эти методы предполагают наличие значимого этапа начальной предобработки изображения. Например, в пекинском методе из [4] для удаления шумов из простых изображений применяется бинаризация Otsu [5]. Для сложных изображений предлагается использовать два многослойных перцептрона [6] с последующей бинаризацией Сауволя [7]. После предобработки изображения генерируются шаблоны символов кандидатов, которые затем классифицируются с помощью полиномиального классификатора.

В шанхайском методе из [4] цифры сегментируются на основе адаптивной бинаризации и анализа связанных компонентов. Собственно классификация реализуется с использованием машины опорных векторов (SVM) [8] и самоорганизующихся карт Кохонена (SOM) [9]. Результаты распознавания для каждой цифры последовательности упорядочиваются по их достоверности.

Сингапурский метод [4] реализует предварительную обработку изображения с целью извлечения векторов последовательных признаков. Для этой цели используется гистограмма ориентированных градиентов (HOG) [10]. Распознавание цифровых последовательностей реализуется рекуррентной нейронной сетью (RNN) [11].

Бразильский метод [4] предварительно обрабатывает изображения с целью обнаружения цифровых последовательностей, для распознавания которых используется классификатор, состоящий из многомерной рекуррентной нейронной сети (MDRNN) [12] и машины опорных векторов (SVM) [13].

И, наконец, алжирский метод Tébéssa I [4] использует стратегию прямой сегментации, основанную на ориентированном скользящем окне [14]. Распознавание осуществляется с помощью преобразования Радона

[15] в сочетании с SVM [8]. Другой аналогичный метод, Tébéssa П [4], реализует те же действия, что и выше, но с несколькими углами ориентации скользящего окна.

В [16] предложен алгоритм, который обучает k-разрывные автокодеры и использует их в качестве предварительно обученных скрытых слоёв CNN. Для достижения высоких скоростей распознавания применяется техника раздвижных окон.

Комбинация модели ResNet [17] с двунаправленными слоями краткосрочной памяти предложена в [18]. Высокая производительность модели ResNet для извлечения функций из изображений привела к значительному улучшению распознавания цифровых последовательностей, особенно при работе с маркировкой длинных строк.

В [19] представлена CNN для маркировки рукописных строк, в которой для обработки последовательностей произвольной длины строки используется пул пространственных пирамид. Сегментация цифр в строковых последовательностях в этой работе не используется.

Интересная модель для задач маркировки буквенно-цифровых последовательностей приведена в [20]. В основе модели лежит архитектура CRNN–CNN без недостатков RNN. Такая модель немного улучшила результаты из [19], в которой использовались двунаправленные рекуррентные блоки в сочетании с моделью ResNet.

В [21] разработана система, в которой для распознавания цифр на изображениях с низким разрешением используется детектор цели на уровне символов. В основе детектора цели лежат координаты контура символа.

В большинстве приведенных выше работах использовалась комбинация нетривиальных методов предварительной обработки изображения с последующим применением нейросетевых классификаторов. Однако в настоящее время существенные успехи в области нейронных сетей привели к появлению ряда работ, в которых для распознавания в том числе и буквенно-цифровых последовательностей используются только нейронные сети того или иного типа. Этап предварительной обработки или отсутствует или крайне незначителен. Например, в [22] для обнаружения и распознавания элементов буквенно-цифровых последовательностей предложено использовать модель на основе нейронной сети глубокого обучения (DNN) YOLO [23]. В этой работе полностью исключена предварительная обработка и сегментация изображений, кроме того, YOLO хорошо справляется с распознаванием длинных последовательностей.

Помимо YOLO, для распознавания последовательностей машинописных и рукописных цифр, изолированных или контактных, могут быть использованы и CNN, например так, как это сделано в [24]. В этой работе показано, что *распознавание цифровых последовательностей может быть эффективно реализовано без существенной предварительной обработки изображений*. Критически важным для такого подхода является незашумленность (чёткость) изображения.

Ещё один пример эффективного использования CNN для распознавания строк рукописных цифр с низким разрешением приведён в [25]. Повышение точности работы CNN реализуется за счёт добавления двух выходных ветвей для классификации и регрессии по ограничивающим контурам символов. С целью дальнейшего повышения точности используются некатегоризированное подавление и тонкая минипакетная настройка.

Исходя из существующей в настоящее время тенденции распознавания буквенно-цифровых последовательностей, использование CNN с незначительной предварительной обработкой изображений представляется достаточно эффективным. В этой работе был выбран именно такой подход к распознаванию цифровых последовательностей из документов ППК «Роскадастр».

## 2. Формирование и исследование моделей CNN

Как показал анализ результатов сегментирования произвольной выборки из 2000 документов ППК «Роскадастр», последовательности из 5-и и более цифр встречаются крайне редко, рисунок 2. Исключения

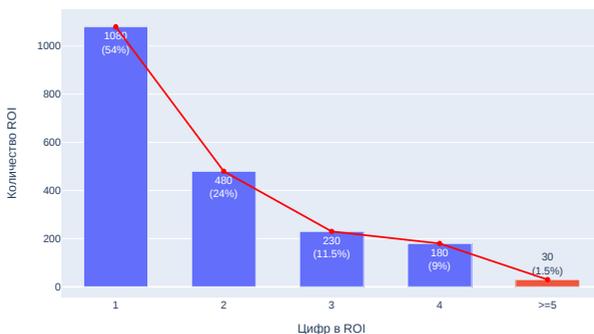


Рисунок 2. Количество ROI в выборке из 2000 отсканированных документов составляют сильно зашумленные документы или документы с неприемлемым качеством сканирования. Для таких документов, как правило,

используется ручной разбор содержимого. Как следствие, было принято решение о распознавании последовательностей, состоящих не более чем из четырёх цифр.

## 2.1. Формирование наборов данных для обучения и тестирования моделей

На начальном этапе формирования наборов данных для обучения и тестирования моделей CNN был использован предложенный в [2] подход формирования изображений цифр от 0 до 9 основными шрифтами кадастровых документов. Далее, в зависимости от количества цифр в последовательности, реализуется их случайный выбор, аугментация, варьирующая необходимым образом угол наклона и масштабирование и слияние изображений в одно, рисунок 3.



Рисунок 3. Примеры сформированных изображений последовательностей, сверху указаны метки классов

При проведении экспериментальных исследований, для последовательностей из двух, трёх и четырёх цифр, были сформированы 1500, 2000 и 3000 изображений соответственно, гарантирующих перебор всех возможных цифровых комбинаций. Для обучения CNN использовались 70% этих изображений и 30% – для валидации и тестирования. Количество изображений цифр (классов) в каждой из цифровых последовательностей одинаково, как следствие, наборы данных для обучения и тестирования являются сбалансированными.

## 2.2. Формирование моделей CNN

Экспериментальные исследования с использованием языка Python и библиотеки Keras [26] показали, что модели CNN, обеспечивающие приемлемые точности распознавания цифровых последовательностей, должны быть последовательными и иметь в двух свёрточных слоях Conv2D<sup>1</sup> по 32 фильтра с размером ядра  $3 \times 3$ , при этом размер выборки в двух дискретизирующих слоях с типом MaxPooling2D<sup>5</sup> должен быть  $2 \times 2$ .

Компиляция всех моделей осуществлялась со следующими параметрами: функция потерь *categorical\_crossentropy*<sup>2</sup>, оптимизатор *adam*<sup>3</sup> и основная метрика *accuracy*<sup>4</sup>, дающая приемлемую оценку точности для сбалансированных наборов данных.

Все модели CNN имеют одинаковые свёрточные и дискретизирующие слои и отличаются только выходными полносвязными слоями Dense<sup>5</sup>, количество которых определяется числом распознаваемых моделью CNN цифр, рисунок 4.

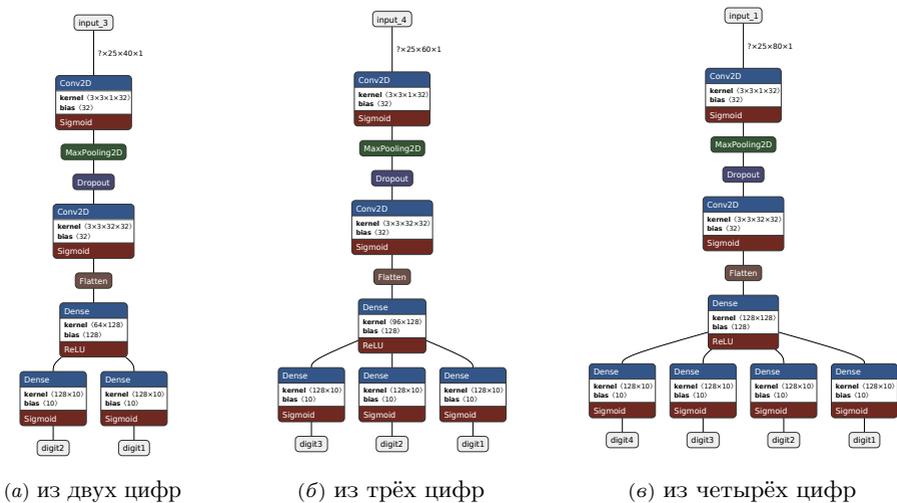


Рисунок 4. Модели CNN для распознавания последовательностей

Каждый из этих слоёв возвращает вероятность принадлежности цифры к одному из 10-ти классов.

<sup>1</sup>см. описание *Keras layers API*<sup>URL</sup>

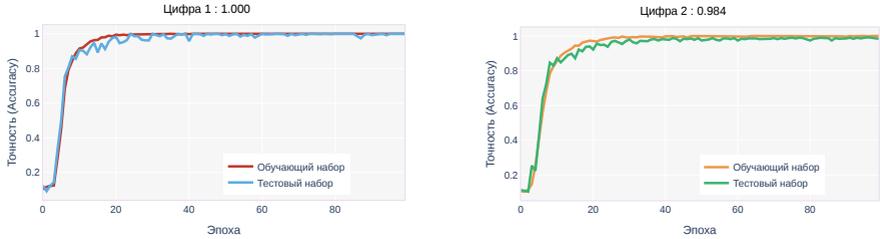
<sup>2</sup>см. описание *Keras.Losses*<sup>URL</sup>

<sup>3</sup>см. описание *Keras.Optimizers*<sup>URL</sup>

<sup>4</sup>см. описание *Keras.Accuracy metrics*<sup>URL</sup>

### 2.2.1. Исследование моделей CNN

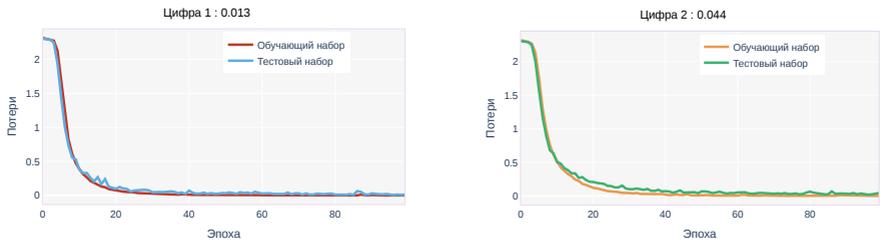
Результаты исследований моделей CNN из рисунка 4 при 100 (оптимальное значение) эпохах их обучения в виде графиков точности, функции потерь и ROC-кривых для распознавания последовательностей из двух цифр приведены на рисунке 5, 6, 7.



(а) для первой цифры

(б) для второй цифры

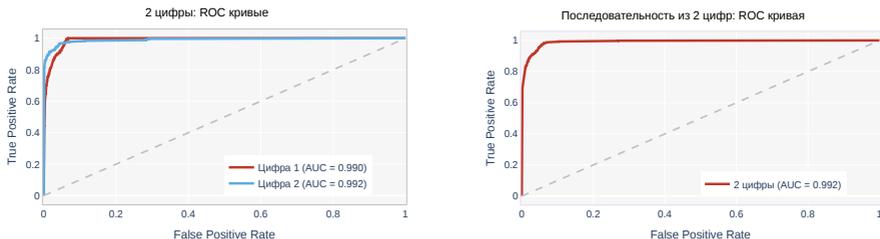
Рисунок 5. Точность распознавания пары цифр



(а) для первой цифры

(б) для второй цифры

Рисунок 6. Потери распознавания пары цифр

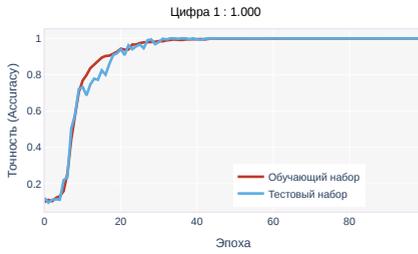


(а) каждой цифры

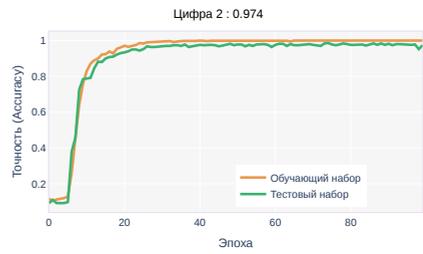
(б) последовательности цифр

Рисунок 7. ROC кривые при распознавании пары цифр

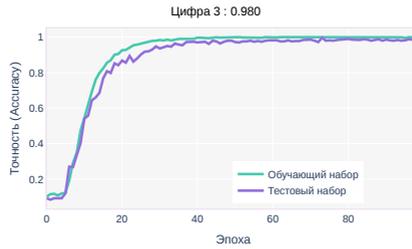
Для распознавания последовательностей из трёх цифр аналогичные результаты приведены на рисунках 8, 9, 10.



(a) для первой цифры

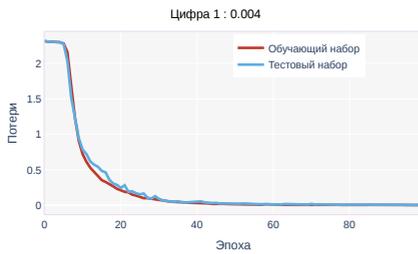


(б) для второй цифры

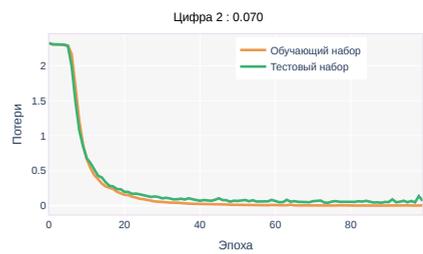


(в) для третьей цифры

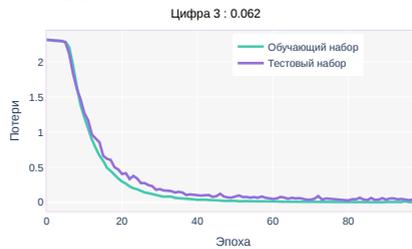
РИСУНОК 8. Точность распознавания тройки цифр



(a) для первой цифры

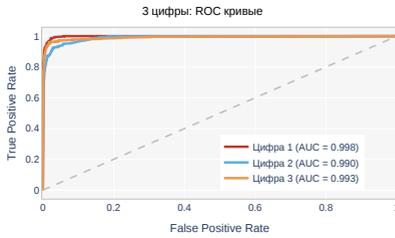


(б) для второй цифры

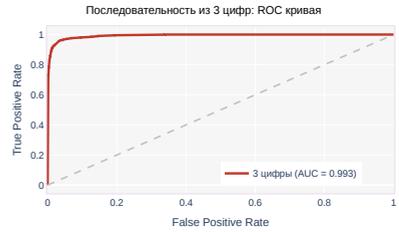


(в) для третьей цифры

РИСУНОК 9. Потери распознавания тройки цифр



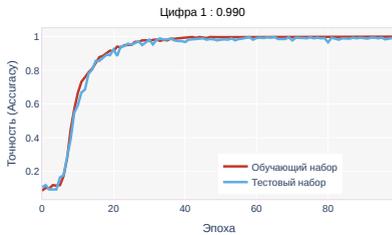
(а) каждой цифры



(б) последовательности цифр

Рисунок 10. ROC кривые при распознавании тройки цифр

Результаты распознавания последовательностей из четырёх цифр представлены на рисунках 11, 12 и 13.



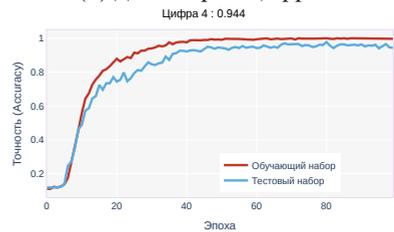
(а) для первой цифры



(б) для второй цифры



(в) для третьей цифры



(г) для четвёртой цифры

Рисунок 11. Точность распознавания четвёрки цифр

Здесь и ниже по тексту над графиками приведено значение исследуемого параметра на последней эпохе обучения CNN.

Количество эпох обучения определено экспериментально и позволяет произвести точную настройку моделей CNN для распознавания всех возможных комбинаций цифр, сгенерированных на этапе их аугментации. С целью устранения возможного переобучения, в каждую из моделей был добавлен слой Dropout<sup>5</sup>.

<sup>5</sup>см. описание *Keras layers API*<sup>URL</sup>

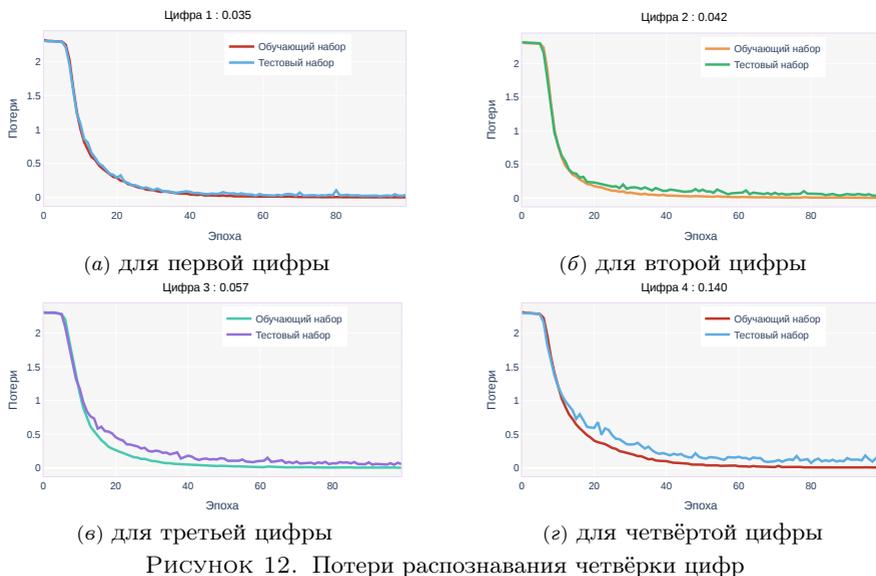


Рисунок 12. Потери распознавания четвёрки цифр

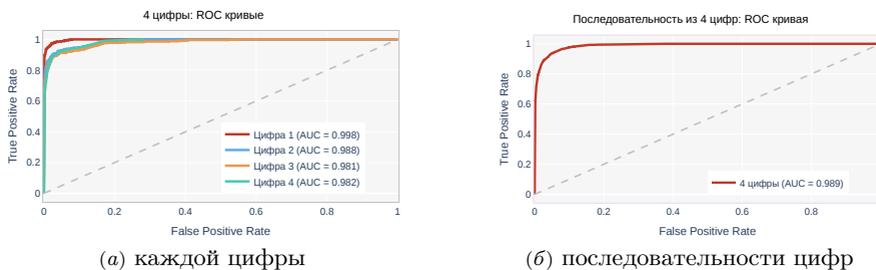


Рисунок 13. ROC кривые при распознавании четвёрки цифр

Наилучшие значения метрик *accuracy*, *recall* и *f1*<sup>6</sup>, полученных в результате проведения экспериментальных исследований моделей, приведены в таблице 1.

ТАБЛИЦА 1. Значения основных метрик точности моделей

Цифр в последовательности	Отдельная цифра ( <i>accuracy</i> , <i>recall</i> , <i>f1</i> )	Последовательность цифр ( <i>accuracy</i> , <i>recall</i> , <i>f1</i> )
2	(0.996, 0.984, 0.908)	(0.988, 0.987, 0.897)
3	(0.989, 0.980, 0.890)	(0.980, 0.984, 0.876)
4	(0.977, 0.970, 0.824)	(0.912, 0.945, 0.822)

<sup>6</sup>см. описание *Keras.Accuracy metrics*<sup>[9]</sup>

Исследование моделей CNN показало вполне ожидаемый результат – с увеличением количества цифр в последовательности снижается точность их распознавания. Например, ошибка распознавания отдельной цифры в последовательностях, приведённых на рисунке 3в, встречается чаще, чем в последовательностях из рисунка 3а и 3б. Пример ошибочного распознавания последовательности из четырёх цифр показан на рисунке 14.

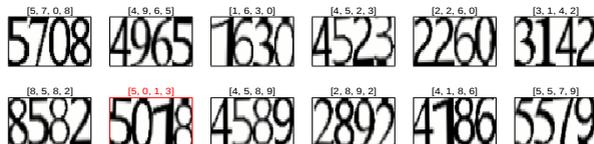


Рисунок 14. Ошибка при распознавании последовательностей из четырёх цифр

Ошибки распознавания не снижает значимости полученных результатов, поскольку они встречаются крайне редко – 1 раз на несколько десятков документов. В текущей реализации ИС ППК «Роскадастр» ошибки распознавания цифровых последовательностей не обрабатываются.

### 3. Реализация распознавания последовательностей цифр в ИС

Описанные выше модели CNN были реализованы в ИС ППК «Роскадастр», предназначенной для автоматизированного распознавания содержимого отсканированных документов, в том числе, и с координатами объектов. Расположение таблиц внутри документа детектируются с использованием сети глубокого обучения YOLOv5s [1]. Далее внутри таблиц выявляются контуры ROI и их размеры – для этой цели используется стандартная функция *findContours()* библиотеки OpenCV [3]. Например, для координат, приведённых на рисунке 1, диапазоны размеров ROI приведены на рисунке 15.



Рисунок 15. Диапазоны значений высоты и ширины ROI

Размеры ROI (в большинстве случаев это только ширина) используются для оценки количества цифр в последовательности с использованием метода кластеризации k-средних [27]. Результаты кластеризации ROI этим методом для координат из рисунка 1 показаны на рисунке 16.

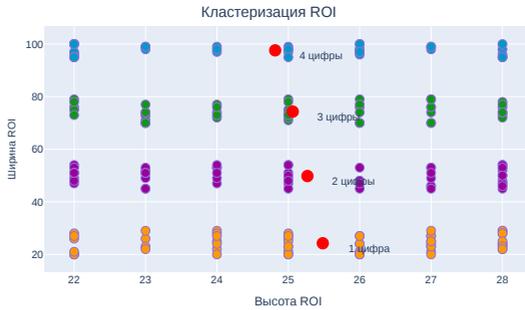


Рисунок 16. Кластеризация ROI. Красным показаны центры кластеров

Собственно процесс распознавания заключается в переборе всех ROI, соотношения их размеров с центроидами четырёх кластеров, и выборе соответствующей модели CNN, рисунок 17.

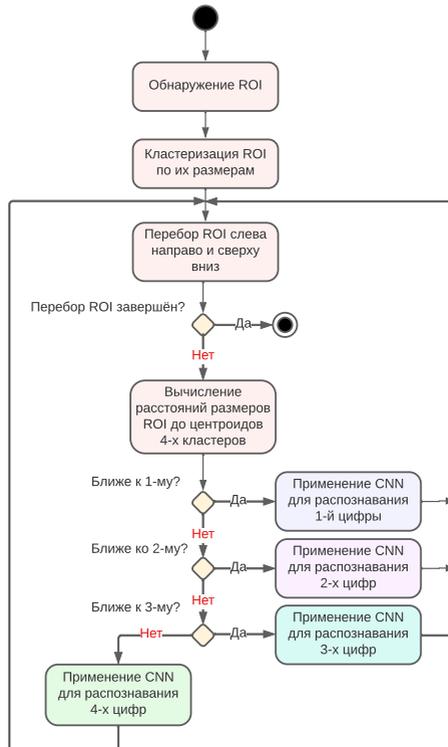


Рисунок 17. Диаграмма основных этапов работы ИС по выбору модели CNN

Пример распознавания координат объекта в ИС ППК «Роскадастр» приведён на рисунке 18.

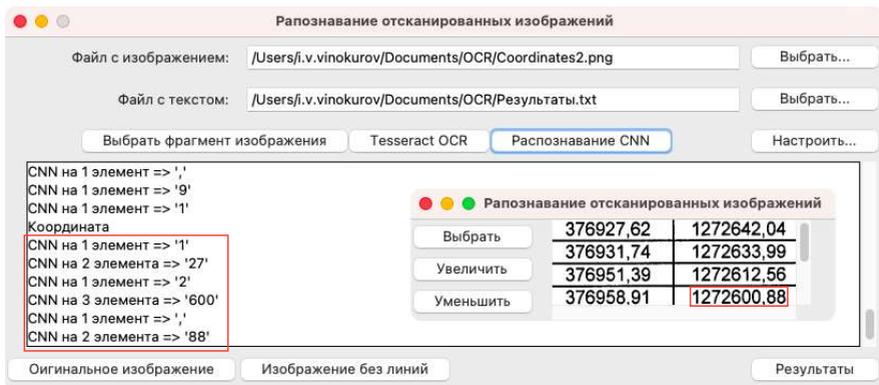


Рисунок 18. Реализация распознавания координат объекта с использованием свёрточных нейронных сетей

#### 4. Достоинства и недостатки предлагаемого подхода

К очевидным достоинствам распознавания цифровых последовательностей с использованием CNN, применительно к документам ППК «Роскадастр», относятся отсутствие необходимости в предварительной обработке изображений, как это было отмечено в [24], и достаточно высокая точность распознавания. Незначительное снижение точности распознавания второй, третьей и четвёртой цифры (см. рисунок 5б, 8в, 11г) объясняется структурой CNN, приводящей к накоплению погрешности распознавания последующих цифр.

К недостатку можно отнести неверную в ряде случаев кластеризацию ROI. Причиной этого является использование в документах ППК «Роскадастр» немоноширинных шрифтов. Неверная кластеризация приводит к тому, что в ИС выбирается модель CNN, не соответствующая цифровой последовательности.

Например, цифровая последовательность из трёх единиц оказывается ближе к центроиду кластера двух цифровых последовательностей, цифровая последовательность из четырёх единиц ближе к центроиду кластера трёх цифровых последовательностей и т.д. Подобные ситуации встречаются крайне редко, тем не менее, их появление является вполне возможным.

## Заключение

В результате проведённых исследований были разработаны модели CNN, позволившие достаточно эффективно решить задачу распознавания последовательностей из двух, трёх и четырёх цифр. Математические и экспериментальные исследования этих моделей показали приемлемую точность распознавания как отдельных цифр, так и их непрерывных последовательностей.

Разработанные модели были использованы в ИС ППК «Роскадастр» для решения достаточно значимой на данный момент задачи преобразования отсканированных документов в их текстовые аналоги. По предварительным оценкам, использование предложенного в работе подхода распознавания цифровых последовательностей с использованием CNN, позволило повысить эффективность (скорость, качество) распознавания кадастровых координат на отсканированных документах невысокого качества на 20%–40% в зависимости от качества последних.

## Список литературы

- [1] Винокуров И. В. *Распознавание табличной информации с использованием свёрточных нейронных сетей* // Программные системы: теория и приложения.– 2023.– Т. 14.– № 1(56).– С. 3–30. [doi](#) [\\*](#) [URL](#) ↑<sub>4, 14</sub>
- [2] Винокуров И. В. *Использование свёрточной нейронной сети для распознавания элементов текста на отсканированных изображениях плохого качества* // Программные системы: теория и приложения.– 2022.– Т. 13.– № 3(54).– С. 29–43. [doi](#) [\\*](#) [URL](#) ↑<sub>4, 8</sub>
- [3] Мюллер А., Гвидо С. *Введение в машинное обучение с помощью Python. Руководство для специалистов по обработке данных.*– СПб.: Изд-во «ООО Альфа-книга».– 2017.– ISBN 978-5-907515-03-1.– 480 с. ↑<sub>4, 14</sub>
- [4] Diem M., Fiel S., Kleber F., Sablatnig R., Saavedra J. M., Contreras D., Barrios J. M., Oliveira L. S. *ICFHR 2014 competition on handwritten digit string recognition in challenging datasets (HDSRC 2014)* // 2014 14th International Conference on Frontiers in Handwriting Recognition (Hersonissos, Greece, 01–04 September 2014).– Pp. 779–784. [doi](#) ↑<sub>5, 6</sub>
- [5] Otsu N. *A threshold selection method from gray-level histograms* // *IEEE Trans. Syst. Man Cybern.*– vol. 9.– 1979.– Pp. 62–66. [doi](#) ↑<sub>5</sub>
- [6] Rabelo J. C. B., Zanchettin C., C. A. B. Mello and B. Bezerra L. D. *A multi-layer perceptron approach to threshold documents with complex background*, 2011 IEEE International Conference on Systems, Man, and Cybernetics (Anchorage, AK, USA, 09–12 October 2011).– Pp. 2523–2530. [doi](#) ↑<sub>5</sub>
- [7] Sauvola J., Pietikäinen M. *Adaptive document image binarization* // *Pattern Recognition.*– 2000.– Vol. 33.– No. 2.– Pp. 225–236. [doi](#) ↑<sub>5</sub>
- [8] Cortes C., Vapnik V. *Support-vector networks* // *Mach. Learn.*– 1995.– Vol. 20.– No. 3.– Pp. 273–297. [doi](#) ↑<sub>5, 6</sub>

- [9] Kohonen T. *Self-organized formation of topologically correct feature maps* // Biol. Cybern.– 1982.– Vol. **43**.– No. 1.– Pp. 59–69.  
- [10] Dalal N., Triggs B. *Histograms of oriented gradients for human detection* // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR'05 (San Diego, CA, USA, 20–25 June 2005).– ISBN 0-7695-2372-2.– Pp. 886–893.  
- [11] Hopfield J. J. *Neural networks and physical systems with emergent collective computational abilities* // Proc. Nat. Acad. Sci. USA.– 1982.– Vol. **79**.– No. 8.– Pp. 2554–2558.  
- [12] Graves A. *Offline handwriting recognition with multidimensional recurrent neural networks* // *Guide to OCR for Arabic Scripts*, eds. Märgner V., El Abed H., London: Springer.– ISBN 978-1-4471-4072-6.– Pp. 297–313.  
- [13] Bezerra B. L. D., Zanchettin C., de Andrade V. B. *A MDRNN-SVM hybrid model for cursive offline handwriting recognition*, Artificial Neural Networks and Machine Learning — ICANN 2012, Lecture Notes in Computer Science.– vol. **7553**, Berlin: Springer.– 2012.– ISBN 978-3-642-33265-4.– Pp. 246–254.  
- [14] Gattal A., Chibani Y. *Segmentation strategy of handwritten connected digits (SSHCD)*, Image Analysis and Processing (ICIAP), Lecture Notes in Computer Science.– vol. **6979**, Berlin–Heidelberg: Springer.– 2011.– ISBN 978-3-642-24087-4.– Pp. 248–254.  
- [15] Hejazi M., G. Shevlyakov and Y.-S. Ho *Modified discrete radon transforms and their application to rotation-invariant image analysis*, 2006 IEEE Workshop on Multimedia Signal Processing, Lecture Notes in Computer Science.– vol. **6979**, Berlin–Heidelberg: Springer.– 2006.– ISBN 978-3-642-24088-1.– Pp. 429–434.  
- [16] Saabni R. *Recognizing handwritten single digits and digit strings using deep architecture of neural networks* // 2016 Third International Conference on Artificial Intelligence and Pattern Recognition, AIPR (Lodz, Poland, 19–21 September 2016).– 2016.– Pp. 1–6.  
- [17] He K., Zhang X., Ren S., Sun J. *Deep residual learning for image recognition* // 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR (Las Vegas, NV, USA, 27–30 June 2016).– Pp. 770–778.  
- [18] Zhan H., Wang Q., Lu Y. *Handwritten digit string recognition by combination of residual network and RNN-CTC*, ICONIP 2017: Neural Information Processing, Lecture Notes in Computer Science.– vol. **10639**, Cham: Springer.– 2017.– ISBN 978-3-319-70136-3.– Pp. 583–591.  
- [19] Wang Q., Lu Y. *A sequence labeling convolutional network and its application to handwritten string recognition* // *Twenty-Sixth International Joint Conference on Artificial Intelligence* (Melbourne, Australia, 19–25 August 2017).– IJCAI.– 2017.– ISBN 978-0-9992411-0-3.– Pp. 2950–2956.  
- [20] Zhan H., Lyu S., Tu X., Lu Y. *Residual CRNN and its application to handwritten digit string recognition*, ICONIP 2019: Neural Information Processing, Communications in Computer and Information Science.– vol. **1143**, Cham: Springer.– 2019.– ISBN 978-3-030-36802-9.– Pp. 49–56.  

- [21] Ma Y., Guo J., Wei W. *An exceedingly fast model for low resolution handwritten digit string recognition // 2019 IEEE 7th International Conference on Computer Science and Network Technology, ICCSNT (Dalian, China, 9–20 October 2019).*– Pp. 282–288. doi ↑6
- [22] Hochuli A. G., Britto A. S., Barddal J. P., Sabourin R., Oliveira L. E. S. *An End-to-End approach for recognition of modern and historical handwritten numeral strings // 2020 International Joint Conference on Neural Networks, IJCNN (Glasgow, UK, 19–24 July 2020).*– 2020.– Pp. 1–8. doi ↑6
- [23] Redmon J., Divvala S., Girshick R., Farhadi A. *You only look once: Unified real-time object detection // 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR (Las Vegas, NV, USA, 27–30 June 2016).*– Pp. 779–788. doi ↑6
- [24] Hochuli A. G., Oliveira L. S., Britto A. S., Sabourin R. *Handwritten digit segmentation: Is it still necessary?* Pattern Recognition.– 2018.– Vol. 78.– Pp. 1–11. doi ↑7, 16
- [25] Guo J., Xu Y. *Low resolution handwritten digit string recognition based on object detection network // 2020 International Joint Conference on Neural Networks, IJCNN (Glasgow, UK, 19–24 July 2020).*– Pp. 1–7. doi ↑7
- [26] Moolayil J. *Learn Keras for Deep Neural Networks. A Fast-Track Approach to Modern Deep Learning with Python.*– Berkeley: Apress.– 2019.– ISBN 978-1-4842-4240-7.– xv+182 pp. doi ↑9
- [27] Datta S. *Learning OpenCV 3 Application Development.*– Packt Publishing.– 2016.– ISBN 9781784391454.– 294 pp. ↑14

Поступила в редакцию 14.04.2023;  
одобрена после рецензирования 04.07.2023;  
принята к публикации 04.07.2023;  
опубликована онлайн 13.08.2023.

Рекомендовал к публикации

д.ф.-м.н. А. М. Елизаров

## Информация об авторе:



### Игорь Викторович Винокуров

Кандидат технических наук (PhD), ассоциированный профессор в Финансовом Университете при Правительстве Российской Федерации. Область научных интересов: информационные системы, информационные технологии, технологии обработки данных.

ID 0000-0001-8697-1032  
e-mail: [igvvinokurov@fa.ru](mailto:igvvinokurov@fa.ru)

Автор заявляет об отсутствии конфликта интересов.

UDC 004.932.75'1+004.89

 10.25209/2079-3316-2023-14-3-3-36

# Recognition of digital sequences using convolutional neural networks

Igor Victorovich **Vinokurov**<sup>✉</sup>

Financial University under the Government of the Russian Federation, Moscow, Russia

<sup>✉</sup>[igvvinokurov@fa.ru](mailto:igvvinokurov@fa.ru)

**Abstract.** The relevance of identifying tabular information and recognizing its contents for processing scanned documents is shown. The formation of a data set for training, validation and testing of a deep learning neural network (DNN) YOLOv5s for the detection of simple tables is described. The effectiveness of using this DNN when working with scanned documents is shown. Using the Keras Functional API, a convolutional neural network (CNN) was formed to recognize the main elements of tabular information — numbers, basic punctuation marks and Cyrillic letters. The results of a study of the work of this CNN are given. The implementation of the identification and recognition of tabular information on scanned documents in the developed IS updating information in databases for the Unified State Register of Real Estate system is described.

**Key words and phrases:** Convolutional Neural Networks, Deep Learning Neural Networks, CNN, DNN, YOLOv5s, Keras, Python

2020 *Mathematics Subject Classification:* 68T20; 68T07, 68T45

**For citation:** Igor V. Vinokurov. *Recognition of digital sequences using convolutional neural networks*. Program Systems: Theory and Applications, 2023, 14:3(58), pp. 3–36. [https://psta.psir.ru/read/psta2023\\_3\\_3-36.pdf](https://psta.psir.ru/read/psta2023_3_3-36.pdf)



FIGURE 1. Segmentation of object coordinates

## Introduction

The process of transition to electronic document management, currently implemented by many organizations, involves the conversion of electronic copies of documents into their text analogues. One of such organizations is the «Roskadastr» PLC. In this organization, the main documents subject to recognition and translation into text formats (txt, json and others) are images of tables with cadastral coordinates of objects. Table detection on document images, as it was shown in [1], is realized quite effectively using YOLO deep learning networks. Single characters of typewritten text [2] can be recognized equally effectively using CNNs. However, if the quality of the scanned document is poor, its segmentation using both standard (e.g., *findContours()* of the OpenCV [3] library) and proprietary detection methods can lead to the detection of regions (ROI, *Region Of Interest*) containing more than 1 character. In these cases, the approach proposed in [2] cannot be used, so there is a need to develop a new one that allows to recognize continuous sequences of several digits in image fragments.

An example of image segments (fragments) containing from 1 to 3 digits is given in Figure 1. Detection of regions containing more than 1 digit, as shown in Figure 1b, is explained by the presence of grayscale colors, which the threshold subsystem of color transformation in the IS [2] perceives as image elements and, as a consequence, relates them to the current ROI.

The solution of this problem, with mass automated conversion of scanned images into their text analogs, is currently urgent and significant for the «Roskadastr» PLC.

Below, in section 1 the existing approaches to the recognition of sequences of digits are given. The section 2 is devoted to the formation and investigation of CNN models for recognizing sequences of 2, 3, and 4 digits. Section 3 describes the peculiarities of selecting the generated CNN models for recognizing documents with cadastral coordinates of objects.

## 1. Analysis of possible approaches to solving the problem

At present, there are quite a large number of works devoted to the task of recognising alphanumeric sequences. The interest in this problem is explained by the fact that the results of its solution can be used in non-trivial computer vision to recognise the contents of electronic copies of documents, to process forms with survey results, to identify postal codes and addresses, etc.

Of definite interest from the point of view of finding a possible approach to solving the problem are the methods of research groups from different countries for recognising handwritten alphanumeric sequences presented at the International Conference on Handwriting Recognition (ICFHR) 2014 [4]. All these methods involve a meaningful initial image preprocessing step. For example, the Beijing method from [4] applies Otsu [5] binarisation to remove noise from simple images. For complex images, two multilayer [6] perceptrons followed by Sauvola [7] binarisation is proposed. After image preprocessing, candidate character templates are generated and then classified using a polynomial classifier.

In the Shanghai method from [4] digits are segmented based on adaptive binarisation and connected component analysis. The actual classification is implemented using a support vector machine (SVM) [8] and self-organising Kohonen maps (SOM) [9]. The recognition results for each digit of the sequence are ordered according to their confidence.

The Singapore [4] method implements image preprocessing to extract sequence feature vectors. The histogram of oriented gradients (HOG) [10] is used for this purpose. The recognition of digital sequences is realised by a recurrent neural network (RNN) [11].

The Brazilian [4] method pre-processes the images to detect digital sequences, for the recognition of which a classifier consisting of a multidimensional recurrent neural network (MDRNN) [12] and a support vector machine (SVM) [13] is used.

Finally, the Algerian Tébessa I [4] method uses a direct segmentation strategy based on an oriented sliding window [14]. Recognition is

performed using Radon transform [15] combined with SVM [8]. Another similar method, Tébéssa II [4], implements the same steps as above, but with multiple sliding window orientation angles.

In [16], an algorithm is proposed that trains k-breaking autoencoders and uses them as pre-trained hidden layer CNNs. A sliding window technique is applied to achieve high recognition rates.

A combination of the ResNet model [17] with bidirectional short-term memory layers is proposed in [18]. The high performance of the ResNet model for feature extraction from images led to a significant improvement in digital sequence recognition, especially when dealing with long string labelling.

A CNN for handwritten string labelling is presented in [19], which uses a pool of spatial pyramids to handle sequences of arbitrary string length. Segmentation of digits in string sequences is not used in this work.

An interesting model for the task of labelling alphanumeric sequences is given in [20]. The model is based on the CRNN architecture — CNN without the drawbacks of RNN. This model slightly improved the results from [19], which used bidirectional recurrent blocks in combination with the ResNet model.

In [21], a system is developed in which a character-level target detector is used to recognise digits in low-resolution images. The target detector is based on the symbol contour coordinates.

Most of the above works used a combination of non-trivial methods of image preprocessing followed by the application of neural network classifiers. However, at present, significant progress in the field of neural networks has led to a number of works in which only neural networks of one or another type are used to recognise alphanumeric sequences. The preprocessing stage is either absent or extremely insignificant. For example, in [22], a model based on the deep learning neural network (DNN) YOLO [23] is proposed to detect and recognise alphanumeric sequence elements. In this work, image preprocessing and segmentation are completely eliminated, and in addition, YOLO performs well in recognising long sequences.

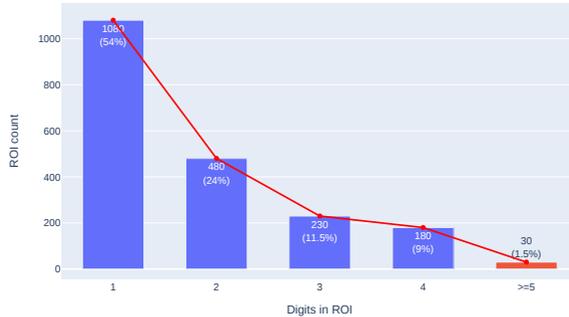


FIGURE 2. Number of ROIs in a sample of 2000 scanned documents

Besides YOLO, CNNs can also be used to recognise sequences of typewritten and handwritten digits, isolated or contacted, for example as done in [24]. In this work, it is shown that *digit sequence recognition can be efficiently implemented without significant image preprocessing*. Critical to this approach is the unnoised (sharpness) of the image.

Another example of effective use of CNNs for low-resolution handwritten digit string recognition is given in [25]. The accuracy improvement of the CNN is realised by adding 2 output branches for classification and regression on bounding character contours. In order to further improve the accuracy, unclassified suppression and fine minipacket tuning are used.

Based on the current trend of alphanumeric sequence recognition, the use of CNNs with minor image preprocessing appears to be quite effective. In this work, exactly such an approach was chosen for recognising numeric sequences from «Roskadastr» PLC documents.

## 2. Formation and exploration of CNN models

As the analysis of the results of segmentation of a random sample of 2000 documents of the «Roskadastr» PLC has shown, sequences of 5 and more digits are extremely rare, Figure 2. Exceptions are highly noisy documents or documents with unacceptable scanning quality. For such

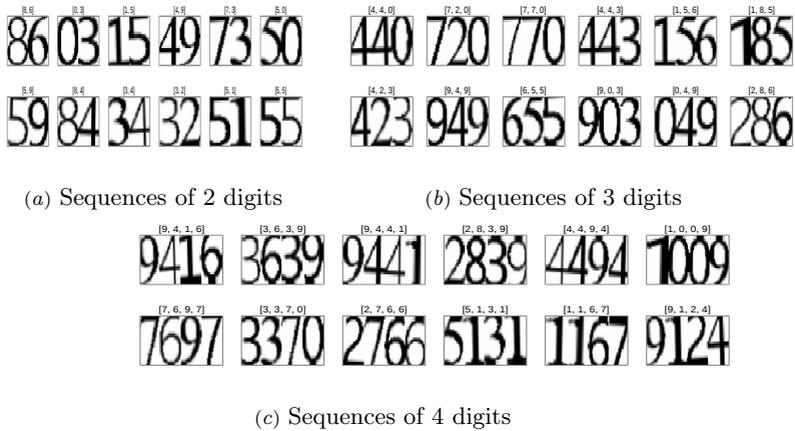


FIGURE 3. Examples of generated image sequences of 2, 3 and 4 digits. The class labels are shown at the top

documents, manual content parsing is usually used. As a consequence, it was decided to recognise sequences consisting of no more than 4 digits.

## 2.1. Generating datasets for training and testing models

At the initial stage of datasets generation for training and testing of CNN models, the approach proposed in [2] to form images of digits from 0 to 9 by the main fonts of cadastral documents was used. Then, depending on the number of digits in the sequence, their random selection, augmentation, varying the slope angle as necessary and scaling and merging the images into one, Figure 3.

In experimental studies, for sequences of 2, 3, and 4 digits, 1500, 2000 and 3000 images were generated, respectively, guaranteeing that all possible digit combinations were enumerated. About 70% of these images were used for training the CNN and 30% were used for validation and testing. The number of digit images (classes) in each of the digit sequences is the same; as a consequence, the training and testing datasets are balanced.

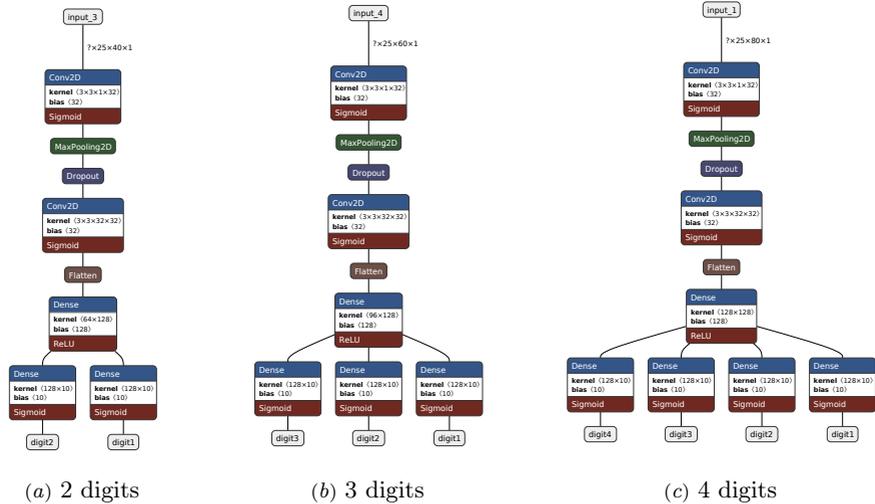


FIGURE 4. CNN models for recognising sequences of 2, 3 and 4 digits

## 2.2. Shaping the CNN models

Experimental researches using the Python language and the Keras [26] library have shown that CNN models that provide acceptable digital sequence recognition accuracies should be sequential and have 32 filters in 2 convolutional Conv2D<sup>1</sup> layers with a kernel size of  $3 \times 3$ , while the sample size in 2 sampling layers with the MaxPooling2D<sup>1</sup> type should be  $2 \times 2$ .

All models were compiled with the following parameters: loss function — *categorical\_crossentropy*<sup>2</sup>, optimiser — *adam*<sup>3</sup>, and basic metric — *accuracy*<sup>4</sup>, which gives an acceptable accuracy estimate for balanced datasets.

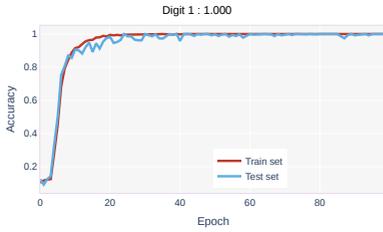
All CNN models have the same convolution and discretisation layers and differ only in the output full-link Dense<sup>1</sup> layers, the number of which is determined by the number of digits recognised by the CNN model, Figure 4. Each of these layers returns the probability that a digit belongs to one of 10 classes.

<sup>1</sup>see the description *Keras.layers.API*<sup>URL</sup>

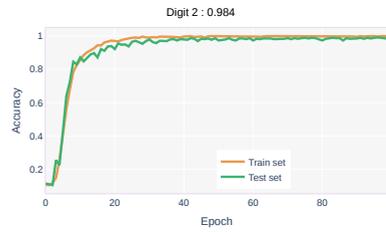
<sup>2</sup>see the description *Keras.Losses*<sup>URL</sup>

<sup>3</sup>see the description *Keras.Optimizers*<sup>URL</sup>

<sup>4</sup>see the description [https://keras.io/api/metrics/accuracy\\_metrics/](https://keras.io/api/metrics/accuracy_metrics/)<sup>URL</sup>

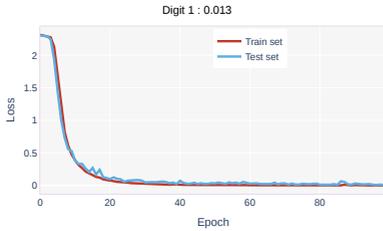


(a) Recognition accuracy of the 1st digit

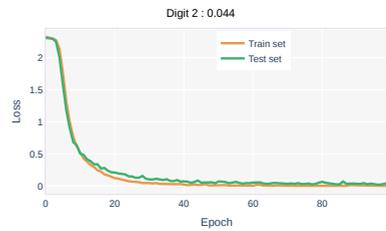


(b) Recognition accuracy of the 2nd digit

FIGURE 5. Recognition accuracy of 2 digits

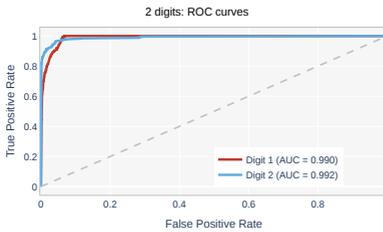


(a) Recognition loss of the 1st digit

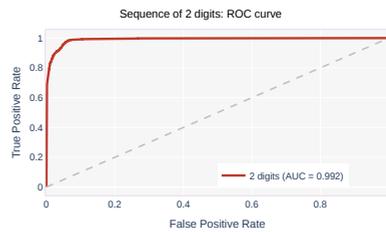


(b) Recognition loss of the 2nd digit

FIGURE 6. Recognition loss of 2 digits



(a) ROC when each of the 2 digits is recognised

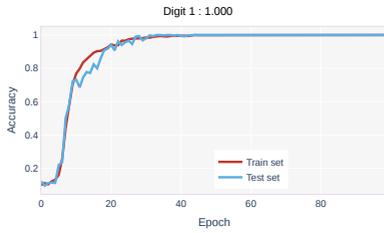


(b) ROC in recognition of sequences of 2 digits

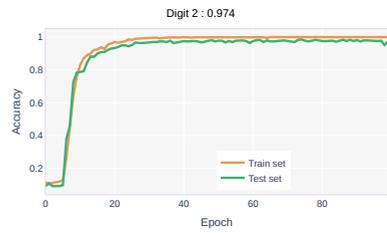
FIGURE 7. ROC curves for 2-digit recognition

### 2.2.1. CNN models research

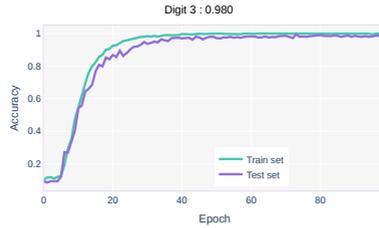
The results of the CNN models from Figure 4 at 100 (optimal) training epochs in the form of accuracy plots, loss functions and ROC curves for recognising sequences of 2 digits are shown in Figures 5, 6, 7.



(a) Recognition accuracy of the 1st digit

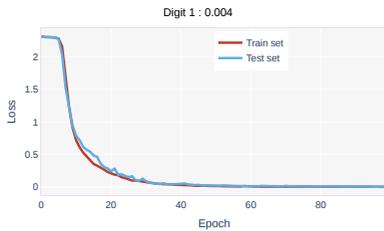


(b) Recognition accuracy of the 2nd digit

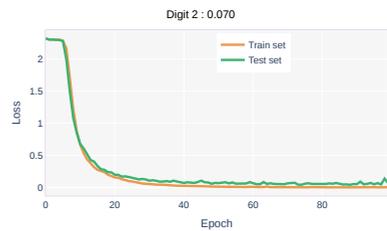


(c) Recognition accuracy of the 3rd digit

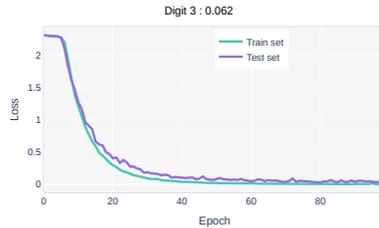
FIGURE 8. Recognition accuracy of 3 digits



(a) Recognition loss of the 1st digit



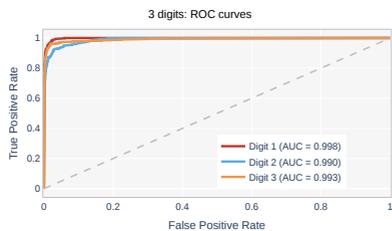
(b) Recognition loss of the 2nd digit



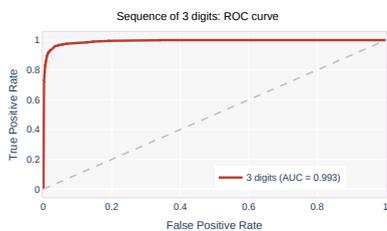
(c) Recognition loss of the 3rd digit

FIGURE 9. Recognition loss of 3 digits

For the recognition of sequences of 3 digits the results are shown in Figures 8, 9, 10.

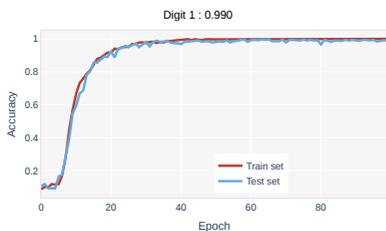


(a) ROC when each of the 3 digits is recognised

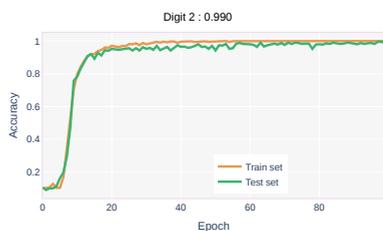


(b) ROC in recognition of sequences of 3 digits

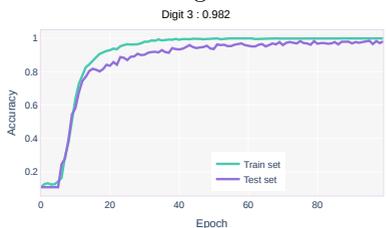
FIGURE 10. ROC curves for 3-digit recognition



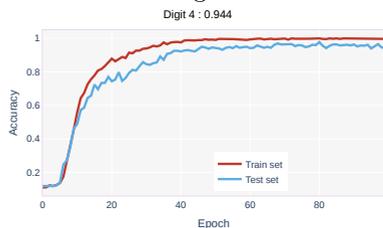
(a) Recognition accuracy of the 1st digit



(b) Recognition accuracy of the 3rd digit



(c) Recognition accuracy of the 3rd digit



(d) Recognition accuracy of the 4th digit

FIGURE 11. Recognition accuracy of 4 digits

The results of recognising sequences of 4 digits are shown in Figures 11, 12 and 13. Here and below the text above the graphs, the value of the investigated parameter at the last CNN training epoch is given.

The number of training epochs is determined experimentally and allows us to fine-tune the CNN models to recognise all possible combinations of digits generated at the stage of their augmentation. In order to eliminate possible overtraining, a Dropout<sup>5</sup> layer was added to each of the models.

<sup>5</sup>see the description *Keras layers API*<sup>URL</sup>

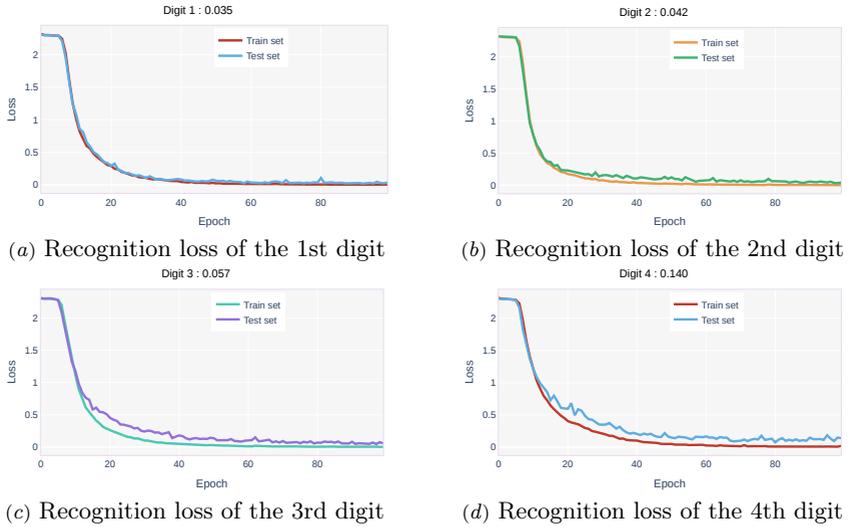


FIGURE 12. Recognition loss of 4 digits

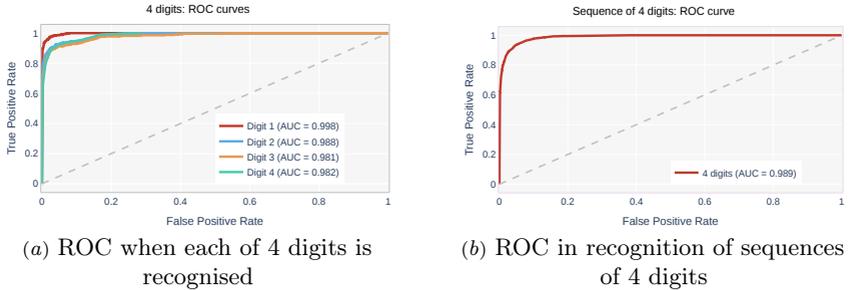


FIGURE 13. ROC curves for 4-digit recognition

TABLE 1. Values of the main metrics of model accuracy

Digits in sequences	An individual digit (accuracy, recall, f1)	A sequence of digits (accuracy, recall, f1)
2	(0.996, 0.984, 0.908)	(0.988, 0.987, 0.897)
3	(0.989, 0.980, 0.890)	(0.980, 0.984, 0.876)
4	(0.977, 0.970, 0.824)	(0.912, 0.945, 0.822)

The best values of the accuracy, recall and f1<sup>6</sup> metrics obtained from the experimental studies of the models are summarised in Table 1.

<sup>6</sup>see the description [https://keras.io/api/metrics/accuracy\\_metrics/](https://keras.io/api/metrics/accuracy_metrics/)<sup>(URL)</sup>

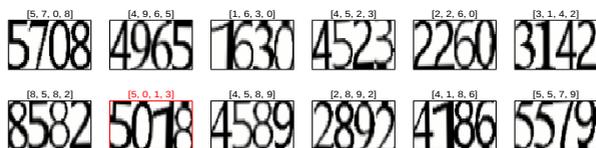


FIGURE 14. Error in recognising 4-digit sequences



FIGURE 15. ROI height and width value ranges

The study of CNN models has shown the expected result — as the number of digits in the sequence increases, the accuracy of their recognition decreases. For example, the recognition error of a single digit in the sequences shown in Figure 3c is more common than the sequences in Figure 3a and autoreffig:inputs3. An example of erroneous recognition of a sequence of 4 digits is shown in Figure 14. Recognition errors do not reduce the significance of the obtained results, as they occur extremely rarely — 1 time per several dozens of documents. In the current implementation of the «Roskadastr» PLC IS digital sequence recognition errors are not processed.

### 3. Realisation of digit sequence recognition in IS

The CNN models described above have been implemented in the «Roskadastr» PLC IS designed for automated recognition of the contents of scanned documents, including those with object coordinates. The location of tables inside the document are detected using YOLOv5s deep learning network [1]. Next, ROI contours and their sizes are detected within the tables – the standard function *findContours()* of the OpenCV [3] library is used for this purpose. For example, for the coordinates shown in Figure 1, the ranges of ROI sizes are given in Figure 15.

The dimensions of the ROI (in most cases only the width) are used to estimate the number of digits in the sequence using the K-means clustering method [27]. The results of ROI clustering by this method for the coordinates from Figure 1 are shown in Figure 16.

The actual recognition process consists of enumerating all ROIs,

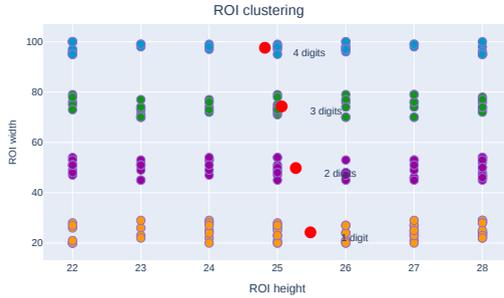


FIGURE 16. ROI clustering. Centroids of clusters are shown in red

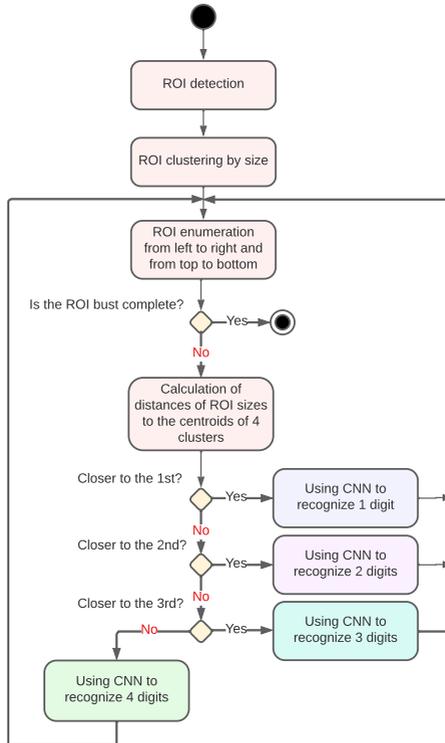


FIGURE 17. Diagram of the main stages of the IS's work on CNN model selection

correlating their sizes with the centroids of the 4 clusters, and selecting the corresponding CNN model, Figure 17.

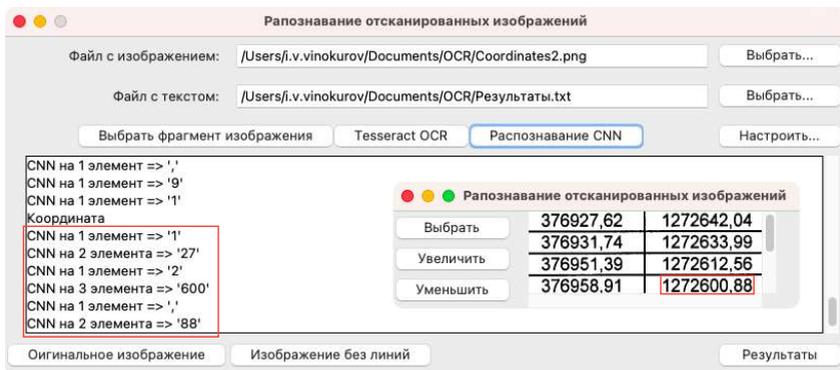


FIGURE 18. Implementation of object coordinates recognition using CNN models

An example of recognizing the coordinates of an object in the «Roskadastr» PLC IS is shown in Figure 18.

#### 4. Advantages and disadvantages of the proposed approach

The obvious advantages of recognition of digital sequences using CNN, as applied to the documents of the «Roskadastr» PLC, include the absence of the need for preliminary image processing, as it was noted in [24], and rather high recognition accuracy. A slight decrease in the recognition accuracy of the 2nd, 3rd and 4th digits (see Figures 5*b*, 8*c*, 11*d*) can be explained by the structure of the CNN, which leads to the accumulation of the recognition error of subsequent digits.

A drawback is that ROI clustering is incorrect in some cases. The reason for this is the use of non-monospaced fonts in «Roskadastr» PLC documents. Incorrect clustering leads to the fact that the IS selects a CNN model that does not correspond to the numeric sequence.

For example, a numerical sequence of 3 units turns out to be closer to the centroid of a cluster of 2 numerical sequences, a numerical sequence of 4 units is closer to the centroid of a cluster of 3 numerical sequences, etc. Such situations are extremely rare, nevertheless, their occurrence is quite possible.

## Conclusion

As a result of this research, CNN models were developed that allowed us to solve the problem of recognising sequences of 2, 3 and 4 digits quite effectively. Mathematical and experimental studies of these models have shown acceptable accuracy of recognition of both individual digits and their continuous sequences.

The developed models were used in the «Roskadastr» PLC IS to solve a rather significant at the moment task of converting scanned documents into their text analogues. According to preliminary estimates, the use of the proposed approach of digital sequence recognition using CNN allowed to increase the efficiency (speed, quality) of cadastral coordinates recognition on scanned documents of low quality by 20%–40% depending on the quality of the latter.

## References

- [1] I. V. Vinokurov. “Tabular information recognition using convolutional neural networks”, *Program Systems: Theory and Applications*, **14**:1(56) (2023), pp. 3–30 (in Russian).   [↑](#)<sub>21, 31</sub>
- [2] I. V. Vinokurov. “Using a convolutional neural network to recognize text elements in poor quality scanned images”, *Program Systems: Theory and Applications*, **13**:3(54) (2022), pp. 45–59.   [↑](#)<sub>21, 25</sub>
- [3] A. C. Müller, S. Guido. *Introduction to Machine Learning with Python*, O’Reilly Media, Inc., 2016, ISBN 9781449369897, 338 pp. [↑](#)<sub>21, 31</sub>
- [4] M. Diem, S. Fiel, F. Kleber, R. Sablatnig, J. M. Saavedra, D. Contreras, J. M. Barrios, L. S. Oliveira. “ICFHR 2014 competition on handwritten digit string recognition in challenging datasets (HDSRC 2014)”, *2014 14th International Conference on Frontiers in Handwriting Recognition* (Hersonissos, Greece, 01–04 September 2014), pp. 779–784.  [↑](#)<sub>22, 23</sub>
- [5] N. Otsu. “A threshold selection method from gray-level histograms”, *IEEE Trans. Syst. Man Cybern.*, **9**:1 (1979), pp. 62–66.  [↑](#)<sub>22</sub>
- [6] J. C. B. Rabelo, C. Zanchettin, C. A. B. Mello and B. L. D. Bezerra. “A multi-layer perceptron approach to threshold documents with complex background”, *2011 IEEE International Conference on Systems, Man, and Cybernetics* (Anchorage, AK, USA, 09–12 October 2011), pp. 2523–2530.  [↑](#)<sub>22</sub>
- [7] J. Sauvola, M. Pietikäinen. “Adaptive document image binarization”, *Pattern Recognition*, **33**:2 (2000), pp. 225–236.  [↑](#)<sub>22</sub>
- [8] C. Cortes, V. Vapnik. “Support-vector networks”, *Mach. Learn.*, **20**:3 (1995), pp. 273–297.  [↑](#)<sub>22, 23</sub>
- [9] T. Kohonen. “Self-organized formation of topologically correct feature maps”, *Biol. Cybern.*, **43**:1 (1982), pp. 59–69.  [↑](#)<sub>22</sub>

- [10] N. Dalal, B. Triggs. “Histograms of oriented gradients for human detection”, *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, CVPR’05 (San Diego, CA, USA, 20–25 June 2005), ISBN 0-7695-2372-2, pp. 886–893. doi ↑22
- [11] J. J. Hopfield. “Neural networks and physical systems with emergent collective computational abilities”, *Proc. Nat. Acad. Sci. USA*, **79**:8 (1982), pp. 2554–2558. doi ↑22
- [12] A. Graves. “Offline handwriting recognition with multidimensional recurrent neural networks”, *Guide to OCR for Arabic Scripts*, eds. Märgner V., El Abed H., Springer, London, ISBN 978-1-4471-4072-6, pp. 297–313. doi ↑22
- [13] B. L. D. Bezerra, C. Zanchettin, V. B. de Andrade. “A MDRNN-SVM hybrid model for cursive offline handwriting recognition”, *Image Analysis and Processing (ICIAP)*, Lecture Notes in Computer Science, vol. **6979**, Springer, Berlin–Heidelberg, 2011, ISBN 978-3-642-24087-4, pp. 248–254. doi ↑22
- [14] A. Gattal, Y. Chibani. “Segmentation strategy of handwritten connected digits (SSHCD)”, *Image Analysis and Processing (ICIAP)*, 2011, pp. 248–254. doi ↑22
- [15] M. Hejazi, G. Shevlyakov and Y.-S. Ho. “Modified discrete radon transforms and their application to rotation-invariant image analysis”, *2006 IEEE Workshop on Multimedia Signal Processing*, Lecture Notes in Computer Science, vol. **6979**, Springer, Berlin–Heidelberg, 2006, ISBN 978-3-642-24088-1, pp. 429–434. doi ↑23
- [16] R. Saabni. “Recognizing handwritten single digits and digit strings using deep architecture of neural networks”, *2016 Third International Conference on Artificial Intelligence and Pattern Recognition*, AIPR (Lodz, Poland, 19–21 September 2016), 2016, pp. 1–6. doi ↑23
- [17] K. He, X. Zhang, S. Ren, J. Sun. “Deep residual learning for image recognition”, *2016 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR (Las Vegas, NV, USA, 27–30 June 2016), pp. 770–778. doi ↑23
- [18] H. Zhan, Q. Wang, Y. Lu. “Handwritten digit string recognition by combination of residual network and RNN-CTC”, *ICONIP 2017: Neural Information Processing*, Lecture Notes in Computer Science, vol. **10639**, Springer, Cham, 2017, ISBN 978-3-319-70136-3, pp. 583–591. doi ↑23
- [19] Q. Wang, Y. Lu. “A sequence labeling convolutional network and its application to handwritten string recognition”, *Twenty-Sixth International Joint Conference on Artificial Intelligence* (Melbourne, Australia, 19–25 August 2017), IJCAI, 2017, ISBN 978-0-9992411-0-3, pp. 2950–2956. doi ↑23
- [20] H. Zhan, S. Lyu, X. Tu, Y. Lu. “Residual CRNN and its application to handwritten digit string recognition”, *ICONIP 2019: Neural Information Processing, Communications in Computer and Information Science*, vol. **1143**, Springer, Cham, 2019, ISBN 978-3-030-36802-9, pp. 49–56. doi ↑23
- [21] Y. Ma, J. Guo, W. Wei. “An exceedingly fast model for low resolution handwritten digit string recognition”, *2019 IEEE 7th International Conference on Computer Science and Network Technology*, ICCSNT (Dalian, China, 9–20 October 2019), pp. 282–288. doi ↑23

- [22] A. G. Hochuli, A. S. Britto, J. P. Barddal, R. Sabourin, L. E. S. Oliveira. “An End-to-End approach for recognition of modern and historical handwritten numeral strings”, *2020 International Joint Conference on Neural Networks, IJCNN* (Glasgow, UK, 19–24 July 2020), 2020, pp. 1–8.  [↑23](#)
- [23] J. Redmon, S. Divvala, R. Girshick, A. Farhadi. “You only look once: Unified real-time object detection”, *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR* (Las Vegas, NV, USA, 27–30 June 2016), pp. 779–788.  [↑23](#)
- [24] A. G. Hochuli, L. S. Oliveira, A. S. Britto, R. Sabourin. Handwritten digit segmentation: Is it still necessary? *Pattern Recognition*, **78** (2018), pp. 1–11.  [↑24, 33](#)
- [25] J. Guo, Y. Xu. “Low resolution handwritten digit string recognition based on object detection network”, *2020 International Joint Conference on Neural Networks, IJCNN* (Glasgow, UK, 19–24 July 2020), pp. 1–7.  [↑24](#)
- [26] J. Moolayil. *Learn Keras for Deep Neural Networks. A Fast-Track Approach to Modern Deep Learning with Python*, Apress, Berkeley, 2019, ISBN 978-1-4842-4240-7, xv+182 pp.  [↑26](#)
- [27] S. Datta. *Learning OpenCV 3 Application Development*, Packt Publishing, 2016, ISBN 9781784391454, 294 pp. [↑31](#)

Received	14.04.2023;
approved after reviewing	04.07.2023;
accepted for publication	04.07.2023;
published online	13.08.2023.

Recommended by

prof. A. M. Elizarov

### Information about the author:



Igor Victorovich Vinokurov

Candidate of Technical Sciences (PhD), Associate Professor at the Financial University under the Government of the Russian Federation. Research interests: information systems, information technologies, data processing technologies.

 0000-0001-8697-1032  
 e-mail: [igvvinokurov@fa.ru](mailto:igvvinokurov@fa.ru)

*The author declare no conflicts of interests.*

УДК 528.854:004.932.2

 10.25209/2079-3316-2023-14-3-37-58

## Анализ изображений растения, полученных с камеры системы автоматизированного ухода, для визуальной оценки изменения его состояния с течением времени

Александр Владимирович **Смирнов**<sup>1</sup>, Егор Сергеевич **Иванов**<sup>2✉</sup>

Институт программных систем им. А. К. Айламазяна РАН, Вельское, Россия

<sup>✉</sup>[egor.s.ivanov@gmail.com](mailto:egor.s.ivanov@gmail.com)

**Аннотация.** Предложен метод анализа изображений растений, полученных с одной камеры, для определения кроны растения и детектирования её отдельных цветовых оттенков. Рассмотрена возможность визуальной оценки состояния растений с использованием среднего арифметического взвешенного расстояния до эталонов. Приведено описание экспериментальной установки автоматизированного ухода за растениями, с помощью которой были собраны анализируемые данные.

**Ключевые слова и фразы:** Анализ изображений, определение кроны растения, цветовое пространство, компьютерное зрение, визуальная оценка состояния растений, автоматизированный уход за растениями, оценка точности

**Для цитирования:** Смирнов А.В., Иванов Е.С. *Анализ изображений растения, полученных с камеры системы автоматизированного ухода, для визуальной оценки изменения его состояния с течением времени* // Программные системы: теория и приложения. 2023. Т. 14. № 3(58). С. 37–58.  
[https://psta.psir.ru/read/psta2023\\_3\\_37-58.pdf](https://psta.psir.ru/read/psta2023_3_37-58.pdf)

## Введение

В настоящее время автоматизация затронула многие аспекты жизни, в частности выращивание различных растений с помощью систем автоматического ухода. Для таких систем крайне важно иметь механизм анализа их состояния, чтобы контролировать динамику роста, а также обеспечить взаимодействие по типу обратной связи между растением и системой.

Наиболее наглядным способом определения состояния растения является его визуальный анализ. Задача визуального анализа растений актуальна и отражена во многих исследованиях, где применялись различные подходы. В обзорной работе [1] рассматриваются базовые подходы анализа и обработки изображений растений, среди которых можно выделить следующие: определение интересующих объектов, анализ цвета, анализ формы, анализ роста растения и анализ текстуры. В итоге авторы изложили базовый алгоритм анализа изображений растений, а также представили ряд концепций для будущих исследований с использованием баз данных и облачных вычислений.

В работе [2] рассматривается применение мультиспектральной сегментации, где авторы делают упор на сегментацию изображений, используя различные цветовые пространства. Использовалась модификация пороговой сегментации с применением нескольких моделей цветового пространства. Точность сегментации в работе достигала 92%.

Работы [3, 4] также посвящены оценке состояния растения путём анализа снимков. Однако в данных работах используются нейростетевые технологии, которые позволяют достичь более достоверных результатов, но требуют набора обучающих данных.

Авторы статьи [5] разработали метод оценки динамики роста растения Резуховидка-Тая. В данном исследовании учитываются площадь, периметр, округлость и соотношение сторон растения. Стоит отметить, что обработку изображений авторы делали через стороннее программное обеспечение.

В настоящей статье предложен метод анализа изображений растений, рассчитано значение среднего арифметического взвешенного расстояния (САВР) цветов кроны к эталону для предварительной оценки состояния растений по их визуальному отображению, а также приведено описание экспериментальной установки автоматизированного ухода за растениями, с помощью которой были собраны анализируемые данные.

Следует отметить, что исследования и разработки, представленные в настоящей статье, проводились на растении томата (*Solanum lycopersicum*) и основывались на следующих закономерностях:

- (1) Крона растения в нормальном состоянии имеет естественный зелёный цвет, который характерен практически для всех растений, благодаря наличию зелёного пигмента – хлорофилла. За нормальное состояние принимается то состояние, при котором растение способно расти и развиваться без явного изменения цвета листьев.
- (2) Кроне растения в ненормальном состоянии характерно изменение цвета листьев в сторону жёлтого спектра. Такое состояние прежде всего характеризуется изменением цвета растения, что может указывать на недостаточный уход за ним: нехватка или избыток полива, нехватка света, несоблюдение температурного режима.

## **1. Экспериментальная система автоматизированного ухода за растениями**

Для сбора данных и наблюдения за ростом растения была разработана экспериментальная система автоматизированного ухода. В общем случае она представляет совокупность аппаратных и программных решений, направленных на обеспечение капельного полива в зависимости от влажности почвы и других внешних факторов.

Капельный полив осуществляется за счёт применения насосов и клапанно-релейного модуля, который управляется основным контроллером. Основной контроллер обеспечивает сбор данных с сенсоров влажности почвы посредством Wi-Fi соединения, а также управление поливом с использованием анализа данных от сенсоров и обмена данными с серверным узлом. Помимо сенсоров влажности почвы, система может быть оснащена датчиками температуры / влажности воздуха и освещённости.

Серверный узел (сервер) выполняет хранение и обработку поступающих от основного контроллера данных с установленных сенсоров. Помимо численных значений, сервер также может получать изображения растений с внутренней программируемой Wi-Fi камеры, которая устанавливается в специально спроектированный бокс. Такой бокс оснащается системой подсветки, разъемом для подключения магистрали полива и водонепроницаемой тканью, уложенной в место засыпки грунта.

Схематичное отображение функционирования экспериментальной системы представлено на рисунке 1.

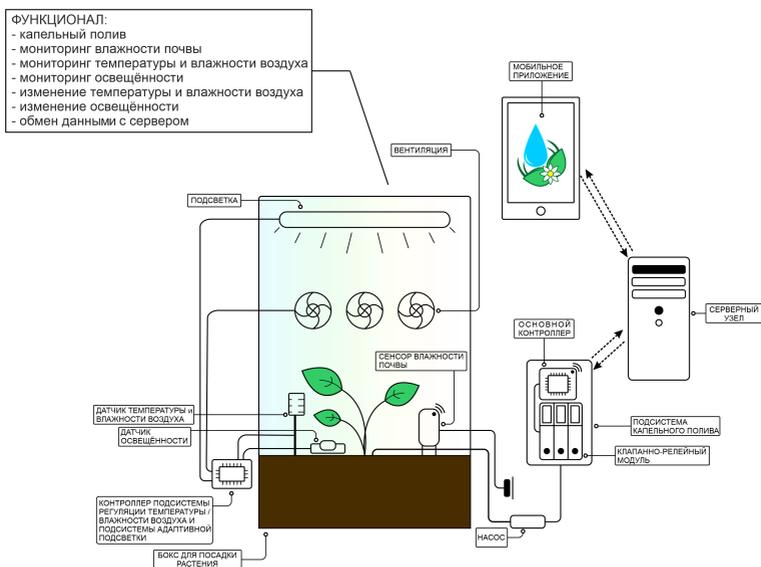


Рисунок 1. Схема функционирования экспериментальной системы автоматизированного ухода

Программируемая Wi-Fi камера с разрешением 1600x1200 пикселей установлена на верхней грани бокса и направлена вниз, на растение. Фотосъемка происходила раз в десять часов. Примеры изображений, полученных с камеры, показаны на рисунке 2.



Рисунок 2. Изображения одного растения в разные моменты времени

С помощью бокса и установленной камеры удалось собрать ряд изображений растения, на которых было зафиксировано его состояние в разные моменты времени.

## 2. Анализ изображений растения

Метод анализа полученных изображений растения состоит из следующих основных этапов:

- Определение кроны растения с применением цветового пространства CIELab.
- Квантование цвета кроны с использованием метода  $k$ -ближайших соседей.
- Определение меры близости цветов кроны к эталонам с использованием евклидовой метрики.
- Расчёт значения САВР для основных цветов кроны растения.

Указанные выше этапы будут более подробно рассмотрены в следующих пунктах настоящей статьи. Программная реализация разработанного метода выполнена на языке программирования *Python*<sup>URL</sup> с использованием библиотеки компьютерного зрения *OpenCV*<sup>URL</sup> и библиотеки визуализации инфографики *Matplotlib*<sup>URL</sup>.

### 2.1. Определение кроны растения

Данная операция необходима для того, чтобы отделить растение от фона, и позволяет проводить дальнейший анализ непосредственно участков изображения, принадлежащих растению, а не всего изображения в целом. Также благодаря определению размеров кроны, можно визуально наблюдать рост растения, а именно – увеличение площади, занимаемой им на изображении.

Алгоритм определения кроны состоит из 5 основных этапов:

1. Перевод обрабатываемого изображения из цветового пространства RGB<sup>1</sup> в цветовое пространство CIELab<sup>2</sup>. В цветовом пространстве CIELab по оси L измеряется светлота (в диапазоне от 0 до 100), отображая

---

<sup>1</sup>Общий обзор цветовых пространств см. в *skillbox*<sup>URL</sup>

<sup>2</sup>см. Цветовое пространство LAB на сайте "Спектрофотомеры"<sup>URL</sup>

коэффициент спектрального отражения, по оси  $a$  измеряется красный-зелёный оттенок, по оси  $b$  жёлтый-синий оттенок (в диапазонах от -128 до +127) (рисунок 3).

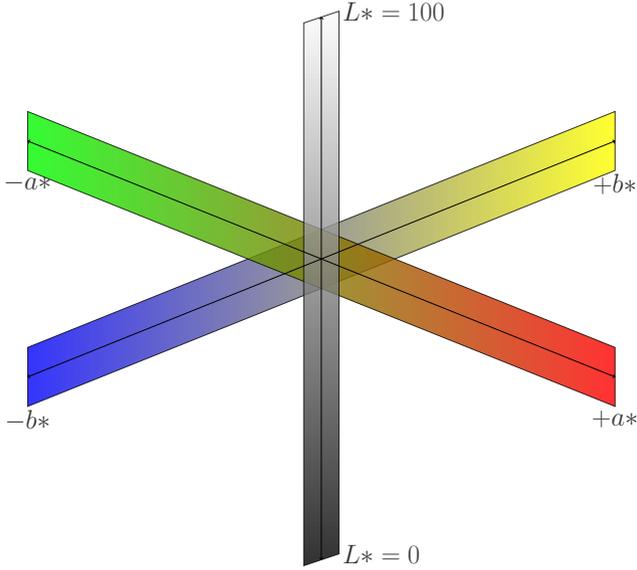


РИСУНОК 3. Цветовое пространство CIE Lab

Таким образом каналы  $a$  и  $b$  (рисунок 4) преобразованных из RGB в CIE Lab изображений будут содержать объекты, цвет которых варьируется от красного до зелёного (для канала  $a$ ) и от синего до жёлтого (для канала  $b$ ), что является искомым диапазоном цветов для детектирования растения.

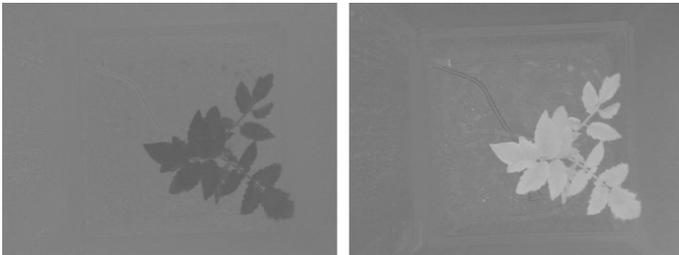


РИСУНОК 4. Визуализация каналов  $a$  и  $b$  в CIE Lab

2. Пороговая бинаризация полутоновых изображений каналов  $a$  и  $b$  в цветовом пространстве CIELab с применением метода Оцу<sup>3</sup> для автоматического вычисления порога бинаризации. В указанном методе [6] ищется порог, уменьшающий дисперсию внутри класса, которая определяется как взвешенная сумма дисперсий двух классов:

$$\sigma_{\omega}^2 = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t),$$

где

$$q_1(t) = \sum_{i=1}^t P(i), q_2(t) = \sum_{i=t+1}^I P(i),$$

$$\mu_1(t) = \sum_{i=1}^t \frac{iP(i)}{q_1(t)}, \mu_2(t) = \sum_{i=t+1}^I \frac{iP(i)}{q_2(t)},$$

$$\sigma_1^2(t) = \sum_{i=1}^t [i - \mu_1(t)]^2 \frac{P(i)}{q_1(t)}, \sigma_2^2(t) = \sum_{i=t+1}^I [i - \mu_2(t)]^2 \frac{P(i)}{q_2(t)},$$

где веса  $q_i$  – это вероятности двух классов, разделённых порогом  $t$ ,  $P(i)$  – гистограмма изображения

3. Выполнение операции побитового «ИЛИ» над полученными на предыдущем шаге бинарными изображениями. Результатом является изображение с кроной растения, сочетающее в себе объекты с каналов  $a$  и  $b$ . На рисунке 5 показан пример такого изображения.



РИСУНОК 5. Бинарное изображение кроны растения с шумом

<sup>3</sup>см. *How does Otsu's Binarization work?* в *OpenCV: Image Thresholding* URL<sup>URL</sup>

4. Контурный анализ (КА) бинарного изображения кроны растения для удаления шумов. На полученном ранее изображении помимо самой кроны также присутствуют небольшие фрагменты белого цвета, которые не принадлежат растению и являются шумом. Для фильтрации подобных шумов был применён КА.

Процедура КА предполагает нахождение некоей границы объекта, которая отделяет его от фона (других объектов), и выполняется при помощи функций используемой библиотеки компьютерного зрения (КЗ). Далее выполняется сортировка найденных контуров по возрастанию площадей, описанных ими объектов.

В ходе экспериментов было выявлено, что контур, описывающий объект наибольшей площади, полностью описывает крону растения на бинарном изображении. Имея такой контур, можно создать контурную маску путём заливки области под контуром.

После получения бинарной маски контура, вновь выполняется операция побитового «ИЛИ», только уже над полученной маской и бинарным изображением кроны. Таким образом, удалось получить результирующее бинарное изображение кроны растения (рисунок 6) без шума.



Рисунок 6. Результирующее бинарное изображение кроны растения

5. Вычисление габаритного прямоугольника кроны растения и извлечение соответствующего фрагмента из исходного снимка. Данная операция также выполняется с использованием функционала библиотеки КЗ. Результат определения кроны растения показан на рисунке 7.



Рисунок 7. Крона растения

Расчёт точности предложенного метода определения кроны растения заслуживает особого внимания. Задача определения кроны во многом схожа с задачей сегментации изображений. Под сегментацией изображений обычно подразумевается разбиение изображений на некие области, которые являются однородными по некоторому признаку. В рамках данной статьи над изображением растения выполняется операция семантической сегментации<sup>4</sup> на два класса: крона растения и фон. Основной проблемой оценки точности сегментации является поиск/генерация эталонных данных, которые, в свою очередь, зависят от детализации сегментации.

Тем не менее, авторами настоящей статьи был предпринят ряд мероприятий по оценке точности разработанного метода определения кроны растения. В частности были созданы эталонные данные, которые представляют собой бинарные изображения (маски) кроны растения. Эталонные маски создавались вручную с использованием программного обеспечения для редактирования растровой графики. На рисунке 8 показан пример эталонной маски кроны растения.

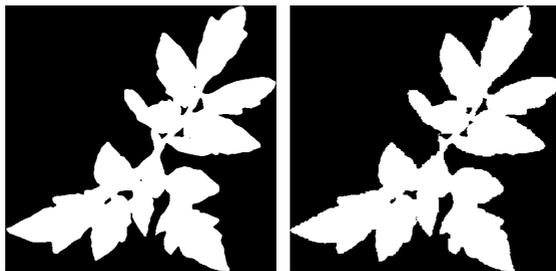


Рисунок 8. Эталонная и полученная маски кроны растения

---

<sup>4</sup>см., например, *Semantic Segmentation* в [paperswithcode](https://paperswithcode.com/)<sup>UR</sup>

Далее с использованием бинарных изображений кроны растения (эталонным и полученным) был рассчитан коэффициент Жаккара [7] – бинарная мера сходства, которая рассчитывается по формуле:

$$K_j = \frac{c}{a + b - c},$$

где  $a$  – количество белых пикселей на полученном изображении,  $b$  – количество белых пикселей на эталонном изображении,  $c$  – количество общих белых пикселей изображений.

В таблице 1 представлен результат подсчёта коэффициента Жаккара с использованием эталонных и полученных масок для нескольких изображений растения.

Таблица 1. Результат подсчета меры сходства

Порядковый номер изображения	Коэффициент Жаккара, %
1	95.4
2	95.0
3	96.4
4	96.3
5	96.4

На основании значений коэффициента Жаккара, полученных выше, средняя точность предложенного метода определения кроны растения составляет не менее 95.9%.

## 2.2. Квантование цвета кроны растения

Квантование цвета (color quantization) используется для получения малого числа характерных цветов на изображении, что приведет к упрощению последующего анализа снимка, в частности – снижает вероятность появления артефактов на следующем этапе обработки. Задачу квантования в данном случае можно сформулировать как выбор заданного количества «наилучших» цветов, имеющихся в полноцветном изображении, и замены всех остальных цветов изображения подходящими заместителями из этого списка.

Квантование цвета происходило при помощи метода  $k$ -средних<sup>5</sup>. Это метод кластерного анализа, целью которого является разделение  $m$  наблюдений (из пространства  $R^n$ ) на  $k$  кластеров, при этом каждое

<sup>5</sup>см., например, *Кластеризация: метод k-средних* на портале знаний [statistica.ru](http://statistica.ru)<sup>®</sup>

наблюдение относится к тому кластеру, к центру (центроиду) которого оно ближе всего. В качестве меры близости используется евклидово расстояние

$$p(x, y) = \|x - y\| = \sqrt{\sum_{p=1}^n (x_p - y_p)^2},$$

где  $x, y \in R^n$ .

Если имеется ряд наблюдений  $(x^1, x^2, \dots, x^m)$ ,  $x^j \in R^n$ , то метод  $k$ -средних разделяет  $m$  наблюдений на  $k$  групп (или кластеров) ( $k \leq m$ )  $S = S_1, S_2, \dots, S_k$  так, чтобы минимизировать суммарное квадратичное отклонение точек кластеров от центроидов этих кластеров:

$$\min \left[ \sum_{i=1}^k \sum_{x^{(j)} \in S_i} \|x^{(j)} - \mu_i\|^2 \right],$$

где  $x^{(j)} \in R^n$ ,  $\mu_i \in R^n$ ,  $\mu_i$  – центроид для кластера  $S_i$ .

Для изображений кроны растения характерны оттенки зелёного и жёлтого цветов. В ходе экспериментов было выявлено, что необходимое количество цветов палитры для квантования равно 5. Уменьшение количества цветов приводит к слиянию некоторых кластеров, в свою очередь увеличение приводит к разбитию некоторых кластеров, которые почти не отличаются по оттенку цвета.

Используемая библиотека компьютерного зрения не только позволяет выполнить процедуру квантования цвета, но также позволяет разложить результирующее изображение на цветовые кластеры и оперировать ими по отдельности. Результат выполнения квантования цвета показан на рисунке 9.

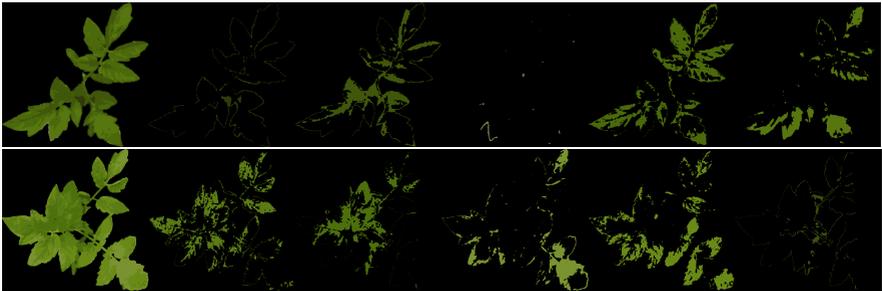


Рисунок 9. Результат выполнения квантования цвета. Справа на лево: результирующее изображение после квантования; отдельные цветовые кластеры (оттенки цветов кроны растения)

После выполнения квантования изображения кроны растения, для каждого из цветов палитры было посчитано его процентное соотношение относительно всей кроны. На рисунке 10 представлены круговые диаграммы, показывающие это отношение.

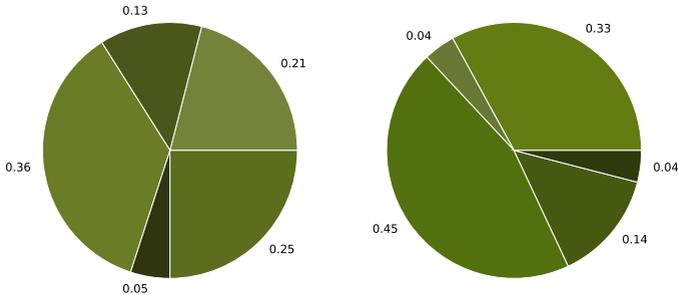


Рисунок 10. Круговые диаграммы, показывающие процентное соотношение цветов палитры квантования на примере двух изображений кроны растения

Цвет секторов, представленных выше круговых диаграмм, соответствует цвету отдельного кластера после квантования, а его размер отражает процентное соотношение данного цвета относительно всей кроны растения. Для удобства восприятия, цвета в диаграмме представлены в цветовом пространстве RGB, однако определение меры близости цвета к эталону будет выполнено с использованием цветового пространства CIE<sub>Lab</sub>.

### 2.3. Сравнение с эталонными цветами

Исходя из описанных выше правил, для определения состояния растения необходимо провести анализ его цвета, то есть определить, насколько он близок к жёлтому или зелёному. Однако измерять меру близости цвета к чисто жёлтому или зелёному не целесообразно, так как цвет листьев растения, по наблюдениям, может варьироваться от оранжево-жёлтого до тёмно-зелёного оттенков. Следовательно, для анализа цвета растения необходимо иметь свои эталонные цвета (оттенки) жёлтого и зелёного, которые являлись бы собирательными цветами для большинства растений.

Для создания эталонных цветов был выполнен поиск изображений растений, схожих с участвующим в исследовании (томат) в сети интернет. Из найденных изображений были вырезаны фрагменты, на которых преобладал интересующий цвет. В последствии, из данных фрагментов были созданы выборки для генерации эталонных жёлтого и зелёного цветов. На рисунках 11–12 показаны используемые фрагменты.

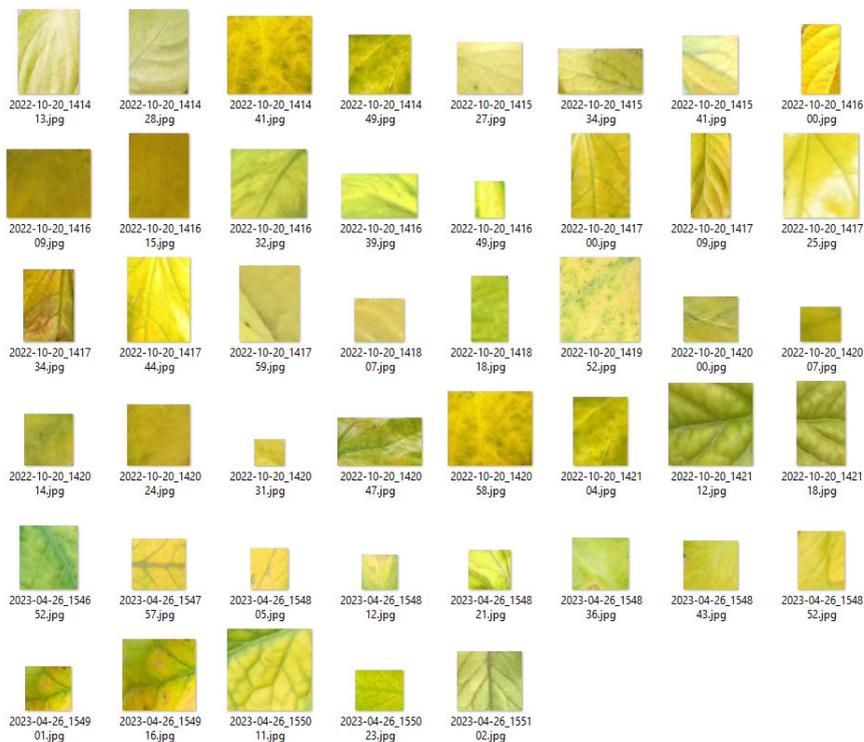


РИСУНОК 11. Фрагменты, полученные из изображений растений для генерации эталонного жёлтого цвета

Генерация эталонного цвета представляла собой процедуру получения среднеарифметических значений составляющих компонент цвета для каждого фрагмента выборки путём перебора значений пикселей в цветовом пространстве CIE Lab. Затем по аналогии рассчитывались среднеарифметические значения цвета для всей выборки.



Рисунок 12. Фрагменты, полученные из изображений растений для генерации эталонного зелёного цвета

Описанная процедура была применена к обеим полученным ранее выборкам, вследствие чего были получены эталонный жёлтый и эталонный зелёный цвета в формате CIE Lab. Для визуальной оценки полученные цвета были переведены в цветовое пространство RGB B (рисунок 13).



(а) жёлтый, #C5B638



(б) зелёный, #669348

Рисунок 13. Эталонные цвета

Расчёт меры близости цветов кроны после квантования к эталонным цветам происходил с использованием евклидовой метрики (евклидова расстояния) – расстояния между двумя точками евклидова пространства, вычисляемое по теореме Пифагора. Для точек  $p = (p_1, \dots, p_n)$  и  $q = (q_1, \dots, q_n)$  евклидово расстояние определяется по следующей формуле

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{k=1}^n (p_k - q_k)^2}.$$

В качестве координат точек использовались значения каналов  $a$  и  $b$ , которые отвечают за оттенок цвета в цветовом пространстве CIE Lab. Использование только каналов  $a$  и  $b$  позволят снизить влияния освещённости на результат расчёта меры близости, так как на уровень яркости в данном цветовом пространстве указывает значение канала  $L$ . Расчёт меры близости выполнялся со всеми цветам после квантования изображений кроны растения. Результатом выполнения данной процедуры являются данные о расстоянии между цветами кроны и эталонными цветами.

Далее в настоящей статье будет использоваться величина, которая рассчитывается путём нормализации полученного расстояния между цветами по следующей формуле:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \cdot 100,$$

где  $X$  – полученное евклидово расстояние,  $X_{min}$  – значение минимального расстояния (в данном случае равно 0),  $X_{max}$  – значение максимального расстояния (в данном случае равное евклидову расстоянию между точками с координатами (-128, -128) и (127, 127), что является минимальным и максимальным значением каналов  $a$  и  $b$  цветового пространства CIE Lab).

В итоге было получено число в диапазоне от 0 до 100, отражающее дистанцию между исследуемыми цветами, наибольшее значение которого соответствует цвету, наиболее отдаленному от эталона. Эти числа используются в качестве меры близости цветов кроны растения к эталонам. На рисунке 14 показаны гистограммы значений меры близости цветов для пары изображений кроны растения.

По полученным данным можно сделать вывод о том, что в большинстве случаев цвета кроны растения после квантования гораздо ближе к эталону

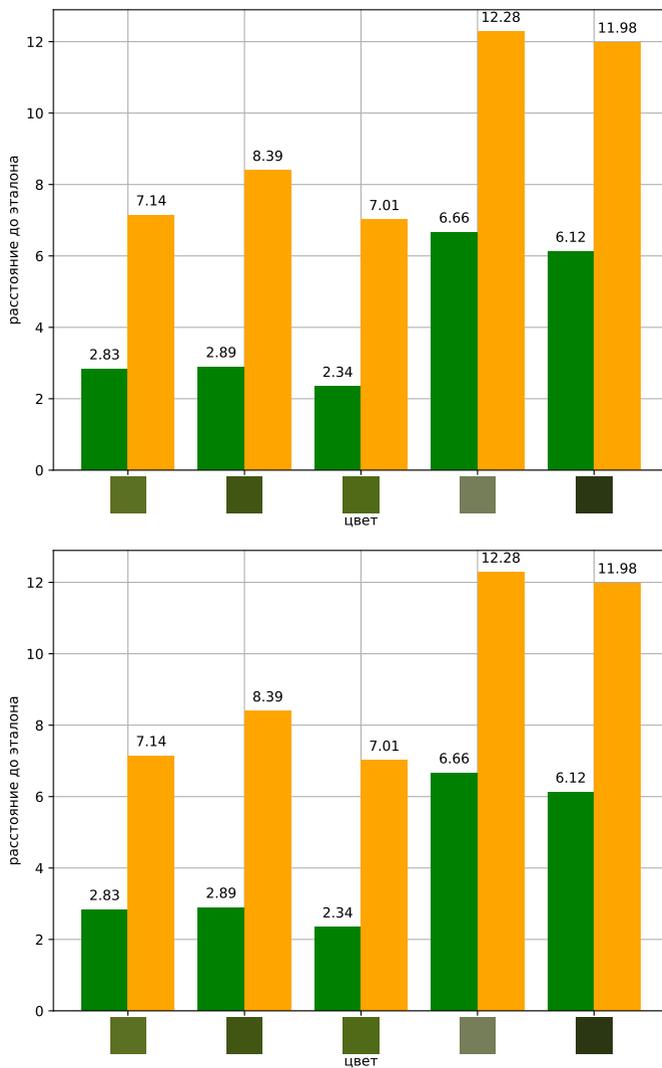


Рисунок 14. Гистограммы значений меры близости цветов кроны после квантования к эталонным цветам: зелёные столбцы – расстояние до эталона зелёного цвета; оранжевые столбцы – расстояние до эталона жёлтого цвета

зелёного цвета, чем к эталону жёлтого цвета. Это свидетельствует об адекватном росте растения и незначительном изменении его состояния, что может являться результатом работы системы автоматизированного ухода.

### 3. Оценка состояния растения

Для предварительной оценки состояния растения было рассчитано значение САВР с использованием полученных значений меры близости цветов и их процентное содержание в кроне растения (для жёлтого и зелёного цветов отдельно). Расчёт САВР происходил по формуле

$$\bar{x} = \frac{\sum_{i=1}^n \omega_i \cdot x_i}{\sum_{i=1}^n \omega_i},$$

где  $\omega_i$  – процентное содержание цвета в кроне растения,  $x_i$  – значение близости цвета к эталону.

Смысл расчёта САВР заключается в том, что полученные в результате показатели (для жёлтого и зелёного цветов) будут отражать степень близости общего цвета кроны растения к эталонам. Зная значения данных показателей, можно сделать предположение о текущем состоянии растения. Например, низкое значение САВР для эталона жёлтого цвета будет свидетельствовать о том, что данный цвет имеет доминирующее положение в цветовом балансе кроны растения, из-за чего растение будет казаться более жёлтым. В соответствии с определёнными закономерностями, представленными в разделе «Введение» настоящей статьи, данная ситуация может быть расценена как ухудшение состояния растения.

Для исследования были использованы 20 изображений растений снятых при помощи специального бокса с Wi-Fi камерой. Полученные данные значений САВР показаны в виде гистограммы на рисунке 15.

Оценка состояния растения происходит путём анализа полученных данных. На протяжении всего исследования значение САВР для зелёного цвета было ниже, чем САВР для жёлтого, что в свою очередь свидетельствует о преобладании в кроне растения оттенка цвета, близкого к зелёному эталону. Данная ситуация не является противоречием, так как существует множество зелёно-жёлтых оттенков цвета, которые одновременно содержат как зелёную, так и жёлтую составляющую.

Следует отметить тот факт, что за время проведения исследований, авторам настоящей статьи не удалось запечатлеть растение (томат)

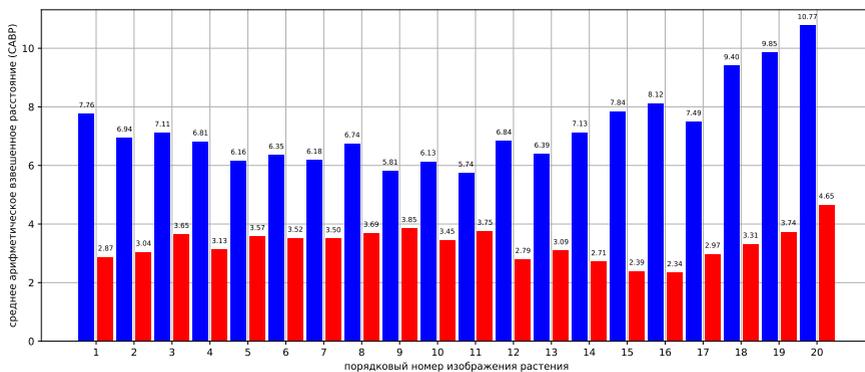


Рисунок 15. Значения САВР изображений кроны растений: столбцы синего цвета – для жёлтого эталона, красного цвета – для зелёного эталона

в ненормальном состоянии с высокой степенью повреждения кроны. Следовательно, испытать предложенный метод на полном спектре состояний растения при изменении цветового баланса его кроны в настоящий момент не представляется возможным. Несмотря на это, по теоретическим расчётам при наблюдении за растением в ненормальном состоянии, значение САВР для жёлтого цвета должно быть заметно ниже, чем для зелёного, что может послужить неким сигналом к началу мероприятий, направленных на улучшение состояния растения.

Таким образом, предварительная оценка состояния растения по его изображению заключается в сравнительном анализе САВР для зелёного и жёлтого цветов. Если значение САВР для жёлтого заметно ниже, чем для зелёного, то общий цвет кроны растения будет ближе к оттенкам жёлтого цвета, что в соответствии с закономерностью №2 настоящей статьи, свидетельствует об ухудшении состояния растения.

Подводя итог, можно сделать вывод о том, что разработанный метод оценки состояния растения выявил отсутствие значительных его изменений. Тем не менее, данный метод можно использовать в качестве некоего сенсора состояния растения при проектировании систем автоматизированного ухода, а значение САВР можно использовать как метку для набора данных при обучении нейронных сетей. Однако следует помнить тот факт, что сам по себе разработанный метод может лишь указать на изменение состояния

растения в ту или иную сторону, но не способен распознать причину такого изменения. На данном этапе исследования оценку состояния растения, полученную при использовании разработанного метода, должен изучить эксперт-ботаник для составления стратегии дальнейших действий.

#### 4. Заключение

В настоящей статье был предложен метод анализа изображений растений для определения кроны растения и детектирования её отдельных цветовых оттенков с использованием цветового квантования. Выполнен расчёт меры близости цветов кроны растения к эталонным цветам, а также подсчитано их процентное соотношение.

Рассмотрена возможность визуальной оценки состояния растений в соответствии с определёнными закономерностями на основании САВР, рассчитанного авторами настоящей статьи. Полученные в результате исследования данные о состоянии растения могут быть использованы в качестве сенсорных данных систем автоматизированного ухода за растениями.

Приведено описание экспериментальной установки автоматизированного ухода за растениями, с помощью которой были собраны анализируемые данные.

#### Список литературы

- [1] S.D. Gupta, Y. Ibaraki *Image analysis for plants: Basic procedures and techniques* // *Plant Image Analysis: Fundamentals and Applications*, Boca Raton: CRC Press.– 2014.– ISBN 9780429072345.– Pp. 25–40.  [↑38](#)
- [2] А.Г. Зотин, Е.Ю. Золотарева *Применение мультиспектральной сегментации для анализа состояния растений по данным видеорегистратора* // Программные продукты и системы.– 2011.– № 4.– С. 113–117.  [↑38](#)
- [3] М.Я. Брагинский, Д.В. Тараканов *Оценка состояния растений с использованием сверточных нейронных сетей* // Вестник кибернетики.– 2021 10.34822/1999-7604-2021-1-41-50.– № 1(41).– С. 41–50.  [↑38](#)
- [4] J. Huixian *The analysis of plants image recognition based on deep learning and artificial neural network* // IEEE Access.– 2020.– Vol. 8.– Pp. 68828–68841.  [↑38](#)
- [5] F. Vasseur, J. Bresson, G. Wang, R. Schwab, D. Weigel *Image-based methods for phenotyping growth dynamics and fitness components in Arabidopsis thaliana* // Plant Methods.– 2018.– Vol. 14.– id. 63.– 11 pp.  [↑38](#)

- [6] N. Otsu *A threshold selection method from gray-level histograms* // IEEE Transactions on Systems, Man, and Cybernetics.– 1979.– Vol. **9**.– No. 1.– Pp. 62–66.  ↑43
- [7] P. Jaccard *Distribution de la flore alpine dans le Bassin des Dranses et dans quelques regions voisines* // Bull. Soc. Vaudoise Sci. Nat.– 1901.– Vol. **37**.– No. 140.– Pp. 241–272.  ↑46

Поступила в редакцию 19.05.2023;  
 одобрена после рецензирования 22.06.2023;  
 принята к публикации 29.08.2023;  
 опубликована онлайн 13.09.2023.

Рекомендовал к публикации

*к.т.н. В. П. Фраленко*

### Информация об авторах:



#### Александр Владимирович Смирнов

Младший научный сотрудник лаборатории методов обработки и анализа изображений Института программных систем имени А. К. Айламазяна РАН. Область научных интересов: компьютерное зрение, нейронные сети, обработка изображений.

 0000-0002-7104-1462  
 e-mail: [asmirnov\\_1991@mail.ru](mailto:asmirnov_1991@mail.ru)



#### Егор Сергеевич Иванов

Младший научный сотрудник лаборатории методов обработки и анализа изображений Института программных систем имени А. К. Айламазяна РАН. Область научных интересов: компьютерное зрение, анализ данных, обработка изображений.

 0000-0002-5593-4404  
 e-mail: [egor.s.ivanov@gmail.com](mailto:egor.s.ivanov@gmail.com)

*Все авторы сделали эквивалентный вклад в подготовку публикации.  
 Авторы заявляют об отсутствии конфликта интересов.*

UDC 528.854:004.932.2

 10.25209/2079-3316-2023-14-3-37-58

## Analysis of images of a plant obtained from a camera of an automated care system to visually assess the change in its state over time

Alexandr Vladimirovich **Smirnov**<sup>1</sup>, Egor Sergeevich **Ivanov**<sup>2</sup>

Ailamazyan Program Systems Institute of RAS, Ves'kovo, Russia

<sup>2✉</sup>[egor.s.ivanov@gmail.com](mailto:egor.s.ivanov@gmail.com)

**Abstract.** A method is proposed for analyzing plant images obtained from a single camera to determine the crown of a plant and detect its individual color shades. The possibility of visual assessment of the state of plants is considered. A description is given of the experimental installation for automated plant care, with the help of which the analyzed data were collected. (*In Russian*).

**Key words and phrases:** Image analysis, plant crown determination, color space, computer vision, plant state evaluation, automated plant care, accuracy evaluation

2020 *Mathematics Subject Classification:* 68U10; 68T40

**For citation:** Alexandr V. Smirnov, Egor S. Ivanov. *Analysis of images of a plant obtained from a camera of an automated care system to visually assess the change in its state over time.* Program Systems: Theory and Applications, 2023, **14**:3(58), pp. 37–58. (*In Russ.*). [https://psta.psiras.ru/read/psta2023\\_3\\_37-58.pdf](https://psta.psiras.ru/read/psta2023_3_37-58.pdf)

## References

- [1] Gupta S.D., Ibaraki Y.. “Image analysis for plants: Basic procedures and techniques”, *Plant Image Analysis: Fundamentals and Applications*, CRC Press, Boca Raton, 2014, ISBN 9780429072345, pp. 25–40. 
- [2] Zotin A.G., Zolotareva E.Yu.. “Application of multispectral segmentation for the green vegetation status analysis based on video”, *Programmnyye produkty i sistemy*, 2011, no. 4, pp. 113–117 (in Russian). 
- [3] Braginskij M.Ya., Tarakanov D.V.. “Estimation of plants health using convolutional neural networks”, *Vestnik kibernetiki*, 2021 10.34822/1999-7604-2021-1-41-50, no. 1(41), pp. 41–50 (in Russian). 
- [4] Huixian J.. “The analysis of plants image recognition based on deep learning and artificial neural network”, *IEEE Access*, **8** (2020), pp. 68828–68841. 
- [5] Vasseur F., Bresson J., G. Wang, Schwab R., Weigel D.. “Image-based methods for phenotyping growth dynamics and fitness components in *Arabidopsis thaliana*”, *Plant Methods*, **14** (2018), id. 63, 11 pp. 
- [6] Otsu N.. “A threshold selection method from gray-level histograms”, *IEEE Transactions on Systems, Man, and Cybernetics*, **9**:1 (1979), pp. 62–66. 
- [7] Jaccard P.. “Distribution de la flore alpine dans le Bassin des Dranses et dans quelques regions voisines”, *Bull. Soc. Vaudoise Sci. Nat.*, **37**:140 (1901), pp. 241–272. 

УДК 004.652.3, 616-079.4

 10.25209/2079-3316-2023-14-3-59-94

## Построение этиопатогенетического образа концептов метатезауруса UMLS с использованием графовых метрик

Павел Андреевич Астанин<sup>1✉</sup>, Светлана Евгеньевна Раузина<sup>2</sup>,  
Татьяна Васильевна Зарубина<sup>3</sup>

Российский национальный исследовательский медицинский университет им. Н. И. Пирогова, Москва, Россия

<sup>1✉</sup>med\_cyber@mail.ru

**Аннотация.** Разработка средств информационной поддержки принятия клинических решений (ППКР) является актуальной задачей медицинской информатики. Довольно часто в системах ППКР используются информационно-поисковые алгоритмы, важным этапом проектирования которых служит создание средств автоматического распознавания этиопатогенетического образа заболеваний при работе с неструктурированным текстом. В настоящей статье произведены обзор и сравнительная характеристика аналитических метрик, применимых для построения образа концептов метатезауруса Unified Medical Language System (UMLS), представленного в виде графовой информационной модели. Предложен собственный вариант графовой метрики, показавший наибольшую эффективность при решении данной задачи.

**Ключевые слова и фразы:** медицинская информационная система, информационно-поисковый алгоритм, база знаний, теория графов, UMLS

**Благодарности:** работа выполнена за счет средств стратегического проекта «Приоритет-2030» на базе Института цифровой трансформации медицины (ИЦТМ) ФГАОУ ВО «Российский национальный исследовательский медицинский университет имени Н. И. Пирогова» Минздрава России.

Для цитирования: Астанин П. А., Раузина С. Е., Зарубина Т. В. *Построение этиопатогенетического образа концептов метатезауруса UMLS с использованием графовых метрик* // Программные системы: теория и приложения. 2023. Т. 14. № 3(58). С. 59–94. [https://psta.psiras.ru/read/psta2023\\_3\\_59-94.pdf](https://psta.psiras.ru/read/psta2023_3_59-94.pdf)

## Введение

Unified medical language system (UMLS) является крупнейшим сводом биомедицинских справочников и словарей, применимых в работе с неструктурированными данными [1]. Актуальная версия UMLS (2022AB) обеспечивает терминологический охват свыше 4.6 млн концептов – уникальных междисциплинарных понятий, классифицированных по тематической принадлежности на 127 групп. Семантически близкие концепты соединены связями, однозначно отнесенными к 9 основным (и 2 дополнительным) типам и 992 уточняющим необязательным подтипам (уточнениям). Практически каждый концепт UMLS связан хотя бы с одним другим концептом, что позволяет представить данный свод терминов в виде ориентированного мультиграфа с 98 млн уникальных связей.

Организация данных в виде графовых информационных моделей имеет ряд преимуществ, среди которых следует выделить наличие больших возможностей для оптимизации аналитических операций и существование средств, обеспечивающих наглядную интерпретацию структуры знаний на пользовательском уровне [2]. Однако в настоящее время не существует единого подхода к извлечению релевантных знаний из графовых информационных моделей [3]. Применение простых инструментов автоматического извлечения знаний из UMLS (например, фильтров на типы связей и тематические группы терминов) не приводит к клинически значимому результату по причине недостатка прямых связей между концептами метатезауруса и неоднородности структуры знаний. Необходимо ансамблирование разнородных аналитических инструментов с использованием сложных систем весовых коэффициентов, оптимизированных валидированных аналитических инструментов и метамоделей – унифицированных сводов правил технической реализации моделей знаний.

Одним из элементов системы весовых коэффициентов сущностей UMLS могут стать значения, отражающие степень принадлежности каждого термина к отдельным клиническим профилям (пульмонологии, кардиологии, гастроэнтерологии), а также свидетельствующие о наличии этиологической или патогенетической связи с соответствующими областями. Каждый профиль включает патологические состояния, относящиеся к определенному классу заболеваний. Набор подобных значений для отдельного концепта формирует его этиопатогенетический образ – вектор ненулевых значений функции принадлежности термина к характерным клиническим профилям.

Создание этиопатогенетического образа для концептов UMLS позволит определять степень их значимости в контексте решаемых задач и обеспечивать значительное сокращение ширины поиска при выполнении графовых запросов. Вычисление мер близости между векторами функции

принадлежности сущностей к соответствующим медицинским областям даст возможность ранжирования симптомов и заболеваний по степени их клинического сходства, что может стать основой автоматического формирования дифференциально-диагностических рядов. Автоматическое построение этиопатогенетического образа терминов UMLS позволит сравнить существующие графовые метрики и обосновать их применение при решении задач поиска релевантных знаний. В свою очередь, успешное построение системы весовых коэффициентов для концептов метатезауруса UMLS обеспечит значительное улучшение информационно-поисковых алгоритмов за счет повышения математической контрастности получаемых результатов.

*Цель* настоящего исследования заключается в построении и оценке этиопатогенетического образа концептов метатезауруса UMLS с использованием графовых метрик для анализа связности узлов.

## 1. Постановка задач

Для достижения цели вся работа была разделена на несколько последовательных этапов, соответствующих задачам исследования. Первая задача заключалась в подготовке графовой информационной модели UMLS и подключении русскоязычных справочников терминов. Вторая задача состояла в создании семантических и логических правил включения формулировок заболеваний в свод для построения этиопатогенетического образа концептов. В рамках третьей задачи осуществлялись проектирование и программная реализация алгоритма вычисления функций принадлежности концептов к клиническим профилям. Наконец, в рамках четвертой задачи выполнены обзор и сравнительная оценка различных графовых метрик, применявшихся для построения этиопатогенетического образа концептов-нозологий. Также в ходе данного этапа был разработан собственный вариант графовой метрики, предназначенный для оценки связности концептов UMLS.

## 2. Графовая информационная модель UMLS

Техническая реализация графовой информационной модели UMLS осуществлялась с использованием стандарта Ecore. В построении модели были задействованы исходно переведенные на русский язык термины из справочников MedDRA (Medical Dictionary for Regulatory Activities Terminology), LOINC (Logical Observation Identifiers Names and Codes) и MeSH (Medical Subject Headings). MedDRA позволяет достичь терминологического покрытия большинства возможных вариантов описания клинической картины заболеваний, лабораторных и инструментальных находок, используемых в практике [4]. LOINC является унифицированным

стандартом в области описания клинико-лабораторных исследований [5]. MeSH содержит вертикальную иерархию заголовков и их синонимов, применяемых для распределения научных статей по предметным рубрикам, и пригодных к использованию в качестве словаря для описания биомедицинских областей [6].

Графовая информационная модель, содержащая русскоязычные формулировки концептов UMLS и связи между ними, оказалась значительно меньше исходной англоязычной версии: всего в полученном графе содержалось 144 тыс. (3.1%) концептов и 7.9 млн (8.0%) связей между ними. Все концепты относились к одной или нескольким группам из 124 (97.6%) возможных. Отсутствующие в русскоязычной версии UMLS тематические группы не имеют прямого отношения к описанию возможных вариантов клинических проявлений и инструментально-лабораторных методов исследования.

Для разметки терминов по клиническим направлениям оценивалась графовая связность их англоязычных аналогов с концептами-нозологием, соотнесенными с кодами справочника международной классификации болезней 10 пересмотра (МКБ-10) в UMLS. Использовано сопоставление кодов МКБ-10 с актуальной версией справочника из реестра нормативно-справочной информации Минздрава России.

### 3. Подготовка перечня нозологий для разметки концептов

Задача предварительной диагностики не предполагает точной формулировки диагноза, однако привязка симптомов к нозологическим группам требует введения ограничений на их перечень. Подготовлен набор правил отсечения, позволивших снизить количество нозологических единиц для поиска. К их числу относятся наименования, содержащие слова «неуточненный», «классифицированный», «идентифицированный», «другой», а также общие формулировки, содержащие в подстроках другие нозологические единицы. Например, код «K29» – «Гастрит и дуоденит» можно считать избыточным по причине существования кодов «K29.7» – «Гастрит» и «K29.8» – «Дуоденит», служащих для него подстроками.

Для дополнительного сокращения списка потенциальных нозологий экспертным способом были отобраны клинические симптомы и синдромы, не являющиеся самостоятельными заболеваниями (например, K92.0 – «Кровавая рвота»). Далее для всех кодов МКБ-10, с которыми соотнесено более одного узла, осуществлен отбор концептов с наибольшим числом прямых связей. Указанные процедуры позволили снизить степень потенциального смещения оценок принадлежности терминов к классам и сократить перечень заболеваний с 11232 до 1577. Концептам, соотнесенным с кодами МКБ-10, автоматически присваивалась метка принадлежности к определенному клиническому профилю, согласно таблице 1.

Таблица 1. Основные клинические профили, использованные для построения этиопатогенетического образа концептов UMLS

№	Класс МКБ-10 и диапазон кодов	Наименование профиля
1	V(F00-F99)	Психиатрический
2	VI(G00-G99)	Неврологический
3	VII(H00-H59)	Офтальмологический
4	VIII(H60-H95)	Сурдологический
5	IX(I00-I99)	Кардиологический
6	X(J00-J99)	Пульмонологический
7	XI(K00-K93)	Гастроинтестинальный
8	XII(L00-L99)	Дерматологический
9	XIII(M00-M99)	Ортопедический
10	XIV(N00-N99)	Урогенитальный

Для разметки не использовались состояния, представленные в классах:

- I некоторые инфекционные и паразитарные болезни,
- II новообразования,
- III болезни крови, кроветворных органов и отдельные нарушения, вовлекающие иммунный механизм,
- IV болезни эндокринной системы, расстройства питания и нарушения обмена веществ,
- XV беременность, роды и послеродовой период,
- XVI отдельные состояния, возникающие в перинатальном периоде,
- XVII врожденные аномалии [пороки развития], деформации и хромосомные нарушения,
- XVIII симптомы, признаки и отклонения от нормы, выявленные при клинических и лабораторных исследованиях, не классифицированные в других рубриках,
- XIX травмы, отравления и некоторые другие последствия воздействия внешних причин,
- XX внешние причины заболеваемости и смертности,
- XXI факторы, влияющие на состояние здоровья населения и обращения в учреждения здравоохранения и
- XXII коды для особых целей.

#### 4. Алгоритм построения этиопатогенетического образа концептов-симптомов UMLS

Построение этиопатогенетического образа для концептов-симптомов осуществлялось итеративным циклом, состоящим из нескольких шагов. Первый шаг заключался в агрегации всех терминов, напрямую или косвенно связанных с корневым концептом одним или несколькими из следующих

типов связей: SIB (sibling relationships – горизонтальные связи между близкородственными концептами), RO (other relationships – вертикальные связи несинонимичных терминов), CHD (child relationships) и RN (narrower relationships) – вертикальные связи родительских терминов с дочерними, RQ (related and possibly synonymous relationships – горизонтальные и вертикальные связи близкородственных или синонимичных терминов) и SY (synonymous relationships – строго синонимичные связи). Краткая характеристика основных типов связей UMLS дана в таблице 2 [7].

ТАБЛИЦА 2. Характеристика основных типов связей между концептами UMLS

Класс связей	Тип связи	Пример связи	
		Корневой термин	Концевой термин
Ассоциативные	RO	Боль в спине	Поясничный радикулит
	RQ	Боль в спине	Ощущения дискомфорта в спине
	RL	Гиперурикемия	Боль в пальце ноги
	AQ	Боль в спине	Диагностический аспект
	QB	Диагностический аспект	Боль в спине
Иерархические	SIB	Боль в спине	Артралгия
	PAR	Боль в спине	Дорсопатия
	CHD	Боль в спине	Боль в пояснице
	RB	Боль в спине	Боль
	RN	Боль в спине	Боль в верхней части спины
Синонимичные	SY	Боль в спине	Боль: спина

Необходимо отметить, что иерархические связи PAR (parent relationships) и RB (broader relationships) – вертикальные связи дочерних терминов с родительскими – ведут к обобщающим терминам, искажающим этиопатогенетический образ. По этой причине данные типы связей не использовались для агрегации терминов. В свою очередь, связи RL («like» relationships – горизонтальные или вертикальные связи между близкородственными терминами), AQ (allowed qualifier – вертикальные связи дочерних терминов с родительскими) и QB (can be qualified by – вертикальные связи родительских терминов с дочерними) не были представлены в справочниках терминов UMLS, переведенных на русский язык.

На втором шаге отбирались термины, принадлежащие к тематическим группам, входящим в надкласс Disorders (расстройств) семантической сети UMLS. Коды и расшифровки групп терминов указанного надкласса представлены в таблице 3. Данное поисковое условие позволило исключить группы промежуточных концептов, не имеющих прямого отношения к этиопатогенезу заболеваний: таксономические и хронологические

Таблица 3. Тематические группы терминов, используемые для поиска связей с нозологиями справочника МКБ-10

№	Код	Наименование тематической группы и адаптированный перевод
1	T019	Congenital abnormality (врожденные аномалии)
2	T020	Acquired abnormality (приобретенные аномалии)
3	T033	Finding (клинические находки)
4	T037	Injury or poisoning (травмы или отравления)
5	T046	Pathologic function (патологические функции)
6	T047	Disease or syndrome (заболевания или синдромы)
7	T048	Mental or behavioral dysfunction (расстройства мышления и психики)
8	T049	Cell or molecular dysfunction (клеточные или молекулярные нарушения)
9	T050	Experimental model of disease (экспериментальные модели заболеваний)
10	T184	Sign or symptom (признаки или симптомы)
11	T190	Anatomical abnormality (анатомические нарушения)
12	T191	Neoplastic process (опухолевые процессы)

сущности, экономико-юридические и географические термины, а также узкоспециализированные понятия из биологии и химии.

На третьем шаге определялось число незамкнутых графовых путей между корневым (симптоматическим) и концевым (нозологическим) концептами, связанными друг с другом не более, чем через два промежуточных узла, относящихся хотя бы к одной тематической группе из таблицы 3. Необходимо отметить, что проведенные ранее исследования продемонстрировали возможность нахождения подавляющего большинства релевантных концептов при указанной глубине поиска [8].

В ходе третьего шага извлекались нозологические концепты (соотнесенные с кодами МКБ-10) и рассчитывались различные метрики их связности с корневым узлом. Далее по каждому клиническому профилю полученных концептов значения графовых метрик суммировались. Полученный числовой ряд сортировался по убыванию и подвергался минимаксной нормализации, где единице соответствовала максимальная близость к соответствующей медицинской области, а нулю – минимальная. Значения сохранялись в базе данных и могли использоваться для построения вектора функций принадлежности этиопатогенетического образа отдельных концептов или набора концептов.

## 5. Оценка качества построения образа нозологий

Для проведения сравнительной характеристики метрик оценки связности при построении этиопатогенетического образа полученная для концепта метка образа нозологической единицы сравнивалась с фактической (класс МКБ-10). Образ считался корректным, если метка класса

МКБ-10 присутствовала в перечне трех наиболее характерных классов в векторе этиопатогенетического образа концепта. Важно отметить, что образы были построены только для концептов, не соотнесенных с кодами МКБ-10. Для нозологий этиопатогенетические образы строились путем суммирования значений функций принадлежности, напрямую связанных с ними симптоматических концептов. Данный подход позволил интегрально оценить эффективность разметки нозологических концептов и сделать вывод о возможности ее применения при решении задачи поиска наиболее подходящих диагнозов на основании любого исходного перечня симптомов.

Для выполнения запросов с большим количеством операций объединения (в первую очередь, *cross join*), лежащих в основе расчета метрик связности узлов, применялась графовая СУБД Neo4j. При работе с указанной СУБД использовался декларативный язык запросов Cypher [9]. Для точечного поиска концептов UMLS применялась объектно-реляционная СУБД PostgreSQL, демонстрирующая наилучшие результаты при извлечении информации из хранилищ данных [10].

Для интеграции аналитических процедур в единый вычислительный алгоритм применялись библиотеки языка программирования Python: Psycopg2 (для работы с СУБД PostgreSQL), Neo4j (для работы с СУБД Neo4j), Scipy, Pandas и NumPy (для реализации вычислительных операций и формирования OLAP-срезов).

## 6. Виды графовых метрик для оценки связности узлов

Универсальным языком разметки графовых моделей знаний является Eclipse Modeling Framework (EMF) – набор инструментов и стандартов, обеспечивающих основу для взаимодействия создаваемых метамodelей с другими программными продуктами. Согласно стандарту Ecose, используемому EMF, в зависимости от системы типизации узлов графовые модели делятся на два класса. Так, модель знаний считается монотипной (*one-dimensional*), если в графе отсутствует деление узлов на типы. При существовании справочника, соотносящего узлы графа с конкретными типами (из множества представленных в модели типов), граф считается многотипным (*multidimensional*) [11]. Следует уточнить, что исходно в UMLS представлено 127 тематических групп концептов, которые правомерно могут считаться типами. Для многотипных моделей знаний существует ряд общепринятых правил, допущений и ограничений, которые необходимо учитывать при абстрактно-логическом анализе. Прежде всего, каждая семантическая сущность (формулировка термина) должна быть сопоставлена с соответствующим узлом графа. Помимо этого, каждый узел графа должен принадлежать хотя бы к одному из выделяемых

в структуре знаний типов. Наконец, между любой парой узлов графа должен существовать как минимум один путь.

Для монотипных и многотипных графовых моделей существуют различные виды метрик оценки связности узлов [12]. Однако важно отметить, что многотипная модель всегда может быть приведена к монотипной при опущении информации о выделяемых типах узлов. Поскольку UMLS полностью удовлетворяет требованиям стандарта Escore для классических многотипных графовых моделей знаний, в настоящей статье рассмотрены метрики для обоих классов.

Важным дополнительным атрибутом графовых метрик является наличие или отсутствие нормализации их вычисляемых значений. Для всех нормализованных метрик интервал допустимых значений лежит в строгих пределах от 0 до 1, где нулевое значение метрики указывает на полное отсутствие связности узла с подграфом, а единичное – на достижение максимальной связности [13]. Очевидным преимуществом нормализованных метрик является возможность проведения сравнительных оценок мощности и валидности. Помимо этого, нормализованные метрики могут быть объединены в единый экспертный ансамбль для повышения эффективности решения задач извлечения знаний из графовых моделей. Тем не менее, ненормализованные метрики также могут представлять практическую значимость, благодаря возможности их адаптации под индивидуальные требования работы с отдельными графовыми моделями знаний, а также за счет доступности их модификации путем проведения нормализации с использованием собственных аналитических подходов.

Наиболее информативными структурами UMLS являются узлы (термины), в то время как связи между ними служат в качестве вспомогательных элементов, хранящих сведения о природе взаимоотношений соответствующих понятий. В связи с этим особую значимость представляют аналитические метрики, основанные на оценке связности узлов в подграфе (таблица 4) [14].

Таблица 4. Графовые метрики для оценки связности узлов в подграфе

№	Оригинальное название	Смысловой перевод на русский язык	Тип модели	Наличие нормализации
1	Clustering coefficient	Коэффициент кластеризации	А	Да
2	Dimensional clustering coefficient	Коэффициент многотипной кластеризации	Б	Да
3	Node activity	Коэффициент типовой принадлежности	Б	Нет
4	Multiplex participation coefficient	Коэффициент взвешенной типовой принадлежности	Б	Да
Примечания: А – Монотипная графовая модель, Б – Многотипная графовая модель				

Согласно данным из таблицы 4, к числу метрик для оценки связности узлов в монотипных графовых моделях знаний относится коэффициент

кластеризации (КК). КК предусматривает нормализацию выходных значений [15] и определяется по формуле

$$(1) \quad \text{КК}_n = \frac{2 \cdot C_n}{(R_n^в + R_n^из) \cdot (R_n^в + R_n^из - 1)},$$

где  $C_n$  – количество графовых (геометрических) контуров, в образовании которых участвует узел  $n$ ;  $R_n^в$  – число рёбер (связей), входящих в узел  $n$ , а  $R_n^из$  – число рёбер (связей), исходящих из узла  $n$ .

К метрикам оценки значимости узлов в многотипных графовых моделях знаний относятся коэффициент многотипной кластеризации, коэффициент типовой принадлежности и коэффициент взвешенной типовой принадлежности.

Коэффициент многотипной кластеризации (КМК) является обобщенной версией рассмотренного ранее КК для монотипных моделей, также рассчитывается по формуле (1) и имеет три возможных вариации. Первый вариант КМК предполагает включение в расчет узлов только того типа, к которому относится корневой узел. Второй вариант данной метрики отличается от предыдущего включением в расчет узлов исключительно тех типов, к которым не относится корневой узел. Третий вариант КМК не предполагает введения ограничений на соответствие или несоответствие типов корневого узла и связанных с ним узлов.

Коэффициент типовой принадлежности (КТП) равняется числу уникальных типов узлов, с которыми связан корневой узел. Необходимо отметить, что КТП не предполагает нормализации значений и не учитывает структуру встречаемости различных типов связанных узлов.

Коэффициент взвешенной типовой принадлежности (КВТП) устраняет проблемы, связанные с применением классического КТП, и определяется по формуле

$$(2) \quad \text{КВТП}_n = \frac{T}{T-1} \cdot \left[ 1 - \sum_{d \in D} \left( \frac{R_n(d) + R_n^из(d)}{R_n(D) + R_n^из(D)} \right)^2 \right],$$

где  $T$  – число уникальных типов узлов в подграфе для узла  $n$ ;  $R_n(d)$  – число прямых связей (исходящих или входящих) между корневым узлом  $n$  и другими узлами того же типа ( $d$ );  $R_n(D)$  – число прямых связей (исходящих или входящих) между корневым узлом  $n$  и другими узлами любого типа ( $D$ ).

## 7. Эмпирический закон «ранг–частотность»

Необходимо отметить, что структура связей между концептами UMLS является неоднородной. При большом числе представленных типов узлов (тематических групп), а также многообразии типов и атрибутов связей

между терминами целесообразным является применение наиболее мощных многотипных нормализованных графовых метрик – третьего варианта КМК и КВТП. Применение этих аналитических инструментов для работы с огромными массивами данных может стать причиной длительного поиска знаний, в связи с чем необходимо заниматься параллельной разработкой более простых способов оценки связности терминов в семантической сети. К числу таких способов следует отнести подсчет числа прямых связей между исследуемым корневым узлом и другими узлами, а также подсчет геометрических контуров, в которых присутствует корневой узел.

При построении монотонно убывающей функции распределения частоты встречаемости узлов формируется классическая кривая Парето (рисунок 1) [16]. Согласно данным из рисунка, для любой точки на кривой можно определить количество концептов UMLS с числом прямых связей не менее заданного. Подобное распределение плотности структуры графовых элементов встречается при анализе языковых моделей и известно как закон Ципфа [17]. Настоящее исследование показывает, что схожая обратная закономерность характерна для графовой информационной модели UMLS.

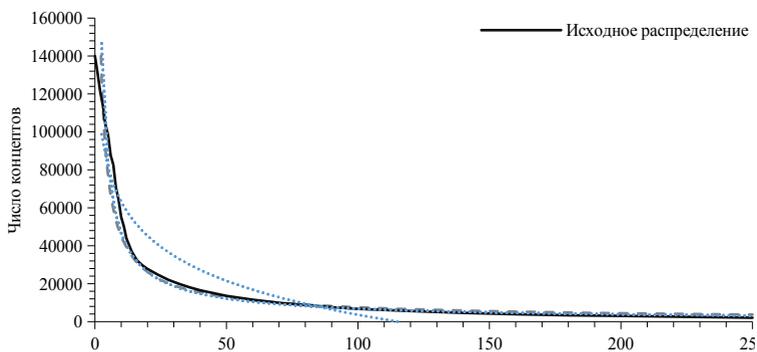


Рисунок 1. Кривая зависимости минимального числа прямых связей и числа концептов UMLS с соответствующим числом прямых связей

Математическое выражение, аппроксимирующее описанную ранее зависимость, представлено формулой

$$(3) \quad y = 2 \cdot 10^6 \cdot x^{-1.290},$$

где  $x$  – число прямых связей термина,  $y$  – количество концептов с числом прямых связей не менее  $x$ . Коэффициент детерминации ( $R^2$ ) для данного эмпирического закона составил 0.939, что свидетельствует о высоком качестве аппроксимации ( $p < 0.001$ ).

## 8. Взвешенный коэффициент кластеризации

Вышеупомянутый закон может стать фундаментальной основой для создания нормализованных метрик ранжирования концептов графовой модели UMLS. Подобные математические инструменты позволят учесть плотность прямого окружения концептов при решении задачи поиска релевантных знаний и избежать потенциального смещения оценок при ранжировании узлов графа.

С использованием формулы (3) было получено выражение для расчета взвешенного коэффициента кластеризации термина (ВКК) внутри извлекаемого подграфа UMLS

$$(4) \quad \text{ВКК}_n = \frac{C_n}{R_n^{1.290}},$$

где  $C_n$  – число геометрических контуров или незамкнутых путей, в образовании которых участвует узел  $n$ ,  $R_n$  – число прямых связей между узлом  $n$  и любыми другими вершинами графа.

Для масштабирования значений ВКК, рассчитанных по формуле (4), необходимо применить минимаксную нормализацию

$$(5) \quad \text{ВКК}'_n = \text{ВКК}_n - \frac{\min(\text{ВКК}N)}{\max(\text{ВКК}N) - \min(\text{ВКК}N)},$$

где  $\text{ВКК}'_n$  – значение ВКК для узла  $n$ ;  $\text{ВКК}N$  – набор рассчитанных значений ВКК для набора узлов  $N$ , для которого необходимо произвести ранжирование.

## 9. Варианты графовых контуров и их применение в анализе структуры UMLS

Необходимо отметить, что геометрическими контурами в теории графов принято называть любые замкнутые фигуры – структуры, в которых один и тот же узел является корневым и конечным. На рисунке 2 представлены некоторые возможные типы геометрических контуров в абстрактном подграфе. Наиболее сбалансированным и оптимизированным по вычислительным затратам является использование поиска графовых треугольников. Именно этот тип контуров рассчитывается в классических версиях КК и КМК, рассмотренных ранее. В зависимости от количества включаемых узлов в каждый уровень подграфа могут быть использованы различные виды геометрических контуров. В настоящем исследовании предлагается использовать три разновидности в зависимости от размера и числа слоев извлекаемого подграфа терминов UMLS. Так, подсчет числа внешних треугольников (у которых одна вершина обращена к корневому

концепту, а две – к концевому) целесообразно при условии

$$(6) \quad N_{\text{внутр.}} < N_{\text{внешн.}},$$

где  $N_{\text{внутр.}}$  – число концептов в слое, расположенном ближе к корневому концепту, а  $N_{\text{внешн.}}$  – число концептов на более внешнем (терминальном) слое по отношению к  $N_{\text{внутр.}}$ . Подобная ситуация характерна для подграфа, представленного на рисунке 2, поскольку в его промежуточном слое насчитывается всего 4 концепта, а в терминальном слое – 7.

При изменении знака неравенства в формуле (6) на противоположный рекомендуется использовать подсчет внутренних треугольников – контуров, у которых одна вершина обращена к терминальному слою концептов, а две – к корневому узлу). Подсчет числа графовых ромбов целесообразен при приблизительном равенстве размеров слоев при необходимости учета наличия непрямых связей между терминами. Использование варьирования различных типов графовых контуров в зависимости от размера слоев извлекаемого подграфа позволит повысить контрастность расчетов и улучшить качество ранжирования концептов.

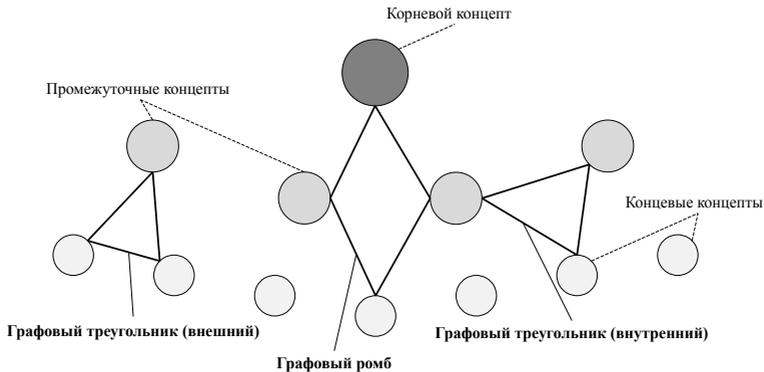


РИСУНОК 2. Варианты геометрических контуров

## 10. Результаты оценки качества разметки нозологических концептов UMLS

При использовании метрик из таблицы 4, а также эмпирически полученного взвешенного коэффициента кластеризации удалось построить этиопатогенетический образ для 134 тыс. симптоматических (93.9%) концептов русскоязычной версии UMLS (оставшиеся концепты не имели прямых и непрямых связей с нозологиями МКБ-10 при глубине поиска третьего уровня).

Далее с использованием образов симптоматических концептов осуществлена оценка качества построения образов для нозологических

концептов при использовании различных графовых метрик (таблица 5).

Таблица 5. Сравнительная характеристика результатов классификации нозологических концептов UMLS ( $n = 1577$ )

Тип графовой метрики	Тип графовых контуров		P
	гамильтоновы	негамильтоновы	
Коэффициент кластеризации (КК)	$63 \pm 1.2$	$64 \pm 1.2$	0.361
Коэффициент многотипной кластеризации (КМК)	$72 \pm 2.2$	$74 \pm 2.2$	0.297
Коэффициент типовой принадлежности (КТП)	$24 \pm 2.4$	$24 \pm 2.4$	0.965
Коэффициент взвешенной типовой принадлежности (КВТП)	$84 \pm 1.6$	$88 \pm 1.6$	0.024
Взвешенный коэффициент кластеризации (ВКК)	$86 \pm 1.7$	$91 \pm 1.4$	0.009

В ячейках таблицы представлены значения долей концептов МКБ-10, для которых эталонная метка присутствовала в перечне трех наиболее характерных классов в векторе этиопатогенетического образа ( $p \pm m$ , где  $p$  – значение доли,  $m$  – 95% доверительный интервал). Согласно данным, представленным в таблице, наилучший результат показал подход с использованием эмпирически полученного ВКК. Значение метрики качества классификации составило  $86 \pm 1.7\%$  при использовании только гамильтоновских графовых контуров (со строго не повторяющимися гранями) [18]. Использование негамильтоновских графовых контуров (с возможностью повторения одних и тех же граней в контуре дважды) позволило статистически значимо повысить качество классификации и добиться значений метрики качества в  $91 \pm 1.4\%$  ( $p = 0.009$ ).

Клиническая интерпретация этиопатогенетического образа терминов UMLS может быть продемонстрирована на следующем примере. Так, для концепта-симптома с русскоязычной формулировкой «боль в спине» при использовании алгоритма классификации, основанного на применении взвешенного коэффициента кластеризации, был получен следующий этиопатогенетический образ: [«Ортопедический профиль»: 1.000, «Неврологический профиль»: 0.921, «Кардиологический профиль»: 0.652, «Урогенитальный профиль»: 0.560]. Указанный термин имеет наиболее выраженную связь с патологиями костно-мышечной (1.000) и нервной (0.921) систем, сердечно-сосудистыми патологиями (0.652) и заболеваниями урогенитального тракта (0.560). Полученный образ согласуется с современными данными об этиопатогенезе боли в спине [19–21]. Связь с ортопедическим и неврологическим профилями может указывать на наличие неспецифических (деструктивных, дисфункциональных или дистрофических) изменений опорно-двигательного аппарата с возможностью вторичного повреждения смежных структур периферической нервной системы. Связь с заболеваниями сердечно-сосудистой системы заключается в возможности возникновения боли с иррадиацией

в спину при остром инфаркте миокарда, когда боль может отдавать в спину под левую лопатку или между лопатками. Наконец, боль в спине может возникать при некоторых заболеваниях почек и проявляться в виде симптома Пастернацкого, что позволяет считать полученный образ корректным.

### **Заключение**

Вопрос о возможности использования графовых информационных моделей в системах ППКР характеризуется низкой степенью научной проработанности [22–25]. Реализация информационной поддержки принятия клинических решений с использованием метатезауруса UMLS требует применения метрик оценки связности концептов для их ранжирования по степени значимости в контексте решаемой задачи. Извлечение релевантных знаний из UMLS может стать возможным при разработке специализированных аналитических метрик и создании системы весовых коэффициентов в структуре графовой информационной модели.

В настоящем исследовании проведена сравнительная характеристика графовых метрик при решении задачи построения этиопатогенетического образа концептов метатезауруса UMLS. Среди найденных в литературе метрик оценки связности узлов в графе допустимые результаты показали коэффициент многотипной кластеризации и коэффициент взвешенной типовой принадлежности. Доли совпадений меток алгоритма с фактическим классом заболевания из справочника МКБ-10 составили 72–74% и 84–88%, соответственно. Наилучшее качество классификации продемонстрировал авторский взвешенный коэффициент кластеризации (86–91%), основанный на использовании эмпирически полученной функции распределения плотности структуры прямых связей узлов. Определение числа негамильтоновских графовых треугольников, образованных целевым концептом, позволяет с высокой степенью точности осуществлять оценку степени принадлежности любого концепта метатезауруса UMLS к соответствующим клиническим направлениям.

Исходя из результатов многоклассовой классификации нозологий МКБ-10 по клиническим направлениям, был сделан вывод о возможности использования взвешенного коэффициента кластеризации для решения задач разметки концептов UMLS. В дальнейшем предполагается использование полученных образов симптоматических концептов при дифференциальной диагностике заболеваний на основании данных, получаемых из неструктурированного текста с использованием алгоритмов извлечения именованных сущностей. Соотнесение формулировок из текстов с клинически размеченными концептами UMLS позволит строить уникальные этиопатогенетические варианты образов по клинической картине и сопоставлять их с образами заболеваний, что сделает возможным осуществление качественной работы информационно-поисковых алгоритмов.

## Список литературы

- [1] Астанин П. А., Ронжин Л. В., Раузина С. Е., Зарубина Т. В. *Алгоритмы семантического анализа данных и возможности их применения в разработке медицинских информационных систем // Цифровая статистика. Новые задачи и траектория движения – 2022, Материалы IV Съезда медицинских статистиков Москвы (21–23 сентября 2022 года).*– 2022.– С. 6–9.  [↑60](#)
- [2] Кукарцев В. В., Колмакова З. А., Мельникова О. Л. *Системный анализ возможностей по извлечению именованных сущностей с применением технологии Text Mining // Перспективы науки.*– 2019.– Т. **120.**– № 9.– С. 18–20.  [↑60](#)
- [3] Berlingerio M., Coscia M., Giannotti F., Monreale A., Pedreschi D. *Multidimensional networks: foundations of structural analysis // World Wide Web.*– 2013.– Vol. **16.**– Pp. 567–593.  [↑60](#)
- [4] Клабукова Д. Л., Давыдовская М. В. *Внедрение международной терминологической базы MedDRA в практику фармаконадзора в Российской Федерации // Московская медицина.*– 2020.– Т. **35.**– № 1.– С. 64–69.  [↑61](#)
- [5] Кузьмин А. Г., Умаров М. Ф. *Интеграция современных медицинских информационных технологий // Вестник Вологодского государственного университета. Серия: Технические науки.*– 2021.– Т. **12.**– № 2.– С. 32–35.  [↑62](#)
- [6] Зацман И. М., Золотарев О. В., Хакимова А. Х., Дунсяо Гу *Модель и технология извлечения новых терминов из медицинских текстов // Информ. и её примен.*– 2022.– Т. **16.**– № 4.– С. 80–86.  [↑62](#)
- [7] Mougín F., Grabar N. *Auditing the multiply-related concepts within the UMLS // Journal of the American Medical Informatics Association.*– 2014.– Vol. **21.**– No. e2.– Pp. e185–e193.  [↑64](#)
- [8] Астанин П. А. *Применение автоматизированного анализа семантической сети UMLS для решения задачи поиска релевантных знаний о ревматических заболеваниях // Математическое моделирование систем и процессов, Сборник материалов Международной научно-практической конференции (г. Псков, 9–11 ноября 2022 г.), Псков: Псковский государственный университет.*– 2022.– ISBN 978-5-00200-102-6.– С. 6–12. [↑65](#)
- [9] Ямашкин С. А., Скворцов М. А., Большакова М. В., Ямашкин А. А. *Сравнительный анализ подходов к управлению базами данных для организации хранилища репозитория нейросетевых моделей // Современные наукоемкие технологии.*– 2021.– № 6-1.– С. 108–113.  [↑66](#)
- [10] Елисеева Е. А., Горячкин Б. С., Виноградова М. В., Черненко М. В. *Оценка времени выполнения поисковых запросов в NoSQL и объектно-реляционной базах данных // Динамика сложных систем – XXI век.*– 2022.– Т. **23.**– № 2.– С. 44–51.  [↑66](#)
- [11] Еремеев А. П., Мунтян Е. Р. *Разработка онтологии на основе графов с множественными и разнотипными связями // Искусственный интеллект и принятие решений.*– 2021.– № 3.– С. 3–18.  [↑66](#)
- [12] Nicosia V., Latora V. *Measuring and modeling correlations in multiplex networks // Physical Review E.*– 2015.– Vol. **92.**– id. 032805.  [↑67](#)
- [13] Battiston F., Nicosia V., Latora V. *Structural measures for multiplex networks // Physical review E.*– 2014.– Vol. **89.**– id. 032804.  [↑67](#)

- [14] Szárnyas G., Kővári Z., Salánki A., Varro D. *Towards the characterization of realistic models: evaluation of multidisciplinary graph metrics* // *Proceedings of the ACM/IEEE 19th International Conference on Model Driven Engineering Languages and Systems* (Saint-malo, France, October 2–7, 2016), New York: ACM.– 2016.– ISBN 978-1-4503-4321-3.– Pp. 87–94. doi ↑67
- [15] Nascimento M. C. V., Carvalho A. C. P. L. F. *A graph clustering algorithm based on a clustering coefficient for weighted graphs* // *Journal of the Brazilian Computer Society.*– 2011.– Vol. 17.– Pp. 19–29. doi ↑68
- [16] Пулькин И. С., Татаринцев А. В. *Достаточная статистика для параметра распределения Парето* // *Russian Technological Journal.*– 2021.– Т. 9.– № 3.– С. 88–97. doi ↑69
- [17] Синицын В. Ю., Кашпарова В. С. *Частотные свойства лексики научных текстов и законы Ципфа высших порядков* // *Вестник РГТУ. Серия: Информатика. Информационная безопасность. Математика.*– 2022.– № 4.– С. 75–91. doi ↑69
- [18] Ghaderpour E., Morris D. *Cayley graphs on nilpotent groups with cyclic commutator subgroup are hamiltonian* // *Ars Mathematica Contemporanea.*– 2011.– Vol. 7.– No. 1, Special Issue Bled'11.– Pp. 55–72. doi ↑72
- [19] Каратеев А. Е. *Хроническая боль в спине как проявление остеоартрита позвоночника: обоснование и практика применения симптоматических средств замедленного действия* // *Современная ревматология.*– 2022.– Т. 16.– № 4.– С. 88–97. doi ↑72
- [20] Higuchi H., Harada T., Hiroshige J. *Evaluation of the usefulness of costovertebral angle tenderness in patients with suspected ureteral stone* // *J. Gen. Fam. Med.*– 2023.– Vol. 24.– No. 1.– Pp. 56–58. doi ↑72
- [21] Се Л., Ду Ч., Ин Ч., Вэй Я. *Острое расслоение аорты с правосторонней болью в грудной клетке и спине, сопровождающейся левосторонней дискинезией конечности* // *Кардиология.*– 2022.– Т. 62.– № 6.– С. 74–76. doi ↑72
- [22] Дуга С. В., Труфанов А. И. *Сеть знаний как концепция систем поддержки принятия решения в предварительном следствии* // *Безопасность информационных технологий.*– 2020.– Т. 27.– № 3.– С. 54–65. doi ↑73
- [23] Мосалов О. П. *Векторные представления рёбер графа онтологии как инструмент для анализа и генерации новых данных* // *Информационно-технологический вестник.*– 2021.– Т. 27.– № 1.– С. 93–101. \* ↑73
- [24] Ананьева Е. А. *Уровни представления маршрутных пассажирских транспортных сетей в виде графовых моделей* // *Colloquium-Journal.*– 2019.– Т. 35.– № 11-1.– С. 63–68. \* ↑73
- [25] Близнякова Е. А., Куликов А. А., Куликов А. В. *Сравнительный анализ методов поиска кратчайшего пути в графе* // *Архитектура, строительство, транспорт.*– 2022.– № 1.– С. 80–87. doi ↑73

Поступила в редакцию	30.03.2023;
одобрена после рецензирования	05.05.2023;
принята к публикации	18.06.2023;
опубликована онлайн	07.10.2023.

Рекомендовал к публикации

к.т.н. Я. И. Гулиев

## Информация об авторах:



### Павел Андреевич Астанин

аспирант кафедры медицинской кибернетики и информатики имени С. А. Гаспаряна, аналитик (лаборатория семантического анализа медицинской информации РНИМУ им. Н. И. Пирогова). Область интересов: обработка естественного языка (NLP), теория графов, базы данных, нечёткая логика, информационная поддержка принятия решений. Автор более 50 научных работ

 0000-0002-1854-8686

**e-mail:** [med\\_cyber@mail.ru](mailto:med_cyber@mail.ru)



### Светлана Евгеньевна Раузина

к. м. н., доцент, зав. лаб. семантического анализа медицинской информации (РНИМУ им. Н. И. Пирогова). Область интересов: медицинская информатика, проектирование МИС, разработка систем поддержки принятия решений. Автор более 50 научных работ, в том числе 1 учебник, 2 главы в монографиях, 4 сертификата и свидетельства на программные средства, 3 методических пособия

 0000-0002-9535-2847

**e-mail:** [rauzina@mail.ru](mailto:rauzina@mail.ru)



### Татьяна Васильевна Зарубина

д. м. н., профессор, член-корр. РАН, главный внештатный специалист Минздрава России по информационным системам в здравоохранении. Автор более 270 научных работ, в том числе 2 учебника, 5 монографий, 10 сертификатов и свидетельств на программные средства, патенты, методические пособия, сборники под редакцией

 0000-0002-4403-8049

**e-mail:** [t\\_zarubina@mail.ru](mailto:t_zarubina@mail.ru)

*Все авторы сделали эквивалентный вклад в подготовку публикации.  
Авторы заявляют об отсутствии конфликта интересов.*



## Computing of umls concepts etiopathogenetic image using graph metrics

Pavel Andreevich **Astanin**<sup>1✉</sup>, Svetlana Evgen'evna **Rauzina**<sup>2</sup>,  
Tat'yana Vasil'evna **Zarubina**<sup>3</sup>

Pirogov Russian National Research Medical University, Moscow, Russia

<sup>✉</sup>*med\_cyber@mail.ru*

**Abstract.** At present, the development of clinical decision support (CDS) tools is a crucial task in medical informatics. A lot of different information searching algorithms are used in CDS systems. A fundamental step in the design of these algorithms is the creation of an etiopathogenetic image for the analysis of unstructured medical texts. In this paper, we have conducted the literary review and a comparative evaluation of analytical metrics used to compute the etiopathogenetic image of concepts within the graph model of the Unified Medical Language System (UMLS) metathesaurus. Subsequently, we developed and validated our version of a graph metric suitable for the aforementioned task implementation.

**Key words and phrases:** hospital information system, information searching algorithms, knowledge base, graph theory, UMLS

2020 *Mathematics Subject Classification:* 68T30; 92C50

**Acknowledgments:** the current study has been performed within the framework of the Federal program «Priority 2030» based on the Healthcare Digital Transformation Institute (HDTI) in the Pirogov Russian National Research Medical University.

**For citation:** Pavel A. Astanin, Svetlana E. Rauzina, Tat'yana V. Zarubina. *Computing of umls concepts etiopathogenetic image using graph metrics*. Program Systems: Theory and Applications, 2023, 14:3(58), pp. 59–94. [https://psta.psir.ru/read/psta2023\\_3\\_59-94.pdf](https://psta.psir.ru/read/psta2023_3_59-94.pdf)

## Introduction

The Unified medical language system (UMLS) metathesaurus is the biggest set of biomedical terms applicable to unstructured text analysis. The latest version of UMLS includes 4.6 million interdisciplinary concepts classified into 127 different thematic groups. Semantically close concepts are connected by relationships uniquely assigned into nine main (and two additional) types with 992 optional clarifying attributes. Almost each UMLS concept is connected with at least one concept. That is why this metathesaurus is able to be presented as oriented multigraph with about 98 million different relationships.

The approach to organizing data as graph informational models offers numerous advantages, including significant opportunities for optimizing analytical operations and the flexibility to choose various tools for data visualization. This approach enables the creation of user-friendly program modules for interpreting the structure of knowledge. However, currently, there is no unified method for the meaningful aggregation of data from graph informational models. The use of simple analytical options, such as filters on relationship types or thematic groups of concepts, does not yield clinically significant results. This challenge is attributed to the heterogeneity of the UMLS Metathesaurus structure and the lack of directed relationships between clinically relevant concepts. Therefore, it is crucial to develop ensembles of different analytical algorithms and combine them with complex systems of weight coefficients. Additionally, the utilization of optimized and validated tools and metamodels, which are unified sets of rules for the technical implementation of knowledge models, is essential.

One of the key elements of the weight system for UMLS can be values indicating the fuzzy membership degree of every concept in clinical branches (such as pulmonology, cardiology and neurology). These values are intended to demonstrate strength of etiologic and pathogenetic relationships between concepts and branches. The set of values described above forms an etiopathogenetic image, which is a vector with nonzero values indicating the degree of concept membership in clinical branches.

The creation of an etiopathogenetic image for UMLS concepts enables us to quantify their significance within a specific task context. This process ensures a reduction in the breadth of information search when

conducting graph queries. Calculating the distance between etiopathogenetic image vectors of different UMLS concepts allows us to rank symptoms and diseases based on their clinical proximity. This approach can be employed for automatically generating lists of diseases for differential diagnosis. Furthermore, it is important to compare various graph metrics and scientifically validate their usage in the aggregation of clinically relevant data from the UMLS graph model. Additionally, computing etiopathogenetic images for UMLS concepts enhances information retrieval algorithms by increasing the precision of mathematical contrasts.

*The aim of this study* is to compute and estimate etiopathogenetic image of UMLS metathesaurus concepts using metrics for graph nodes connectivity analysis.

## 1. Research phases

We have identified sequential key stages in our study. The first stage includes deploying the UMLS graph model and implementing Russian clinical thesauruses. The second stage involves creating semantic and logic rules for extracting atomic formulations. The third stage implies designing and technically implementing algorithms for calculating the membership degree function values for clinical branches of UMLS concepts. Finally, the fourth stage includes reviewing and comparing graph metrics, using nosological concepts as an example, and computing etiopathogenetic images. Additionally, during this stage, we developed a graph metric for evaluating the connectivity of UMLS concepts.

## 2. Graph model of UMLS

UMLS graph information model has been deployed using ECORE standard and all initially Russian translated concepts. They are presented in MedDRA (Medical Dictionary for Regulatory Activities Terminology), LOINC (Logical Observation Identifiers Names and Codes) and MeSH (Medical Subject Headings) terminological systems. MedDRA is a semantic tool for the majority of common clinical findings and symptoms coverage [4]. LOINC is a unified standard for different variants of medical tests describing [5]. MeSH includes vertical heading hierarchy and their synonyms for clinical articles labeling and rubricating. That is why MeSH is applicable for medical text analysis as atomic formulations thesaurus.

UMLS graph model with Russian translated concepts is significantly smaller than the original metathesaurus. In total, our graph model includes 144 thousand (3.1%) concepts and 7.9 million (8.0%) relationships between them. All concepts belong to 124 (97.6%) thematic types. The three missing concept thematic types are not clinically significant UMLS groups. They may never be applicable in describing clinical presentations or used for diagnostic tests in patients.

The matching of UMLS concepts with various clinical branches have been performed by calculating the connectivity of graph nodes. It has been based on a density analysis of relationships between non-nosological concepts and nosological concepts, which are uniquely mapped to ICD-10 (International Classification of Diseases, 10th Revision) codes in UMLS. In turn, ICD-10 has been mapped to its Russian translated version from the regulatory and reference information register of the Russian Ministry of Health.

### 3. Preparation of diseases list for data labeling

Early diagnostic tasks do not require the formulation of a medical diagnosis. Nevertheless, mapping symptoms to diagnoses requires the introduction of restrictions on the list of diseases. In the current research, a set of rules for filtering out irrelevant nosological concepts has been formulated. These concepts include atomic terms with lemmas such as “unspecified”, “classified”, “identified”, “other”, or terms that contain any diseases as substrings. For example, the code “K29” — “Gastritis and duodenitis” is redundant because there are codes in the Russian version of ICD-10, namely “K29.7” — “Gastritis” and “K29.8” — “Duodenitis”, which are substrings of it.

Then, for additional filtering of nosological concepts, UMLS nodes matched with ICD-10 and interpreted as symptoms or signs but not as unique disorders by expert way have been removed from diseases list (for example, “K92.0” — “Hematemesis”). Only concepts with the highest count of directed relationships are used for ICD-10 codes mapped with more than one UMLS concept. These procedures have resulted in a reduction in the length of the diseases list from 11232 to 1577. These measures are important for reducing potential bias in the calculation of clinical branch membership degree values. UMLS nosological concepts mapped with ICD-10 codes have been automatically associated with their clinical branches, as shown in Table 1. There were not used in concepts labeling several ICD-10 classes:

TABLE 1. Clinical branches used in computing of etiopatho-  
genetic image for UMLS concepts

№	CD-10 class (with codes diapasons)	Clinical branch name
1	V(F00-F99)	Psychiatry
2	VI(G00-G99)	Neurology
3	VII(H00-H59)	Ophthalmology
4	VIII(H60-H95)	Surdology
5	IX(I00-I99)	Cardiology
6	X(J00-J99)	Pulmonology
7	XI(K00-K93)	Gastroenterology
8	XII(L00-L99)	Dermatology
9	XIII(M00-M99)	Orthopedy
10	XIV(N00-N99)	Urology and sexology

- I certain infectious and parasitic diseases,
- II neoplasms,
- III diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism,
- IV endocrine, nutritional and metabolic diseases,
- XV pregnancy, childbirth and the puerperium,
- XVI certain conditions originating in the perinatal period,
- XVII congenital malformations, deformations and chromosomal abnormalities,
- XVIII symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified,
- XIX injury, poisoning and certain other consequences of external causes,
- XX external causes of morbidity and mortality,
- XXI factors influencing health status and contact with health services and
- XXII codes for special purposes.

#### 4. Algorithm for UMLS concepts etiopathogenetic images computing

The computation of etiopathogenetic images for UMLS symptomatic concepts has been implemented using an iterative cycle that includes

TABLE 2. Brief semantical describing of basic UMLS relationships types

Relationship class	Type	An example of relationship	
		Root concept	Leaf concept
Associative	RO	Back pain	Lumbar sciatica
	RQ	Back pain	Discomfort in the back
	RL	Hyperuricemia	Pain in the toe
	AQ	Back pain	Diagnostic aspect
	QB	Diagnostic aspect	Back pain
Hierarchical	SIB	Back pain	Arthralgia
	PAR	Back pain	Dorsopathy
	CHD	Back pain	Lumbar back pain
	RB	Back pain	Pain
	RN	Back pain	Pain in upper part of back
Synonymous	SY	Back pain	Pain: back

several mathematical steps. In the first step, we have collected all graph nodes related to the targeted concept by considering SIB (sibling relationships), RO (other relationships, — largely, vertical relationships between non-synonymous concepts), CHD (child relationships), RN (narrower relationships, — vertical relationships between parent and child concepts), RQ (related and possibly synonymous relationships), and SY (synonymous relationships). A brief description of UMLS relationship types is presented in Table 2.

It is important to note that hierarchical PAR (parent relationships) and RB (broader relationships—vertical relationships between child and parent concepts) connect clinical concepts with non-specific UMLS nodes. Use of them skews results of etiopathogenetic images computing. That is why those relationships types were not used in concepts searching. In turn, RL (*like* relationships—different horizontal and vertical relationships between semantically related concepts), AQ (allowed qualifier—technical vertical ascending relations) and QB (can be qualified by—technical vertical descending relations) did not appear in Russian translated UMLS sources.

In the second step, we have separated concepts non-related to the “Disorders” semantic group in the UMLS network. Thematic types and their names including in the “Disorders” group are described in Table 3.

TABLE 3. UMLS thematic groups used for aggregating of relationships with ICD-10 nosological concepts

№	Thematic group code (tui)	Name of thematic group
1	T019	Congenital abnormality
2	T020	Acquired abnormality
3	T033	Finding
4	T037	Injury or poisoning
5	T046	Pathologic function
6	T047	Disease or syndrome
7	T048	Mental or behavioral dysfunction
8	T049	Cell or molecular dysfunction
9	T050	Experimental model of disease
10	T184	Sign or symptom
11	T190	Anatomical abnormality
12	T191	Neoplastic process

This limitation allows us to eliminate concept groups that are not related with diseases etiology and pathogenesis. These include economical, juridical, taxonomic and chronological concepts, geographic objects and specified biochemical terms.

In the third step, we have calculated count of graph path between root (symptomatic) and leaf (nosological) concepts related by not more than 2 other nodes mapped with groups in table 3. According to the results of our early studies, this graph path length is sufficient to achieve a relevant concept searching sensitivity of 90% [8].

We have collected nosological UMLS concepts (mapped to ICD-10) and calculated values of different graph metrics. Then their values have been summarized for each clinical branch and were normalized data between 0 and 1 range by using the feature scaling (min-max normalization). Maximal value 1 corresponded to the maximum proximity to the corresponding medical area, and zero corresponded to the minimum. All calculated values have been saved in database and were used for etiopathogenetic images computing for UMLS concepts then.

## 5. Estimation of etiopathogenetic images quality

In current study, estimation of graph metrics efficiency based on nodes connectivity analysis in etiopathogenetic images computing. Received labels for nosological concepts have been compared with ICD-10 class (actual labels). According to the research condition, the computed image is correct if ICD-10 label is present in top-3 clinical labels for corresponding disease in its etiopathogenetic image vector. It is important to note that images have been computed for non-mapped with ICD-10 concepts only. For concepts mapped to ICD-10 codes we summarized values of membership function for directly related clinical nodes in UMLS graph model. It provides the integral images estimation for nosological concepts to make conclusions about their applicability in multipurpose symptom checkers development.

Queries with union operations (cross-joins) that underlie node connectivity calculations have been performed by the graph DBMS Neo4j. The declarative query language Cypher is used for Neo4j interaction [9]. However, the object-relational DBMS PostgreSQL is more effective in handling specific queries. That is why this DBMS has also been used for data aggregation from UMLS.

Analytical procedures have been integrated into a single script using Python libraries: Psycopg2 (for executing queries in PostgreSQL DBMS), Neo4j (for executing queries in Neo4j DBMS), Scipy, Pandas, and Numpy (for data processing and statistical analysis).

## 6. Metrics for connectivity analysis of graph nodes

The Eclipse Modeling Framework (EMF) is a set of tools for unifying the development and deployment of metamodels. According to the Ecore standard used by EMF, there are two types of graph models based on node typification. A knowledge graph model is one-dimensional if there is no system for node differentiation by type. Otherwise, the graph model is multidimensional [11]. To clarify, the current version of the UMLS Metathesaurus contains 127 concept types. There are a set of rules, limitations, and assumptions that need to be considered in the abstract-logical analysis of multidimensional graph knowledge models. First and foremost, each semantic entity (atomic formulation) must be associated with the most appropriate graph node. Additionally, each graph node must be mapped to at least one type. Lastly, there must be at least one path between each pair of graph nodes.

TABLE 4. Graph metrics for subgraph nodes connectivity analysis

№	Original name	Using graph model	Values normalization
1	Clustering coefficient	monodimensional	+
2	Dimensional clustering coefficient	multidimensional	+
3	Node activity	multidimensional	–
4	Multiplex participation coefficient	multidimensional	+

Different one-dimensional and multidimensional graph metrics have been described for analyzing node connectivity in previous research [12]. However, it is important to note that multidimensional graph models can be interpreted as one-dimensional if the distribution of node types is not considered. The UMLS Metathesaurus fully complies with the Ecore requirements for multidimensional graph knowledge models. That is why we have reviewed graph metrics for both model classes in the current research.

One of the most significant attributes characterizing graph metrics is value normalization. As a rule, normalized metrics involve scaling values in the range from zero (minimum node connectivity) to one (maximum node connectivity) [13]. One distinct advantage of normalized metrics is that they can be compared and validated, making them more effective than metrics without value normalization. Additionally, normalized metrics can be integrated into a single expert ensemble to improve the quality of data aggregation. However, unnormalized metrics have practical significance due to their adaptability for specific graph model analyses. They are also applicable for modification and integration into analytical algorithms.

Nodes are the most significant knowledge sources in UMLS, while relationships are addition elements that contain data about the semantic nature of concepts' connection. For this reason, there is a great interest in graph metrics for the estimation of node connectivity in graphs (Table 4) [14].

According to the Table 4, one of the nodes connectivity metrics for monodimensional graph models is clustering coefficient (CC). It provides values normalization [15] and is calculated using expression

$$(1) \quad CC_n = \frac{2 \cdot C_n}{(R_n^{\text{in}} + R_n^{\text{out}}) \cdot (R_n^{\text{in}} + R_n^{\text{out}} - 1)}$$

In the Equation 1,  $C_n$  notes the count of graph contours included the node  $n$ ;  $R_n^{\text{in}}$  is the count of relationships directed to the node  $n$ , and  $R_n^{\text{out}}$  means the count of relationships directed from node  $n$ .

Multidimensional graph metrics considered in the current study include dimensional clustering coefficient, node activity and multiplex participation coefficient.

The Dimensional Clustering Coefficient (DCC) is a generalized version of CC for one-dimensional models. It can be calculated using expression 1 in three different variants. The first variant involves using nodes whose type coincides with the root graph node. The second variant of DCC is the opposite of the first because it requires nodes whose type does not coincide with the root graph node. The third type of DCC has no limitations on the types of graph nodes.

The Node Activity (NA) is calculated as the count of unique types of nodes directly related to the root in the graph. It is important to note that NA is an unnormalized metric that does not take into account the frequency of occurrence of different node types.

The Multiplex Participation Coefficient (MPC) does not have any issues related to NA. It can be calculated using the expression

$$(2) \quad \text{MPC}_n = \frac{T}{T-1} \cdot \left[ 1 - \sum_{d \in D} \left( \frac{R_n(d) + R_n^{\text{out}}(d)}{R_n(D) + R_n^{\text{out}}(D)} \right)^2 \right].$$

In the expression 2,  $T$  notes the count of unique types which include graph node  $n$ ;  $R_n(d)$  is the count of direct relationships between root node  $n$  and nodes that type coincides ( $d$ ) with  $n$ ;  $R_n(D)$  means the count of direct relationships between root node  $n$  and any nodes regardless ( $D$ ) of their types.

## 7. Empirical law “rank-frequency”

It should be noted that the UMLS semantic network, which includes clinical concepts, is heterogeneous. Given the presence of 127 types of nodes and 992 relationship attributes in UMLS, the most suitable approach is to use multidimensional normalized graph metrics, specifically the third variant of DCC and MPC. The application of these analytical tools for big data analysis has the potential to result in lengthy knowledge aggregation. Therefore, the development of graph tools optimization ways becomes an essential task to address before implementing knowledge bases in hospital

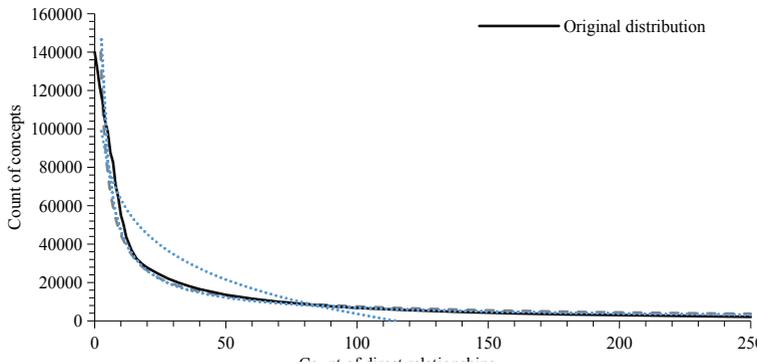


FIGURE 1. Approximation for UMLS concepts direct relationships count and count of nodes with no less neighbors

information systems. This entails the creation of simplified metrics for estimating graph node connectivity, which may involve counting direct relationships of the root node and the number of graph contours involving the root node.

We determined the number of direct relationships for each Russian-translated UMLS concept and constructed an approximation of the number of concepts that have such node neighbors in the graph model. This monotonically decreasing function follows a Pareto distribution [16]. According to Figure 1, it is possible to calculate the count of UMLS concepts with a direct relationship count no less than the corresponding value. The same distribution is characterized, as empirical Zipf’s law applies to the frequency table of words in a text or corpus of natural language [17]. In the current study, we have discovered the application of this law in UMLS graph model analysis.

According to Figure 1, the observed empirical regularity is similar to a Pareto distribution. This pattern is also observed in Zipf’s law, which is applicable to semantic models.

This mathematical pattern is statistically significant and highly determined ( $R^2 = 0.939$ ,  $p < 0.001$ ) by the following mathematical expression:

$$(3) \quad Y = 2 \cdot 10^6 \cdot X^{-1.290}$$

In the expression 3,  $X$  is the number of graph contours that the appropriate graph node is included in and  $Y$  is the actual value of the

concept significance degree. The  $R$ -Squared ( $R^2$ ) for this empirical law is 0.939, indicating the high quality of our approximation function ( $p < 0.001$ ).

## 8. Weighted clustering coefficient

The aforementioned empirical law can serve as the mathematical foundation for normalized graph metrics used in ranking UMLS concepts. These analytical tools take into account the density of the graph structure when measuring concept relevance. This is important for reducing bias in metric values when ranking graph nodes.

Using equation 3, we have created a mathematical expression for calculating the Weighted Clustering Coefficient (WCC) used in ranking concepts within the aggregated UMLS subgraph:

$$(4) \quad WCC_n = \frac{C_n}{R_n^{1.290}}$$

In the expression 4,  $C_n$  is the count of subgraph contours or subgraph paths with node  $n$ ,  $R_n$  is the count of direct relationships between node  $n$  and any graph nodes.

WCC (calculated with expression 4) values scaling is carried out with minimax normalization application:

$$(5) \quad WCC'_n = WCC_n - \frac{\min(WCC_N)}{\max(WCC_N) - \min(WCC_N)}$$

In the equation 5,  $WCC_n$  denotes the WCC value for node  $n$ ;  $WCC_N$  is the set of WCC values for ranked set of nodes ( $N$ ).

## 9. Kinds of graph contours and their usage in UMLS structure analysis

Graph theory maintains that graph contours are any closed geometric shapes, which implies that the root node is also a leaf node in these graph structures. Figure 2 illustrates various types of graph contours. Among these, the most mathematically balanced and computationally optimized are graph triangles. This type of contour is used in the CC and DCC graph metrics described earlier in this study. There are different types of contours that can be used in graph metrics calculations, depending on the number of nodes at each subgraph level. In this study, we have presented three types of graph contours and provided recommendations for their usage.

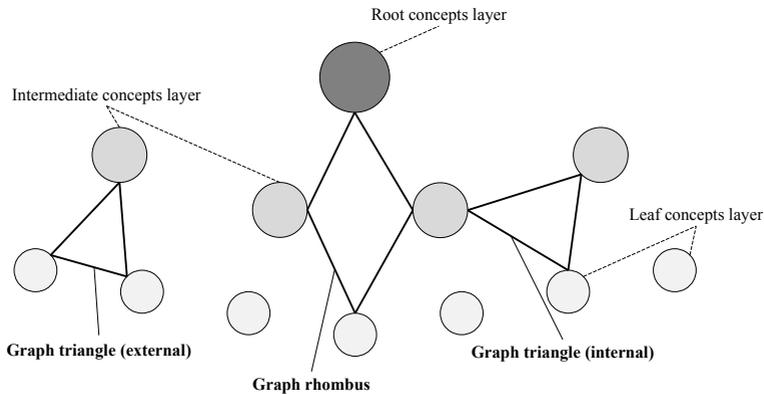


FIGURE 2. Kinds of geometrical contours

Specifically, it is advisable to aggregate external triangles (those with 1 vertex in the internal nodes layer and 2 vertices in the external layer) from the graph if the following condition is met:

$$(6) \quad N_{\text{internal}} < N_{\text{external}}$$

In the inequality 6,  $N_{\text{internal}}$  — count of nodes in internal layer (that is closer to root node),  $N_{\text{external}}$  — count of nodes in external layer (that is closer to leaf node). This condition is met for Figure 2 because in the intermediate subgraph layer, there are only four concepts, while there are seven in the terminal (external) layer.

In the opposite case, we recommend using the count of internal triangles (those with 1 vertex in the external node layer and 2 vertices in the internal layer) for graph metric calculations.

Counting the number of graph diamonds is advisable when the subgraph layer sizes are equal or when it is needed to shortest indirect paths searching between concepts. Utilizing variations of different graph contour types depending on the size of the extracted subgraph provides increasing of mathematical contrast of calculated values and ranking quality of concepts.

## 10. Quality estimation results for etiopathogenetic images of nosological UMLS concepts

Using metrics from Table 4 and empirically obtained WCC, we have computed an etiopathogenetic image for 134 thousand Russian translated symptomatic (93.9%) UMLS concepts. Other nodes are not connected with

TABLE 5. Comparison of nosological UMLS concepts classification results ( $n = 1577$ )

Type of graph metric	Types of graph contours		P
	hamiltonian	non-hamiltonian	
Clustering coefficient (CC)	$63 \pm 1.2$	$64 \pm 1.2$	0.361
Dimensional clustering coefficient (DCC)	$72 \pm 2.2$	$74 \pm 2.2$	0.297
Node activity (NA)	$24 \pm 2.4$	$24 \pm 2.4$	0.965
Multiplex participation coefficient (MPC)	$84 \pm 1.6$	$88 \pm 1.6$	0.024
Weighted clustering coefficient (WCC)	$86 \pm 1.7$	$91 \pm 1.4$	0.009

ICD-10 concepts by direct relationships and indirect paths with length less than 4.

Then we have estimated the quality of etiopathogenetic images calculated by different graph metrics usage for nosological concepts (Table 5).

There are percentages ( $p \pm CI95\%$ ) of concepts for which the basic clinical label (ICD-10 class) is present in the list of the three most inherent classes in the etiopathogenetic image vector presented in Table 5. According to Table 5, the best results have been identified using the empirically obtained WCC. The classification quality metric reached  $86 \pm 1.7\%$  when only Hamiltonian graph contours (with non-repeating edges) were used [18]. The use of non-Hamiltonian graph contours (allowing for the possibility of repeated edges in the contour) statistically significantly improved the classification quality to  $91 \pm 1.4\%$  ( $p=0.009$ ).

Clinical interpretation of the etiopathogenetic images for UMLS concepts can be demonstrated using the following example. There is a beforehand computed etiopathogenetic image vector calculated for symptomatic concept "back pain" using WCC in our database: ["Orthopedy": 1.000, "Neurology": 0.921, "Cardiology": 0.652, "Urology": 0.560]. Thus, the specified concept is most strongly related with musculoskeletal pathology (1.000), neurological disorders (0.921) and a little poor with cardiovascular pathologies (0.652) and urogenital diseases (0.560). Based on this we can conclude that the calculated image aligns with contemporary data on the back pain etiopathogenesis [19–21]. Back pain association with orthopedic and neurology may be explained by the presence of this symptom in patients with nonspecific (degenerative, dysfunctional, or dystrophic) changes

in the musculoskeletal system. Also back pain may be associated with the secondary damage to some structures of the peripheral nervous system. The association with cardiovascular diseases may be explained by back pain presence in patients with acute myocardial infarction when it can refer to the back under the left scapula or between the shoulder blades. Finally, back pain may be related to some kidney diseases, which are manifested by this symptom. That is why the etiopathogenetic image automatically calculated for this UMLS concept is quite correct.

## Conclusion

The lack of knowledge about the possibility of using graph information models in clinical decision support systems (CDSS) is an actual issue [22–25]. UMLS Metathesaurus integration into CDSS requires the application of graph metrics for concepts significance estimation in the specific task context. Clinically relevant UMLS knowledge extraction may become feasible with the development of specialized analytical metrics and automatically calculated weight coefficients system.

In the current study, we have estimated graph metrics in computing of etiopathogenetic images for UMLS concepts. The acceptable results have been taken by the dimensional coefficient and the multiplex participation coefficient described in early studies by other research teams. The percentages of relevant clinical labels for nosological concepts were 72–74% and 84–88%, respectively. The best classification quality has been demonstrated by the weighted clustering coefficient (86–91%), based on the use of an empirically obtained density distribution function of nodes direct relationships. That is why the calculation of non-Hamiltonian graph triangles included root node may be used to relate concepts to the corresponding clinical branches with high quality. Based on the results of multi-class ICD-10 concepts classification, we have concluded that the weighted clustering coefficient could be used for UMLS concepts marking and in particular etiopathogenetic images computing. In the future, it is anticipated that these images will be used for differential diagnosis of diseases with the use of named entity recognition. Text mapping with UMLS clinical concepts provides generation of disease semantic profiles applicable to their integration into CDSS intended for relevant data aggregation.

## References

- [1] P. A. Astanin, L. V. Ronzhin, S. E. Rauzina, T. V. Zarubina. “Semantic analysis algorithms for data processing and possibilities of their usage in medical information systems development”, *Cifrovaya statistika. Novye zadachi i traektoriya dvizheniya — 2022*, Materialy IV S’ezda medicinskix statistikov Moskvy (21–23 sentyabrya 2022 goda), 2022, pp. 6–9 (in Russian). ↑
- [2] V. V. Kukarcev, Z. A. Kolmakova, O. L. Mel’nikova. “System analysis of possibilities to retrieve essentials using text mining technology”, *Perspektivy nauki*, **120**:9 (2019), pp. 18–20 (in Russian). ↑
- [3] M. Berlingerio, M. Coscia, F. Giannotti, A. Monreale, Pedreschi D. . “Multidimensional networks: foundations of structural analysis”, *World Wide Web*, **16** (2013), pp. 567–593.  ↑
- [4] D. L. Klabukova, M. V. Davydovskaya. “Implementation of the MedDRA international terminology base into the pharmacovigilance practice Russia”, *Moskovskaya medicina*, **35**:1 (2020), pp. 64–69 (in Russian). ↑<sup>79</sup>
- [5] A. G. Kuz’min, M. F. Umarov. “Integration of modern medical information technologies”, *Vestnik Volgodskogo gosudarstvennogo universiteta. Seriya: Tearnicheskie nauki*, **12**:2 (2021), pp. 32–35 (in Russian). ↑<sup>79</sup>
- [6] I. M. Zaczman, O. V. Zolotarev, A. X. Xakimova, Gu Dunsyao. “Model and technology for discovering new terms in medical texts”, *Inform. i eyo primen.*, **16**:4 (2022), pp. 80–86 (in Russian).  ↑
- [7] F. Mougín, N. Grabar. “Auditing the multiply-related concepts within the UMLS”, *Journal of the American Medical Informatics Association*, **21**:e2 (2014), pp. e185–e193.  ↑
- [8] P. A. Astanin. “Application of automated analysis of the UMLS semantic network to solve the problem of searching for current knowledge about rheumatic diseases”, *Matematicheskoe modelirovanie sistem i processov*, Sbornik materialov Mezhdunarodnoj nauchno-prakticheskoy konferencii (g. Pskov, 9-11 noyabrya 2022 g.), Pskovskij gosudarstvennyj universitet, Pskov, 2022, ISBN 978-5-00200-102-6, pp. 6–12 (in Russian). ↑<sup>83</sup>
- [9] S. A. Yamashkin, M. A. Skvorczov, M. V. Bol’shakova, A. A. Yamashkin. “Comparative analysis of approaches to database management for organizing a repository of neural network models”, *Sovremennye naukoemkie tekhnologii*, 2021, no. 6-1, pp. 108–113 (in Russian).  ↑<sup>84</sup>
- [10] Eliseeva E.A. , Goryachkin B.S. , M.V. Vinogradova, M.V. Chernen’kij. “Estimating search execution time in NoSQL and object-relational databases”, *Dinamika slozhnyx sistem — XXI vek*, **23**:2 (2022), pp. 44–51 (in Russian).  ↑
- [11] A. P. Ereemeev, E. R. Muntyan. “Development of an ontology based on graphs with multiple edges of different types”, *Iskusstvennyj intellekt i prinyatie reshenij*, 2021, no. 3, pp. 3–18 (in Russian).  ↑<sup>84</sup>
- [12] V. Nicosia, V. Latora. “Measuring and modeling correlations in multiplex networks”, *Physical Review E*, **92** (2015), id. 032805.  ↑<sup>85</sup>
- [13] F. Battiston, V. Nicosia, V. Latora. “Structural measures for multiplex networks”, *Physical review E*, **89** (2014), id. 032804.  ↑<sup>85</sup>

- [14] G. Szárnyas, Z. Kóvári, A. Salánki, D. Varro. “Towards the characterization of realistic models: evaluation of multidisciplinary graph metrics”, *Proceedings of the ACM/IEEE 19th International Conference on Model Driven Engineering Languages and Systems* (Saint-malo, France, October 2–7, 2016), ACM, New York, 2016, ISBN 978-1-4503-4321-3, pp. 87–94. [doi](#) ↑85
- [15] M. C. V. Nascimento, A. C. P. L. F. Carvalho. “A graph clustering algorithm based on a clustering coefficient for weighted graphs”, *Journal of the Brazilian Computer Society*, **17** (2011), pp. 19–29. [doi](#) ↑85
- [16] I. S. Pul’kin, A. V. Tatarincev. “Sufficient statistics for the Pareto distribution parameter.”, *Russian Technological Journal*, **9**:3 (2021), pp. 88–97 (in Russian). [doi](#) ↑87
- [17] V. Yu. Sinicyn, V. S. Kashparova. “Frequency properties of the lexis of scientific texts and Zipf’s laws of higher orders”, *Vestnik RGTU. Seriya: Informatika. Informacionnaya bezopasnost’. Matematika*, 2022, no. 4, pp. 75–91 (in Russian). [doi](#) ↑87
- [18] E. Ghaderpour, D. Morris. “Cayley graphs on nilpotent groups with cyclic commutator subgroup are hamiltonian”, *Ars Mathematica Contemporanea*, **7**:1, Special Issue Bled’11 (2011), pp. 55–72. [doi](#) ↑90
- [19] A. E. Karateev. “Chronic back pain as a spinal osteoarthritis manifestation: rationale and practice of symptomatic slow acting drugs for osteoarthritis use”, *Sovremennaya revmatologiya*, **16**:4 (2022), pp. 88–97 (in Russian). [doi](#) ↑90
- [20] H. Higuchi, T. Harada, J. Hiroshige. “Evaluation of the usefulness of costovertebral angle tenderness in patients with suspected ureteral stone”, *J. Gen. Fam. Med.*, **24**:1 (2023), pp. 56–58. [doi](#) ↑90
- [21] L. Se, Ch. Du, Ch. In, Ya. Vej. “Acute aortic dissection with right-sided chest and back pain accompanied by left-sided limb dyskinesia”, *Kardiologiya*, **62**:6 (2022), pp. 74–76 (in Russian). [doi](#) ↑90
- [22] S. V. Duga, A. I. Trufanov. “The knowledge graph concept of decision support system in preliminary investigation”, *Bezopasnost’ informacionnyx texnologij*, **27**:3 (2020), pp. 54–65 (in Russian). [doi](#) ↑91
- [23] O. P. Mosalov. “Edge embedding of ontology graphs as a tool for analysis and generation of new data”, *Informacionno-tehnologicheskij vestnik*, **27**:1 (2021), pp. 93–101 (in Russian). ↑91
- [24] E. A. Anan’eva. “Levels of presentation of passenger transport route networks in the form of graph models”, *Colloquium-Journal*, **35**:11-1 (2019), pp. 63–68 (in Russian). ↑91
- [25] E. A. Bliznyakova, A. A. Kulikov, A. V. Kulikov. “Comparative analysis of methods for finding the shortest distance in a graph”, *Arxitektura, stroitel’stvo, transport*, 2022, no. 1, pp. 80–87 (in Russian). [doi](#) ↑91

Received  
approved after reviewing  
accepted for publication  
published online

30.03.2023;  
05.05.2023;  
18.06.2023;  
07.10.2023.

Recommended by

Ph.D. Ya. I. Guliev

### Information about the authors:



**Pavel Andreevich Astanin**

postgraduate student of the medical cybernetics and informatics department named after S. A. Gasparyan, data analyst of the Laboratory for semantic analysis of medical information, (HDTI, RNRMU). Research interests: natural language processing (NLP), graph theory, databases, fuzzy logic, CDSS development. An author of more than 45 scientific papers

 0000-0002-1854-8686

*e-mail:* [med\\_cyber@mail.ru](mailto:med_cyber@mail.ru)



**Svetlana Evgen'evna Rauzina**

phD, docent of the medical cybernetics and informatics department named after S. A. Gasparyan, Head of the Laboratory for semantic analysis of medical information (HDTI, RNRMU). Research interests: medical informatics, HIS designing, CDSS development. An author of more than 50 scientific papers includes 1 class-book, 2 monography chapters, 4 certificates confirming the result of intellectual activity, 3 toolkits

 0000-0002-9535-2847

*e-mail:* [rauzina@mail.ru](mailto:rauzina@mail.ru)



**Tat'yana Vasil'evna Zarubina**

phD, professor, head of the medical cybernetics and informatics department named after S. A. Gasparyan, member of RAS, head of HDTI, chief specialist of the Russian Ministry of Health on CDSS. An author of more than 270 scientific papers includes 2 classbook, 5 monographs, 10 certificates confirming the result of intellectual activity

 0000-0002-4403-8049

*e-mail:* [t\\_zarubina@mail.ru](mailto:t_zarubina@mail.ru)

*The authors declare no conflicts of interests.*

УДК 004.932.2:616-073.75

10.25209/2079-3316-2023-14-3-95-113



## Использование нейронных сетей для поиска нарушений укладки пациента на рентгенограммах органов грудной клетки

Александр Александрович **Борисов**<sup>1✉</sup>, Юрий Александрович **Васильев**<sup>2</sup>,  
Антон Вячеславович **Владимирский**<sup>3</sup>, Ольга Васильевна **Омелянская**<sup>4</sup>,  
Серафим Сергеевич **Семенов**<sup>5</sup>, Кирилл Михайлович **Арзамасов**<sup>6</sup>

<sup>1</sup> Российский национальный исследовательский медицинский университет им. Н. И. Пирогова, Москва, Россия

<sup>1-6</sup> Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы, Москва, Россия

✉ [aleksandrborisov10650@gmail.com](mailto:aleksandrborisov10650@gmail.com)

**Аннотация.** В работе представлены результаты применения трансферного обучения глубоких сверточных нейронных сетей для задачи поиска рентгенограмм органов грудной клетки с нарушениями укладки и позиционирования пациента. Оцениваемые нейросетевые архитектуры: InceptionV3, Xception, ResNet152V2, InceptionResNetV2, DenseNet201, VGG16, VGG19, MobileNetV2, NASNetLarge. Для обучения и тестирования использовались рентгенограммы грудной клетки, полученные из открытых наборов данных и Единого радиологического информационного сервиса города Москвы. Все полученные модели имели метрики диагностической точности выше 95%, при этом модели на основе архитектур ResNet152V2, DenseNet201, VGG16, MobileNetV2 имели статистически значимо лучшие метрики, чем другие модели. Наилучшие абсолютные значения метрик показала модель ResNet152V2 (AUC = 0.999, чувствительность = 0.987, специфичность = 0.988, общая валидность = 0.988, F1 мера = 0.988). Модель MobileNetV2 показала наилучшую скорость обработки одного исследования ( $67.8 \pm 5.0$  ms). Широкое использование полученных нами алгоритмов способно облегчить создание больших баз данных качественных медицинских изображений, а также оптимизировать контроль качества при выполнении рентгенографических исследований органов грудной клетки.

**Ключевые слова и фразы:** нейронные сети, глубокое обучение, контроль качества, рентген органов грудной клетки

**Благодарности:** Статья подготовлена в рамках НИОКР "Разработка платформы подготовки наборов данных лучевых диагностических исследований" (№ ЕГИСУ: 123031500003-8)

**Для цитирования:** Борисов А. А., Васильев Ю. А., Владимирский А. В., Омелянская О. В., Семенов С. С., Арзамасов К. М. *Использование нейронных сетей для поиска нарушений укладки пациента на рентгенограммах органов грудной клетки* // Программные системы: теория и приложения. 2023. Т. 14. № 3(58). С. 95–113. [https://psta.pstiras.ru/read/psta2023\\_3\\_95-113.pdf](https://psta.pstiras.ru/read/psta2023_3_95-113.pdf)

© Борисов А. А., Васильев Ю. А., Владимирский А. В., Омелянская О. В., Семенов С. С., Арзамасов К. М. 2023



## Введение

Качество выполненного рентгенологического исследования напрямую влияет на диагностику и лечение заболеваний и является определяющим показателем при постановке диагноза и выборе тактики ведения пациента [3]. Максимальная информативность исследования может быть достигнута только при соблюдении всех установленных критериев качества [1].

Одним из важнейших критериев качества для рентгенологического исследования является правильное положение пациента. Во время укладки и позиционирования пациента при проведении рентгенографии органов грудной клетки необходимо учитывать расположение верхушек легких, реберно-диафрагмальных углов и других границ легких, чтобы быть уверенным, что они будут показаны на рентгенограмме [1]. (см. рисунок 1) Отсутствие данных анатомических зон на рентгенограмме может скрыть такие присутствующие у пациента патологические состояния как плевральный выпот, пневмоторакс и гидроторакс, образования и воспаление в приграничных зонах легких [2]. Отсутствие целостных границ легких на рентгенограмме снижает диагностические возможности как врача – рентгенолога, так и программного обеспечения на основе технологии искусственного интеллекта.



(а) обрезаны часть правого легкого и реберно-диафрагмальные синусы



(б) обрезаны верхушки легких

РИСУНОК 1. Рентгенограммы органов грудной клетки с нарушением укладки

На сегодняшний день дистанционное описание рентгенологических исследований стало неотъемлемой частью диагностической практики [4]. При такой организации диагностического процесса рентгенолаборанты выполняют сканирование пациентов без возможного контроля и консультации

со стороны врачей-рентгенологов, что делает качество исследований полностью зависимым от квалификации среднего медицинского персонала [5]. Наряду с этим, количество проводимых диагностических исследований непрерывно растет, что увеличивает нагрузку на врачей-рентгенологов и рентгенолаборантов и снижает время, уделяемое оценке качества полученных изображений [6]. Автоматизация контроля качества способна сократить число исследований с нарушением критериев качества, что будет способствовать снижению нагрузки на медицинский персонал и предотвращению дополнительной лучевой нагрузки пациента [7].

Алгоритмы глубокого обучения активно применяются в задачах классификации лучевых диагностических изображений [8]. Основным препятствием для применения глубокого обучения является ограниченная доступность больших и качественных наборов данных медицинских изображений [9]. Решением проблемы доступности медицинских данных является трансферное обучение, подразумевающее использование предварительно обученных на неспецифических изображениях сверточных нейронных сетей с дообучением на специфических медицинских изображениях [10].

Целью данного исследования было сравнение эффективности применения трансферного обучения различных типов глубоких сверточных нейросетевых структур для поиска нарушений укладки и позиционирования пациента на рентгенограммах органов грудной клетки

## 1. Материалы и методы

### 1.1. Выбор архитектур

В данном исследовании для сравнения использовались нейросетевые архитектуры, доступные для скачивания со сквозной платформы машинного обучения с открытым исходным кодом *TensorFlow 2*<sup>(URL)</sup>. Все архитектуры, представленные на платформе, были предварительно обучены на открытом наборе данных аннотированных изображений ImageNet, содержащим на данный момент более 14 миллионов изображений. Также для каждой модели известны показатели Top-1 точности и Top-5 точности при тестировании на валидационном наборе данных *ImageNet*<sup>(URL)</sup>. Top-1 точность подразумевает долю правильных предсказаний модели. Top-5 точность рассчитывается аналогично, но считается, что модель сделала верное предсказание, если искомый класс находится среди 5 наиболее вероятных классов при предсказании модели. Метрика Top-5 точность для оценки эффективности модели нам не подходит, так как

мы решаем задачу бинарной классификации, поэтому при отборе мы ориентировались на показатель Топ-1 точность. Для тестирования нами были отобраны 9 нейросетевых архитектур. Критерием отбора являлось максимальное значение показателя Топ-1 точность среди представителей данного типа архитектуры по данным *TensorFlow Keras Applications*<sup>URL</sup>:

*InceptionV3*: Данная архитектура разработана в 2015 году и является разновидностью нейронных сетей Inception. Она содержит 23.9 миллиона параметров, имеет глубину 189 слоев и Топ-1 точность 77.9%. Особенностью данной архитектуры является разложение операции свертки с размером ядра  $n \times n$  на набор операций свертки с размером ядра  $1 \times n$  и  $n \times 1$ , что позволяет достичь углубления, увеличения нелинейности сети и ускорения вычислений [11].

*Xception*: Данная архитектура разработана в 2016 году и является усовершенствованным аналогом сети InceptionV3 за счет введения разделяемой по глубине свертки. Она содержит 22.9 миллиона параметров, имеет глубину 81 слой и Топ-1 точность 79% [12].

*ResNet152V2*: Архитектура ResNet разработана в 2015 году и смогла решить проблему исчезающего градиента при обучении глубоких нейронных сетей. Особенностью является эффективное соединение выходных данных разных слоев, что позволяет лучше извлекать многомерные характеристики изображения. ResNet152V2 содержит 60.4 миллиона параметров, имеет глубину 307 слоев и Топ-1 точность 78% [13].

*InceptionResNetV2*: Данная архитектура разработана в 2016 году и является усовершенствованным аналогом сети InceptionV3 с добавлением связей между выходными данными разных слоев как у ResNet. Она содержит 55.9 миллиона параметров, имеет глубину 449 слоев и Топ-1 точность 80.3% [14].

*DenseNet201*: Архитектура DenseNet является потомком архитектуры ResNet. Ее особенностью является более “плотная” связь между слоями, при которой каждый слой объединяет входные данные всех предыдущих слоев и передает выходные данные каждого слоя всем последующим слоям. DenseNet201 содержит 20.2 миллиона параметров, имеет глубину 402 слоя и Топ-1 точность 77.3% [15].

*VGG19* и *VGG16*: Архитектура VGG появилась в 2014 году. Ее особенностью являлось замена больших фильтров свертки на последовательные свертки с маленькими фильтрами (3x3), что значительно упрощало математические вычисления. У сетей VGG19 и VGG16 одинаковая заявленная Топ-1 точность – 71.3%, поэтому мы протестировали в нашей задаче обе архитектуры. Количество параметров 143.7 и 138.4, количество слоев 19 и 16 соответственно [16].

*MobileNetV2*: Архитектура MobileNet появилась в 2018 году и произвела фурор, существенно снизив количество параметров и вес нейронной сети, позволив активно использовать нейронные сети на мобильных платформах. Отличительной особенностью от MobileNetV1 является наличие свёртки 1x1 с линейной функцией активации, которая понижает число каналов. MobileNetV2 содержит 3.5 миллиона параметров, имеет глубину 105 слоев и Топ-1 точность 71.3% [17].

*NASNetLarge*: Архитектура NASNetLarge использует метод поиска с подкреплением, чтобы найти лучшие конфигурации архитектуры. Она состоит из редуцированных и нормальных слоев свертки. Слои редукции — это ячейки свертки, которые возвращают карту признаков с ее двукратным уменьшением. NASNetLarge содержит 88.9 миллиона параметров, имеет глубину 533 слоев и Топ-1 точность 82.5% [18].

## 1.2. Подготовка данных

Для дообучения выбранных архитектур использовался набор данных из 10734 рентгенограмм органов грудной клетки (РГ ОГК). Из них 5182 РГ ОГК без нарушения укладки и 5552 РГ ОГК с различными нарушениями укладки. Для тестирования моделей использовался набор данных из 2347 РГ ОГК, из них 1096 РГ ОГК без нарушения укладки и 1251 РГ ОГК с различными нарушениями укладки. Для формирования указанных наборов данных использовались отобранные вручную рентгенограммы из открытых наборов данных *CheXpert*<sup>URL</sup> [8], *Chest X-rays*<sup>URL</sup>, *PadChest*<sup>URL</sup> в количестве 6682 рентгенограмм, а также рентгенограммы выгруженные из Единого радиологического информационного сервиса Единой медицинской информационно-аналитической системы города Москвы (ЕРИС ЕМИАС) и конвертированные из формата DICOM в формат PNG в количестве 6399 рентгенограмм.

## 1.3. Условия обучения и тестирования

Для объективного сравнения различных архитектур было решено создать одинаковые условия для обучения. Были сформированы одинаковые по наполнению тестовые и валидационные наборы данных для каждой модели. Для всех архитектур были выбраны: количество эпох обучения -20, размер батча – 64 изображения, размер входного изображения 244x244 пикселя. У каждой архитектуры замораживались внутренние веса, удалялся классификационный слой и добавлялся единый для всех архитектур блок классификации:

- выравнивающий (Flatten) слой
- полносвязный слой из 512 нейронов с функцией активации ReLu
- слой исключения (Dropout), обнуляющий 20% весов
- полносвязный слой из 512 нейронов с функцией активации ReLu
- слой исключения (Dropout), обнуляющий 20% весов
- полносвязный слой из 1 нейрона с функцией активации sigmoid

Обучение и тестирование проводились на персональном компьютере: OS Windows 10, CPU: Intel Core i7-10875H (8 ядер), RAM: 16 Gb, язык программирования Python 3.8, среда: Jupyter Notebook.

Качество работы обученных моделей оценивалось на этапе обучения при помощи метрик: общая валидность (accuracy), чувствительность (recall), ПЦПР - прогностическая ценность положительного результата (precision), функции потерь основанной на бинарной кросс-энтропии. Для объективной оценки обученные модели также оценивали на тестовой отложенной (не участвующей в обучении) выборки при помощи стандартных метрик диагностической точности: чувствительность, специфичность, общая валидность, площадь под характеристической кривой с расчетом 95% доверительных интервалов методом DeLong. Порог отсека для бинаризации выбирался при помощи максимизации индекса Юдена. Данные метрики получены с использованием *открытого отечественного инструмента ROC-анализа из ГБУЗ НПКЦ ДнТ ДЗМ*<sup>[URL](#)</sup>. Помимо этого, для каждой модели рассчитывалась F1 мера и определялось время обработки 1 исследования из тестовой выборки в формате среднее арифметическое  $\pm$  стандартное отклонение.

Решение о достоверности различий между значениями метрик классификации разных моделей определялось на основании не пересечения 95% доверительных интервалов соответствующих метрик. Утверждение о том, что одна модель справилась с задачей классификации лучше, чем другая принималось, если существовали достоверные различия по нескольким или всем метрикам классификации между данными моделями.

## 2. Результаты

Все выбранные архитектуры были развернуты на персональном компьютере и обучены на подготовленном наборе данных. Ниже для каждой исследуемой архитектуры приведены метрики диагностической точности, полученные на этапе обучения и ROC-кривые, полученные на тестовом наборе данных.

- InceptionV3*: Модель прошла обучение в течение 20 эпох. Метрики, полученные на обучающей выборке: функция потерь: 0.2217, общая валидность: 0.9109, чувствительность: 0.9038, ПЦПР: 0.9224.
- Xception*: Модель прошла обучение в течение 20 эпох. Метрики, полученные на обучающей выборке: функция потерь: 0.1160, общая валидность: 0.9555, чувствительность: 0.9570, ПЦПР: 0.9570.
- ResNet152V2*: Модель прошла обучение в течение 16 эпох. После 16 эпохи пошло переобучение модели, поэтому были сохранены веса последней эпохи. Метрики, полученные на обучающей выборке: функция потерь: 0.0999, общая валидность: 0.9633, чувствительность: 0.9625, ПЦПР: 0.9664.
- InceptionResNetV2*: Модель прошла обучение в течение 20 эпох. Метрики, полученные на обучающей выборке: функция потерь: 0.1867, общая валидность: 0.9259, чувствительность: 0.9277, ПЦПР: 0.9289.
- DenseNet201*: Модель прошла обучение в течение 19 эпох. После 19 эпохи пошло переобучение модели, поэтому были сохранены веса последней эпохи. Метрики, полученные на обучающей выборке: функция потерь: 0.0671, общая валидность: 0.9733, чувствительность: 0.9775, ПЦПР: 0.9710.
- VGG16*: Модель прошла обучение в течение 13 эпох. После 13 эпохи пошло переобучение модели, поэтому были сохранены веса последней эпохи. Метрики, полученные на обучающей выборке: функция потерь: 0.0897, общая валидность: 0.9651, чувствительность: 0.9663, ПЦПР: 0.9663.
- VGG19*: Модель прошла обучение в течение 10 эпох. После 10 эпохи пошло переобучение модели, поэтому были сохранены веса последней эпохи. Метрики, полученные на обучающей выборке: функция потерь: 0.1490, общая валидность: 0.9415, чувствительность: 0.9502, ПЦПР: 0.9375.
- MobileNetV2*: Модель прошла обучение в течение 20 эпох. Метрики, полученные на обучающей выборке: функция потерь: 0.0684, общая валидность: 0.9723, чувствительность: 0.9684, ПЦПР: 0.9777.
- NASNetLarge*: Модель прошла обучение в течение 20 эпох. Метрики, полученные на обучающей выборке: функция потерь: 0.1482, общая валидность: 0.9419, чувствительность: 0.9385, ПЦПР: 0.9486.

На втором этапе было проведено тестирование ранее обученных моделей на отложенном тестовом наборе данных с целью объективной оценки эффективности решения задачи классификации. Сводные результаты тестирования на тестовом наборе данных из 2347 РГ ОГК представлены в таблице 1.

Таблица 1. Метрики диагностической точности на тестовой выборке (жирным шрифтом выделены максимальные значения каждой метрики)

Модель	AUC	F1 score	Чувствительность	Специфичность	Общая валидность	Время обработки 1 исследования
<i>InceptionV3</i>	0.992 [0.989;0.994]	0.955 [0.948;0.962]	0.949 [0.937;0.961]	0.957 [0.945;0.969]	0.953 [0.944;0.961]	73.6 ± 8.7 ms
<i>Xception</i>	0.995 [0.994;0.997]	0.969 [0.964;0.975]	0.966 [0.956;0.976]	0.97 [0.96;0.98]	0.968 [0.96;0.975]	121 ± 22 ms
<i>ResNet152V2</i>	<b>0.999</b> [0.998;1.0]	<b>0.988</b> <b>[0.985;0.992]</b>	0.987 [0.981;0.993]	<b>0.988</b> [0.982;0.995]	<b>0.988</b> [0.983;0.992]	153.2 ± 9.9 ms
<i>InceptionResNetV2</i>	0.993 [0.991;0.996]	0.967 [0.961;0.973]	0.961 [0.95;0.972]	0.97 [0.96;0.98]	0.965 [0.958;0.972]	359 ± 30 ms
<i>DenseNet201</i>	0.998 [0.997;0.979]	0.985 [0.98;0.989]	<b>0.99</b> [0.984;0.995]	0.977 [0.968;0.986]	0.984 [0.979;0.9889]	131 ± 6.6 ms
<i>VGG16</i>	0.998 [0.997;0.989]	0.981 [0.977;0.986]	0.978 [0.97;0.9886]	0.982 [0.974;0.989]	0.98 [0.974;0.9886]	90 ± 3.5 ms
<i>VGG19</i>	0.994 [0.992;0.986]	0.964 [0.957;0.97]	0.961 [0.95;0.972]	0.965 [0.954;0.976]	0.963 [0.955;0.971]	110 ± 5.1 ms
<i>MobileNetV2</i>	0.998 [0.997;0.999]	0.978 [0.973;0.982]	0.98 [0.972;0.988]	0.975 [0.966;0.985]	0.978 [0.972;0.984]	<b>67.8 ± 5.0</b> ms
<i>NASNetLarge</i>	0.992 [0.989;0.994]	0.957 [0.95;0.963]	0.945 [0.932;0.957]	0.97 [0.96;0.98]	0.957 [0.948;0.965]	371 ± 29 ms

Согласно результатам, полученным на тестовой выборке, архитектуры ResNet152V2, DenseNet201, VGG16, MobileNetV2 по всем или ряду метрик превосходят архитектуры InceptionV3, Xception, InceptionResNetV2, VGG19, NASNetLarge ( $p < 0.05$ ). Между собой метрики, полученные архитектурами ResNet152V2, DenseNet201, VGG16, MobileNetV2, статистически значимо не различаются ( $p > 0.05$ ), при этом модель на основе архитектуры ResNet152V2 показывает тенденцию к более высоким абсолютным значениям ROC AUC, F1 меры, специфичности и общей валидности, а модель на основе архитектуры DenseNet201 показывает тенденцию к более высокому абсолютному значению чувствительности. Однако, одной из важнейших характеристик для глубоких нейронных сетей является время, затраченное на шаг вывода данных. В нашем исследовании наилучшее время вывода показала модель на основе архитектуры MobileNetV2. При этом MobileNetV2 тратит времени на шаг вывода данных на 24.6% меньше чем VGG16, на 48.4% меньше чем DenseNet201 и на 55.7% меньше чем ResNet152V2.

### 3. Обсуждение

Наше исследование показало, что трансферное обучение глубоких нейросетевых архитектур способно с высокой точностью классифицировать рентгенологические изображения, в том числе решить задачу определения нарушений укладки пациента на рентгенограммах органов грудной клетки. Все полученные нами модели имеют метрики диагностической точности выше 95%, что превышает условный порог для допуска к клинической валидации в 81% [19]. Эффективность нейросетевых архитектур, получивших в нашем исследовании наилучшие показатели (ResNet152V2 и MobileNetV2), подтверждается зарубежными коллегами, использовавшими их в рамках своих исследований [20–23].

Нам удалось улучшить наши собственные результаты по определению нарушения укладки и позиционирования пациента на РГ ОГК прямой проекции, полученные ранее, [24] за счет использования, помимо открытых наборов данных, локальные данные (рентгенограммы выгруженные из Единого радиологического информационного сервиса Единой медицинской информационно-аналитической системы города Москвы). При этом мы использовали меньшее количество изображений для обучения - 10734 РГ ОГК против 13580 РГ ОГК. Это свидетельствует о том, что на точность обученного программного обеспечения (ПО) на основе технологии искусственный интеллект (ТИИ) могут сильно влиять разнообразие данных, используемых при разработке. Это могут

быть: демографические особенности исследуемого населения, настройки диагностической аппаратуры, клинический профиль учреждения в котором получены изображения, способы конвертации форматов, нормализации и эквализации изображений. Наши выводы согласуются с утверждениями, выдвинутыми Европейским Обществом Радиологов [26]. Так мы показали, что AUC, полученный в ходе нашего предыдущего исследования для модели на основе архитектуры ResNet152V2 был повышен с 0.866 до 0.999, что говорит о существенном улучшении классификационной точности модели (см. таблицу 2) [24]. Обобщающая способность (генерализуемость) алгоритма в составе ПО на основе ТИИ будет выше, чем шире вариативность медицинских устройств и организаций, используемых при формировании наборов данных.

Таблица 2. Сравнение метрик диагностической точности для модели ResNet152V2 с разными источниками обучающих данных

Модель обучена	исключительно на открытых данных	на комбинации открытых данных и данных из ЕРИС ЕМИАС
AUC	0.866[0.779;0.954]	<b>0.999[0.998;1.0]</b>
Чувствительность	0.750[0.560;0.940]	<b>0.987[0.981;0.993]</b>
Специфичность	0.852[0.763;0.941]	<b>0.988[0.982;0.995]</b>
Общая валидность	0.827[0.745;0.910]	<b>0.988[0.983;0.992]</b>
F1-мера	0.682[0.566; 0.779]	<b>0.988[0.985;0.992]</b>

Наши результаты использования глубоких сверточных нейронных сетей для определения нарушения укладки и позиционирования пациента на РГ ОГК в прямой проекции превосходят результаты алгоритма Whaley и соавт., основанного на определении геометрических характеристик на изображении. В своем исследовании они столкнулись с низкой специфичностью, полученного алгоритма, описывавшего больше дефектных исследований, чем технологи [25].

Созданные нами модели можно использовать как для автоматической разметки изображений при формировании больших наборов данных, так и в качестве средства поддержки принятия решений рентгенолаборанта, оповещающая его о наличии нарушения укладки и позиционирования на исследовании, что позволит ему с учетом допустимой дозовой нагрузки выполнить исследование повторно, до направления его врачу-рентгенологу на описание.

## Заключение

Нами проведено трансферное обучение и тестирование 9 нейросетевых архитектур для поиска нарушений укладки пациента на рентгенограммах органов грудной клетки. Все полученные нами модели имеют метрики диагностической точности выше 95%, что позволяет использовать их в клинической практике. По результатам нашего исследования можно выделить 2 наилучшие архитектуры для решения поставленной задачи: в условиях ограниченного времени на обработку исследований стоит использовать архитектуру MobileNetV2, имеющую наименьшее время на шаг вывода данных ( $67.8 \pm 5.0$  ms), в условиях неограниченного времени стоит использовать архитектуру ResNet152V2, имеющую наибольшие абсолютные значения метрик качества (AUC = 0.999, чувствительность = 0.987, специфичность = 0.988, общая валидность = 0.988, F1 мера = 0.988). Широкое использование полученных нами алгоритмов способно облегчить создание больших наборов данных качественных медицинских изображений, а также оптимизировать контроль качества при выполнении рентгенографических исследований органов грудной клетки.

## Список литературы

- [1] Lampignano J., Bontrager K. L. *Textbook of Radiographic Positioning and Related Anatomy*, 8th edition. – St. Louis (Mo.): Elsevier Mosby. – 2014. – ISBN 978-0323083881. – 848 pp. [↑](#)<sub>96</sub>
- [2] Broder J. *Imaging the chest: the chest radiograph // Diagnostic Imaging for the Emergency Physician*, chapter 5. – Elsevier. – 2011. – ISBN 978-1-4160-6113-7. – Pp. 185–296. [doi](#) [↑](#)<sub>96</sub>
- [3] Wang J., Li Zh., Pu L., Zhang K., Liu X., Zhou B. *Research and application of orthotopic DR chest radiograph quality control system based on artificial intelligence // Journal of Biomedical Engineering*. – 2020. – Vol. **37**. – No. 1. – Pp. 158–168. [doi](#) [↑](#)<sub>96</sub>
- [4] Морозов С. П., Кузьмина Е. С., Ледихова Н. В., Владимировский А. В., Трофименко И. А., Мокиенко О. А., Панина Е. В., Андрейченко А. Е., Омелянская О. В., Гомбоевский В. А., Полищук Н. С., Шулькин И. М., Решетников Р. В. *Мобилизация научно-практического потенциала службы лучевой диагностики г. Москвы в пандемию COVID-19 // Digital Diagnostics*. – 2020. – Т. **1**. – № 1. – С. 5–11. [doi](#) [↑](#)<sub>96</sub>
- [5] Морозов С. П., Ледихова Н. В., Панина Е. В., Владимировский А. В., Фомичева Е. П. *Качество работы рентгенолаборантов в условиях дистанционного взаимодействия с референс-центром лучевой диагностики с применением телемедицинских технологий // Национальное здравоохранение*. – 2021. – Т. **2**. – № 2. – С. 36–46. [doi](#) [↑](#)<sub>97</sub>
- [6] van Leeuwen K. G., de Rooij M., Schalekamp S., van Ginneken B., Rutten M. J. C. M. *How does artificial intelligence in radiology improve efficiency and health outcomes?* *Pediatr. Radiol.* – 2022. – Vol. **52**. – No. 11. – Pp. 2087–2093. [doi](#) [↑](#)<sub>97</sub>

- [7] Willis Ch. E., Nishino T. K., Wells J. R., Ai H. A., Wilson J. M., Samei E. *Automated quality control assessment of clinical chest images* // Med. Phys.– 2018.– Vol. **45**.– No. 10.– Pp. 4377–4391.  [↑97](#)
- [8] Miotto R., Wang F., Wang Sh., Jiang X., Dudley J. T. *Deep learning for healthcare: review, opportunities and challenges* // Brief Bioinform.– 2018.– Vol. **19**.– No. 6.– Pp. 1236–1246.  [↑97, 99](#)
- [9] Usman M., Zia T., Tariq A. *Analyzing transfer learning of vision transformers for interpreting chest radiography* // J. Digit. Imaging.– 2022.– Vol. **35**.– No. 6.– Pp. 1445–1462.  [↑97](#)
- [10] Kim H. E., Cosa-Linan A., Santhanam N., Jannesari M., Maros M. E., Ganslandt T. *Transfer learning for medical image classification: a literature review* // BMC Med. Imaging.– 2022.– Vol. **22**.– id. 69.– 13 pp.  [↑97](#)
- [11] Szegedy Ch., Vanhoucke V., Ioffe S., Shlens J., Wojna Z. *Rethinking the inception architecture for computer vision*, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Las Vegas, NV, USA, 27–30 June 2016).– 2016.– Pp. 2818–2826.  [↑98](#)
- [12] Chollet F. *Xception: deep learning with depthwise separable convolutions*, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Honolulu, HI, USA, 21–26 July 2017).– 2017.– Pp. 1800–1807.  [↑98](#)
- [13] He K., Zhang X., Ren S., Sun J. *Identity mappings in deep residual networks*, Computer Vision – ECCV 2016 (Amsterdam, The Netherlands, October 11–14, 2016), Lecture Notes in Computer Science.– vol. **9908**, eds. Leibe B., Matas J., Sebe N., Welling M., Cham: Springer.– 2016.– ISBN 978-3-319-46493-0.– Pp. 630–645.  [↑98](#)
- [14] Szegedy Ch., Ioffe S., Vanhoucke V., Alemi A. *Inception-v4, inception-ResNet and the impact of residual connections on learning* (San Francisco, California, 2017) // Proceedings of the AAAI Conference on Artificial Intelligence.– Vol. **31**.– No. 1, Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence AAAI 2017.– Pp. 4278–4284.  [↑98](#)
- [15] Huang G., Liu Z., Van Der Maaten L., Weinberger K. Q. *Densely connected convolutional networks* // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017.– 2017.– ISBN 978-1-5386-0457-1.– Pp. 2261–2269.  [↑98](#)
- [16] Liu S., Deng W. *Very deep convolutional neural network based image classification using small training sample size*, 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR) (Kuala Lumpur, Malaysia, 03–06 November 2015).– 2015.– Pp. 730–734.  [↑98](#)
- [17] Sandler M., Howard A., Zhu M., Zhmoginov A., Chen L.-C. *MobileNetV2: inverted residuals and linear bottlenecks*, 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (Salt Lake City, UT, USA, 18–23 June 2018).– 2018.– Pp. 4510–4520.  [↑99](#)
- [18] Zoph B., Vasudevan V., Shlens J., Le Q. V. *Learning transferable architectures for scalable image recognition* // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (Salt Lake City, UT, USA, 18–23 June 2018).– 2018.– Pp. 8697–8710.  [↑99](#)

- [19] Морозов С. П., Владимировский А. В., Кляшторный В. Г., Андрейченко А. Е., Кульберг Н. С., Гомболевский В. А. *Клинические испытания программного обеспечения на основе интеллектуальных технологий (лучевая диагностика)*, Препринт № ЦДТ-2019-1.– М.– 2019.– 34 с. [URL](#) ↑103
- [20] Kanjanasurat I., Tenghongsakul K., Purahong B., Lasakul A. *CNN-RNN Network Integration for the Diagnosis of COVID-19 Using Chest X-ray and CT Images* // *Sensors* (Basel).– 2023.– Vol. **23**.– No. 3.– id. 1356.– 12 pp. [doi](#) ↑103
- [21] Baccouche A., Garcia-Zapirain B., Elmaghraby A. S. *An integrated framework for breast mass classification and diagnosis using stacked ensemble of residual neural networks* // *Sci. Rep.*– 2022.– Vol. **12**.– No. 1.– id. 12259.– 17 pp. [doi](#) ↑103
- [22] Nguyen T., Do T. H., Pham Q. D. *A deep learning based system for Covid-19 positive cases detection using chest X-ray images* // *Proceedings of the 2022 13th International Conference on Information and Communication Technology Convergence (ICTC)* (Jeju Island, Republic of Korea, 19–21 October 2022).– Pp. 1082–1087. [doi](#) ↑103
- [23] Liu Z., Liu Y., Zhang W., Hong Y., Meng J., Wang J., Zheng Sh., Xu X. *Deep learning for prediction of hepatocellular carcinoma recurrence after resection or liver transplantation: a discovery and validation study* // *Hepatol. Int.*– 2022.– Vol. **16**.– No. 3.– Pp. 577–589. [doi](#) ↑103
- [24] Борисов А. А., Семенов С. С., Арзамасов К. М. *Использование трансферного обучения для автоматизированного поиска дефектов на рентгенограммах органов грудной клетки* // *Медицинская визуализация*.– Т. **27**.– № 1.– С. 158–169. [doi](#) ↑103, 104
- [25] Whaley J. S., Pressman B. D., Wilson J. R., Bravo L., Sehnert W. J., Foos D. H. *Investigation of the variability in the assessment of digital chest X-ray image quality* // *J. Digit. Imaging*.– 2013.– Vol. **26**.– No. 2.– Pp. 217–226. [doi](#) ↑104
- [26] European Society of Radiology (ESR) *What the radiologist should know about artificial intelligence — an ESR white paper* // *Insights Imaging*.– 2019.– Vol. **10**.– No. 1.– id. 44. [doi](#) ↑104

Поступила в редакцию	15.04.2023;
одобрена после рецензирования	05.05.2023;
принята к публикации	18.06.2023;
опубликована онлайн	07.10.2023.

Рекомендовал к публикации

к.т.н. Я. И. Гулиев

## Информация об авторах:



### Александр Александрович Борисов

Младший научный сотрудник отдела медицинской информатики, радиомики и радиогеномики НПКЦ диагностики и телемедицинских технологий ДЗМ, Москва. Аналитик ИЦТМ ФГАОУ ВО РНИМУ им. Н.И. Пирогова, Москва. Научные интересы: лучевая диагностика, компьютерное зрение в области обработки медицинских изображений, машинное обучение, нейронные сети, оценка качества диагностических исследований

 0000-0003-4036-5883

*e-mail:* [aleksandrborisov10650@gmail.com](mailto:aleksandrborisov10650@gmail.com)



### Юрий Александрович Васильев

к.м.н, директор НПКЦ диагностики и телемедицинских технологий ДЗМ. Главный внештатный специалист по лучевой и инструментальной диагностике ДЗМ, член правления Московского регионального отделения Российского общества рентгенологов и радиологов. Научные интересы: оказание медицинской помощи с применением цифровых технологий, технологии искусственного интеллекта в диагностике

 0000-0002-0208-5218

*e-mail:* [VasilevYA1@zdrav.mos.ru](mailto:VasilevYA1@zdrav.mos.ru)



### Антон Вячеславович Владимирский

д.м.н., заместитель директора по научной работе НПКЦ диагностики и телемедицинских технологий ДЗМ. Научные интересы: доказательная телемедицина, организация и оказание медицинской помощи с применением цифровых (информационных) технологий, методология оценки качества телемедицинских технологий, эффективность телемедицины в клинических дисциплинах, методология оценки качества и точности алгоритмов искусственного интеллекта, история телемедицины.

 0000-0002-2990-7736

*e-mail:* [VladimirskijAV@zdrav.mos.ru](mailto:VladimirskijAV@zdrav.mos.ru)



### Ольга Васильевна Омелянская

Руководитель по управлению подразделениями Дирекции наука НПКЦ диагностики и телемедицинских технологий ДЗМ. Научные интересы: оказание медицинской помощи с применением цифровых технологий, технологии искусственного интеллекта в диагностике

 0000-0002-0245-4431

*e-mail:* [OmelyanskayaOV@zdrav.mos.ru](mailto:OmelyanskayaOV@zdrav.mos.ru)



### Серафим Сергеевич Семенов

врач-рентгенолог, младший научный сотрудник сектора разработки систем внедрения медицинских интеллектуальных технологий НПКЦ диагностики и телемедицинских технологий ДЗМ. Научные интересы: радиомика, текстурный анализ изображений, информационные технологии в медицине, технологии искусственного интеллекта и компьютерное зрение в диагностике

 0000-0003-2585-0864

*e-mail:* [SemenovSS3@zdrav.mos.ru](mailto:SemenovSS3@zdrav.mos.ru)



### Кирилл Михайлович Арзамасов

к.м.н., руководитель отдела медицинской информатики, радиомики и радиогеномики НПКЦ диагностики и телемедицинских технологий ДЗМ. Научные интересы: системы поддержки принятия врачебных решений на базе технологии искусственного интеллекта в диагностике, системы анализа медицинских изображений в ультразвуковой и лучевой диагностике, телемедицинские диагностические системы, носимые устройства мониторинга сердечно-сосудистой системы

 0000-0001-7786-0349

*e-mail:* [ArzamasovKM@zdrav.mos.ru](mailto:ArzamasovKM@zdrav.mos.ru)

*Все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.*



## Using of neural networks to search for errors of patient's positioning on chest X-rays

Alexander Aleksandrovich **Borisov**<sup>1✉</sup>, Yuri Aleksandrovich **Vasiliev**<sup>2</sup>,  
Anton Vyacheslavovich **Vladimirsky**<sup>3</sup>, Olga Vasil'evna **Omelyanskaya**<sup>4</sup>,  
Serafim Sergeevich **Serafim**<sup>5</sup>, Kirill Mikhailovich **Arzamasov**<sup>6</sup>

<sup>1-6</sup>Scientific and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Department of Health, Moscow, Russia

<sup>1</sup>Pirogov Russian National Research Medical University, Moscow, Russia

<sup>✉</sup>[aleksandrborisov10650@gmail.com](mailto:aleksandrborisov10650@gmail.com)

**Abstract.** The paper presents the results of the application of transfer learning of deep convolutional neural networks for the task of searching for chest X-rays with errors of patient styling and positioning. Evaluated neural network architectures: InceptionV3, Xception, ResNet152V2, InceptionResnetV2, DenseNet201, VGG16, VGG19, MobileNetV2, NASNetLarge. For training and testing we used chest X-rays from open datasets and the unified radiological information service of the city of Moscow. All the models obtained had diagnostic accuracy metrics above 95%, while models based on the ResNet152V2, DenseNet201, VGG16, MobileNetV2 architectures had statistically significantly better metrics than other models. The best absolute values of metrics were shown by the ResNet152V2 model (AUC = 0.999, sensitivity = 0.987, specificity = 0.988, accuracy = 0.988, F1 score = 0.988). The MobileNetV2 model showed the best processing speed of one study ( $67.8 \pm 5.0$  ms). The widespread use of the algorithms we have obtained can facilitate the creation of large databases of high-quality medical images, as well as optimize quality control when performing chest X-ray examinations. (*In Russian*).

**Key words and phrases:** neural networks, deep learning, quality control, chest X-ray

2020 *Mathematics Subject Classification:* 68T07; 68U10, 92C50

**Acknowledgments:** The article was prepared within the framework of RD "Development of a platform for preparing data sets of radiation diagnostic studies" (EGISU No.: 123031500003-8)

For citation: Alexander A. Borisov, Yuri A. Vasiliev, Anton V. Vladimirsky, Olga V. Omelyanskaya, Serafim S. Serafim, Kirill M. Arzamasov. *Using of neural networks to search for errors of patient's positioning on chest X-rays*. Program Systems: Theory and Applications, 2023, 14:3(58), pp. 95–113. (*In Russ.*). [https://psta.psir.as.ru/read/psta2023\\_3\\_95-113.pdf](https://psta.psir.as.ru/read/psta2023_3_95-113.pdf)

## References

- [1] J. Lampignano, Bontrager K. L. . *Textbook of Radiographic Positioning and Related Anatomy*, 8th edition, Elsevier Mosby, St. Louis (Mo.), 2014, ISBN 978-0323083881, 848 pp.
- [2] J. Broder. “Imaging the chest: the chest radiograph”, *Diagnostic Imaging for the Emergency Physician*, chapter 5, Elsevier, 2011, ISBN 978-1-4160-6113-7, pp. 185–296. [doi](#)
- [3] J. Wang, Zh. Li, L. Pu, K. Zhang, X. Liu, B. Zhou. “Research and application of orthotopic DR chest radiograph quality control system based on artificial intelligence”, *Journal of Biomedical Engineering*, **37**:1 (2020), pp. 158–168. [doi](#)
- [4] S. P. Morozov, E. S. Kuz'mina, N. V. Ledixova, A. V. Vladzimirskij, I. A. Trofimenko, O. A. Mokienko, E. V. Panina, A. E. Andrejchenko, O. V. Omelyanskaya, V. A. Gombolevskij, N. S. Polishhuk, I. M. Shul'kin, R. V. Reshetnikov. “Mobilizing the academic and practical potential of diagnostic radiology during the COVID-19 pandemic in Moscow”, *Digital Diagnostics*, **1**:1 (2020), pp. 5–11 (in Russian). [doi](#)
- [5] S. P. Morozov, N. V. Ledixova, E. V. Panina, A. V. Vladzimirskij, E. P. Fomicheva. “Performance quality of x-ray technicians when they interact remotely with the reference center for diagnostic radiology using telemedicine technologies”, *Nacional'noe zdravooxranenie*, **2**:2 (2021), pp. 36–46 (in Russian). [doi](#)
- [6] K. G. van Leeuwen, de Rooij M., S. Schalekamp, B. van Ginneken, M. J. C. M. Rutten. How does artificial intelligence in radiology improve efficiency and health outcomes? *Pediatr. Radiol.*, **52**:11 (2022), pp. 2087–2093. [doi](#)
- [7] Ch. E. Willis, T. K. Nishino, J. R. Wells, H. A. Ai, J. M. Wilson, E. Samei. “Automated quality control assessment of clinical chest images”, *Med. Phys.*, **45**:10 (2018), pp. 4377–4391. [doi](#)
- [8] R. Miotto, F. Wang, Sh. Wang, X. Jiang, Dudley J. T. . “Deep learning for healthcare: review, opportunities and challenges”, *Brief Bioinform.*, **19**:6 (2018), pp. 1236–1246. [doi](#)
- [9] M. Usman, T. Zia, A. Tariq. “Analyzing transfer learning of vision transformers for interpreting chest radiography”, *J. Digit. Imaging*, **35**:6 (2022), pp. 1445–1462. [doi](#)
- [10] H. E. Kim, A. Cosa-Linan, N. Santhanam, Jannesari M. , M. E. Maros, T. Ganslandt. “Transfer learning for medical image classification: a literature review”, *BMC Med. Imaging*, **22** (2022), id. 69, 13 pp. [doi](#)
- [11] Ch. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna. “Rethinking the inception architecture for computer vision”, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Las Vegas, NV, USA, 27–30 June 2016), 2016, pp. 2818–2826. [doi](#)
- [12] F. Chollet. “Xception: deep learning with depthwise separable convolutions”, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Honolulu, HI, USA, 21–26 July 2017), 2017, pp. 1800–1807. [doi](#)

- [13] K. He, X. Zhang, S. Ren, J. Sun. “Identity mappings in deep residual networks”, *Computer Vision – ECCV 2016* (Amsterdam, The Netherlands, October 11–14, 2016), Lecture Notes in Computer Science, vol. **9908**, eds. Leibe B., Matas J., Sebe N., Welling M., Springer, Cham, 2016, ISBN 978-3-319-46493-0, pp. 630–645. [doi](#)
- [14] Ch. Szegedy, S. Ioffe, V. Vanhoucke, A. Alemi. “Inception-v4, inception-ResNet and the impact of residual connections on learning” (San Francisco, California, 2017), *Proceedings of the AAAI Conference on Artificial Intelligence*, **31**:1, Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence AAAI 2017, pp. 4278–4284. [doi](#)
- [15] G. Huang, Z. Liu, Der Maaten L. Van, K. Q. Weinberger. “Densely connected convolutional networks”, *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 21–26 July 2017, 2017, ISBN 978-1-5386-0457-1, pp. 2261–2269. [doi](#)
- [16] S. Liu, W. Deng. “Very deep convolutional neural network based image classification using small training sample size”, 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR) (Kuala Lumpur, Malaysia, 03–06 November 2015), 2015, pp. 730–734. [doi](#)
- [17] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, Chen L.-C. . “MobileNetV2: inverted residuals and linear bottlenecks”, 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (Salt Lake City, UT, USA, 18–23 June 2018), 2018, pp. 4510–4520. [doi](#)
- [18] B. Zoph, V. Vasudevan, J. Shlens, Q. V. Le. “Learning transferable architectures for scalable image recognition”, *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (Salt Lake City, UT, USA, 18–23 June 2018), 2018, pp. 8697–8710. [doi](#)
- [19] S. P. Morozov et al. *Clinical testing of software based on intelligent technologies (radiation diagnosis)*, Preprint No CzDT-2019-1, M., 2019 (in Russian), 53 pp. [URL](#)
- [20] I. Kanjanasurat, K. Tenghongsakul, B. Purahong, A. Lasakul. “CNN–RNN Network Integration for the Diagnosis of COVID-19 Using Chest X-ray and CT Images”, *Sensors (Basel)*, **23**:3 (2023), id. 1356, 12 pp. [doi](#)
- [21] A. Baccouche, B. Garcia-Zapirain, A. S. Elmaghraby. “An integrated framework for breast mass classification and diagnosis using stacked ensemble of residual neural networks”, *Sci. Rep.*, **12**:1 (2022), id. 12259, 17 pp. [doi](#)
- [22] T. Nguyen, T. H. Do, Q. D. Pham. “A deep learning based system for Covid-19 positive cases detection using chest X-ray images”, *Proceedings of the 2022 13th International Conference on Information and Communication Technology Convergence (ICTC)* (Jeju Island, Republic of Korea, 19–21 October 2022), pp. 1082–1087. [doi](#)
- [23] Z. Liu, Y. Liu, W. Zhang, Y. Hong, J. Meng, J. Wang, Sh. Zheng, X. Xu. “Deep learning for prediction of hepatocellular carcinoma recurrence after resection or liver transplantation: a discovery and validation study”, *Hepatol. Int.*, **16**:3 (2022), pp. 577–589. [doi](#)

- [24] A. A. Borisov, S. S. Semenov, K. M. Arzamasov. “Transfer Learning for automated search for defects on chest X-rays”, *Medicinskaya vizualizaciya*, **27**:1, pp. 158–169 (in Russian). [doi](#)
- [25] J. S. Whaley, B. D. Pressman, J. R. Wilson, L. Bravo, W. J. Sehnert, D. H. Foos. “Investigation of the variability in the assessment of digital chest X-ray image quality”, *J. Digit. Imaging*, **26**:2 (2013), pp. 217–226. [doi](#)
- [26] Society of Radiology (ESR) European. “What the radiologist should know about artificial intelligence — an ESR white paper”, *Insights Imaging*, **10**:1 (2019), id. 44. [doi](#)

УДК 004.724.2 + 004.272.43

 10.25209/2079-3316-2023-14-3-115-138


## Разные неблокируемые самомаршрутизируемые системные сети с прямыми каналами

 Виктор Сергеевич **Подлазов** 

Институт проблем управления имени В. А. Трапезникова РАН, Москва, Россия

 podlazov@ipu.ru

**Аннотация.** Разработаны три новых варианта неблокируемой самомаршрутизируемой сети с прямыми каналами, в которой конфликты пакетов обнаруживаются и разрешаются на входе в сеть. Заблокированные в конфликтах пакеты повторно передаются источниками с минимальными задержками. В сеть добавлены сквозные цепи обратной связи для быстрого обнаружения конфликтов пакетов. Наличие цепей обратной связи позволяет осуществлять произвольный трафик без потери пакетов. Сети разработаны в 1-, 2-, 4-, и 8-каскадном вариантах с масштабированием числа каналов от нескольких сот до многих миллионов с оценкой достигаемого быстродействия сетей на перестановочном трафике. Практическая ориентация предложенных сетей – это системные сети с передачей маршрутной информации в заголовках пакетов с использованием в каждом каскаде управляющей маршрутной информации. Предложенные сети выполнены в расширенном схемном базисе, состоящем из коммутаторов, мультиплексоров и демультиплексоров. В работе представлены характеристики построенных сетей и дана их сравнительная оценка. Новизна разработанных сетей состоит как в самом факте их построения, так и в использованной структуре сетей и их характеристиках.

**Ключевые слова и фразы:** прямые каналы, дуальный коммутатор, пакетный коммутатор, мультиплексоры и демультиплексоры, многокаскадный коммутатор, бесконфликтная маршрутизация, неблокируемый коммутатор, статическая самомаршрутизация, квазиполный оргграф, сложность и быстродействие

Для цитирования: Подлазов В.С. *Разные неблокируемые самомаршрутизируемые системные сети с прямыми каналами* // Программные системы: теория и приложения. 2023. Т. 14. № 3(58). С. 115–138. [https://psta.psiras.ru/read/psta2023\\_3\\_115-138.pdf](https://psta.psiras.ru/read/psta2023_3_115-138.pdf)

## Введение

В статье продолжена работа [1, 2] по решению классической коммутационной задачи построения неблокируемого самомаршрутизируемого распределенного коммутатора с децентрализованным управлением и прямыми каналами. Он образует системную сеть пакетной коммутации с широкой масштабируемостью для многопроцессорных вычислительных систем, в первую очередь для фотонных компьютеров [3] и компьютеров spatial data flow [4]. Эта задача до сих пор не имеет полного решения оптимального по быстродействию и схемной сложности. В работе рассматриваются три новых решения этой задачи с разной оптимизацией по сложности и быстродействию.

Прямой канал между любым источником и любым приемником, содержит только комбинационные элементы и не содержит элементов памяти, используемых обычно для буферизации конфликтных пакетов, и поэтому обеспечивает меньшее время доставки пакетов по сети. Комбинационные элементы в канале образуют  $m$ -канальные коммутаторы  $m \times m$ , демультимплексоры  $1 \times m$  и мультимплексоры  $m \times 1$ . Они осуществляют маршрутизацию пакетов на основе маршрутной информации, передаваемой в заголовке пакета вместе с данными, и делают это без тактовых задержек — на лету. Это всегда можно сделать в комбинационной схеме.

В рассматриваемых решениях используется статическая маршрутизация пакетов, при которой каждый источник порождает маршрутную информацию, задающую путь через сеть к приемнику независимо от других источников. Иначе говоря, под самомаршрутизацией понимается «маршрутизация от источника», при которой источник в заголовке пакета задает его путь через сеть, а не адрес получателя.

Термин неблокируемый коммутатор используется в литературе [5, 6] для перестановочного трафика, при котором каждый источник передает один пакет только одному приемнику. Неблокируемый коммутатор осуществляет параллельную передачу пакетов от всех источников ко всем приемникам без потери пакетов вследствие их конфликтов, и такая передача осуществляется для произвольной перестановки пакетов. В литературе используется также термин условно неблокируемый или перестраиваемый коммутатор, который осуществляет бесконфликтную передачу только при наличии заранее составленного расписания для каждой перестановки.

Возможность построения многокаскадного неблокируемого коммутатора доказана еще Клозом [6, 7]. Однако для такого коммутатора не имеется процедуры параллельной самомаршрутизации. Поэтому он

на практике не используется вследствие его бóльшей сложности, но которая может служить основой для оценки сложности других неблокируемых коммутаторов.

В перестраиваемых коммутаторах самомаршрутизация используется без составления бесконфликтных расписаний, что приводит к потере пакетов при возникновении их конфликтов на произвольных перестановках. Потерянные пакеты повторно передаются источниками со значительными задержками во времени, связанными с протокольными ожиданиями подтверждениями их доставки приемникам. При этом каждая повторная передача конфликтных пакетов сопровождается указанными задержками, что снижают быстродействие перестраиваемых коммутаторов на перестановочном трафике.

Поэтому правомочна постановка задачи о создании неблокируемого коммутатора с меньшими задержками разрешения конфликтов, чем у перестраиваемого коммутатора, и со сложностью не большей сложности неблокируемого коммутатора Клоза. Эта задача может решаться посредством более раннего обнаружения конфликтов пакетов, распараллеливания процедуры их обнаружения и уменьшения задержек повторных передач конфликтных пакетов посредством использования цепей обратной связи.

В работе вводится понятие широкой неблокируемости коммутаторов в смысле отсутствия потерь пакетов. На основе этого понятия оцениваются средние задержки передачи пакетов на перестановочном и его мажорирующем квазиперестановочном трафиках. В работе предлагаются три новых решения задачи построения неблокируемого коммутатора с разной оптимизацией по быстродействию и сложности. Первое решение обладает наибольшим быстродействием и наибольшей схемной сложностью. Второе решение построено посредством использования сетей с обменными связями, в которых номера блоков и каналов смежных каскадов обмениваются значениями. Во втором решении оцениваются средние задержки передачи конфликтных пакетов на перестановочном трафике. В третьем решении однократный прием из второго решения применяется каскадно, что приводит к уменьшению сложности коммутатора.

Предлагаемые решения основываются на использовании дуального пакетного коммутатора, дуальность которого состоит в сочетании шинного и мультиплексного способов разрешения конфликтов пакетов. При шинном способе конфликты разрешаются посредством раздвижения пакетов во времени, а при мультиплексном способе – посредством разведения их по различным каналам.

Решения  $I$  ( $I = 1, 2, 3$ ) оцениваются по критерию  $\aleph_I = S_I/T_I^*$  произведения сложности  $S_I$  на быстродействие для произвольных перестановок, задавая быстродействие как величину обратную среднему числу повторных передач  $T_I^*$ . Разные решения  $I$  и  $J$  сравниваются по отношениям  $\rho_{I,J} = \aleph_I/\aleph_J$ .

В первом разделе строится неблокируемый коммутатор на основе каскадирования коммутаторов с топологией квазиполных орграфов [8], неблокируемых на произвольных перестановках пакетов. Во втором разделе рассматриваются коммутаторы первого раздела с добавленными в него цепями обратной связи (каналами и схемами) и оцениваются их характеристики на квазиперестановочном трафике. В третьем разделе рассматривается коммутатор второго раздела в сочетании с сетями с обменными связями. Там же оцениваются его характеристики на перестановочном трафике и сравниваются с характеристиками коммутаторов второго раздела. В четвертом разделе рассматривается коммутатор с каскадным применением сетей с обменными связями, оцениваются его характеристики на квазиперестановочном трафике и сравниваются с характеристиками коммутаторов второго раздела. В заключении суммируются полученные решения.

## 1. Неблокируемые сети наибольшего быстродействия

В данном разделе строятся неблокируемые самомаршрутизируемые распределенные коммутаторы на базе коммутаторов с топологией квазиполных орграфов [8], не использующие повторных передач пакетов, т.е. обладающие наибольшим быстродействием. Для них возможна любая степень масштабирования числа каналов посредством вложенного каскадирования.

Начнем с 2-мерного гиперкуба, построенного из  $m$ -канальных коммутаторов  $m \times m$ . На рисунке 1 приводится пример двумерного гиперкуба как графа и орграфа.

В орграфе (рисунок 1б) каждый узел содержит  $m$ -портового абонента (круг) и  $m$ -канального коммутатора  $m \times m$  (квадрат), которые связаны  $m$  входными и  $m$  выходными дугами. В таком представлении 2-мерный  $m$ -ичный гиперкуб может быть представлен как двудольный орграф с  $N = m^2$  узлами в каждой доле (рисунок 2), который именуется квазиполным орграфом. Соединения в нем задаются разными таблицами инцидентности для дуг от абонентов и дуг к абонентам (таблицы 1 и 2).

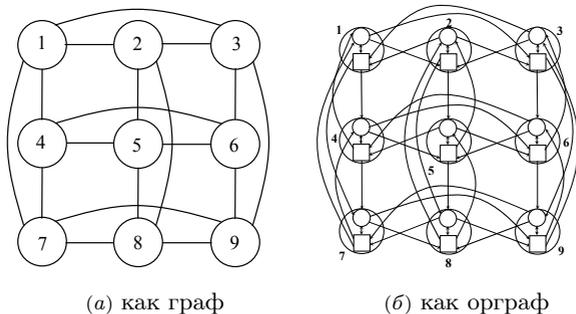


РИСУНОК 1. 2-мерный  $m$ -ичный гиперкуб при  $m = 3$

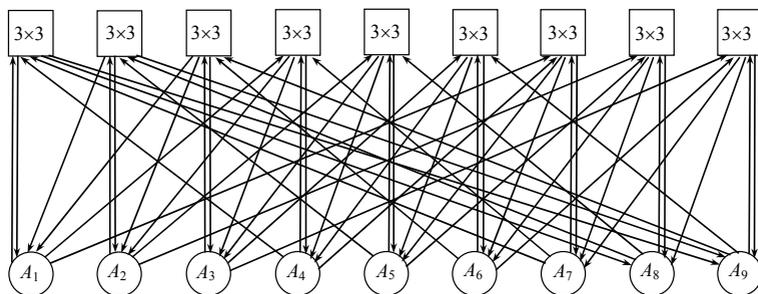


РИСУНОК 2. Квазиполный орграф при  $m = 3$

ТАБЛИЦА 1. Таблица инцидентности для квазиполного орграфа при  $m = 3$

Коммутаторы	Дуги от абонентов			Дуги к абонентам		
	1	2	3	1	4	7
1	1	2	3	2	5	8
2	1	2	3	3	6	9
3	4	5	6	1	4	7
4	4	5	6	2	5	8
5	4	5	6	3	6	9
6	7	8	9	1	4	7
7	7	8	9	2	5	8
8	7	8	9	3	6	9
9	7	8	9	3	6	9

Таблица 2. Таблица инцидентности для квазиполного орграфа при произвольном  $m$ 

Коммутаторы	Дуги от абонентов				Дуги к абонентам			
	1	2	...	$m$	1	$1 + m$	...	$1 + m(m - 1)$
1	1	2	...	$m$	1	$1 + m$	...	$1 + m(m - 1)$
2	1	2	...	$m$	2	$2 + m$	...	$2 + m(m - 1)$
...	1	2	...	$m$	...		...	...
$m$	1	2	...	$m$	$m$	$m + m$	...	$m + m(m - 1)$
$m + 1$	$m + 1$	$m + 2$	...	$m + m$	1	$1 + m$	...	$1 + m(m - 1)$
$m + 2$	$m + 1$	$m + 2$	...	$m + m$	2	$2 + m$	...	$2 + m(m - 1)$
...	$m + 1$	$m + 2$	...	$m + m$	...	...	...	...
$m + m$	$m + 1$	$m + 2$	...	$m + m$	$m$	$m + m$	...	$m + m(m - 1)$
...	...	...	...	...	...	...	...	...
$m^2 - m + 1$	$m^2 - m + 1$	$m^2 - m + 2$	...	$m^2$	1	$1 + m$	...	$1 + m(m - 1)$
...	$m^2 - m + 1$	$m^2 - m + 2$	...	$m^2$	2	$2 + m$	...	$2 + m(m - 1)$
$m^2 - 1$	$m^2 - m + 1$	$m^2 - m + 2$	...	$m^2$	...		...	...
$m^2$	$m^2 - m + 1$	$m^2 - m + 2$	...	$m^2$	$m$	$m + m$	...	$m + m(m - 1)$

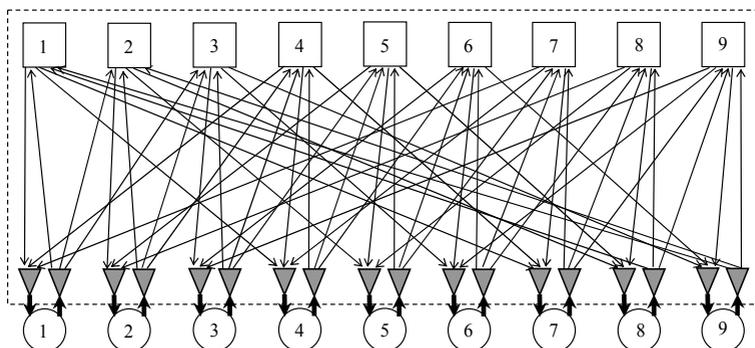


Рисунок 3. Коммутатор с топологией квазиполного орграфа при  $m = 3$

Предполагается, что  $m$  входных и выходных портов каждого 1-портового абонента создается посредством их подсоединения через демультиплексоры  $1 \times m$  и мультиплексоры  $m \times 1$  соответственно (рисунок 3).

Сеть в виде квазиполного орграфа является неблокируемым самомаршрутизируемым распределенным коммутатором. В нем маршрутизации осуществляется от источника по адресам (номерам) выходных портов демультиплексоров и коммутаторов. Бесконфликтная передача пакетов на произвольной их перестановке осуществляется по прямым каналам за один скачок от источников к приемникам, т.е с максимальным быстродействием.

Обозначим  $N_0 = m$  число каналов коммутатора  $m \times m$  и оценим его сложность как  $S_0 = 2m^2 * 2N_0^2$  точек коммутации<sup>1</sup>. Тогда коммутатор  $K_1$  с топологией квазиполного орграфа имеет  $N_1 = N_0^2 = m^2$  каналов и очевидно его сложность задается формулой (1)

$$(1) \quad S_1 = N_1(S_0 + 2m) = 2(m^4 + m^3) = 2(N_1^2 + N_1^{3/2}).$$

Можно построить коммутатор  $K_2$  с топологией квазиполного орграфа на основе коммутатора  $K_1$  и мультиплексоров  $M_1 \times 1$  и демультиплексоров  $1 \times M_1$  с  $M_1 = N_1 = m^2$ , которые имеют сложность  $s_1 = M_1$  точек коммутации. Он имеет  $N_2 = N_1^2 = m^4$  каналов, сложность  $S_2 = N_2(S_1 + 2s_1)$  и по построению сохраняет неблокируемость и самомаршрутизируемость коммутатора  $K_1$ .

Если демультиплексоры  $1 \times M_1$  и мультиплексоры  $M_1 \times 1$  строить как деревья из демультиплексоров  $1 \times m$  и мультиплексоров  $m \times 1$ , то их

<sup>1</sup>Забегая вперед, введен множитель 2 как фактор сбора коммутатора из двух каскадов демультиплексоров мультиплексоров (рисунок 4).

сложность составит  $s_1 = m^2 + m$  точек коммутации, и поэтому сложность коммутатора  $K_2$  задается формулой

$$(2) \quad S_2 = 2(m^8 + m^6 + m^6 + m^5) = 2 \left( N_2^2 + 2N_2^{3/2} + N_2^{5/8} \right).$$

Также, можно построить коммутатор  $K_4$  с топологией квазиполного орграфа на основе коммутатора  $K_2$  и мультиплекторов  $M_2 \times 1$  и демультиплекторов  $1 \times M_2$  с  $M_2 = N_2 = m^4$  со сложностью  $s_2 = M_2$  точек коммутации. Он имеет  $N_4 = N_2^2 = m^8$  каналов, сложность  $S_4 = N_4(S_2 + 2s_2)$  и сохраняет неблокируемость и самомаршрутизируемость коммутатора  $K_2$ .

Если демультиплекторы  $1 \times M_2$  и мультиплекторы  $M_2 \times 1$  строить в виде деревьев из демультиплекторов  $1 \times m$  и мультиплекторов  $m \times 1$ , то их сложность составит  $s_2 = m^2 s_1 + s_1 = m^4 + m^3 + m^2 + m$  точек коммутации, и по построению сложность коммутатора  $K_4$  задается формулой

$$(3) \quad \begin{aligned} S_4 &= N_4(S_2 + 2s_2) = 2(m^{16} + 2m^{12} + m^{12} + m^{11} + m^{10} + m^9) = \\ &= 2 \left( N_4^2 + 3N_4^{3/2} + N_4^{1/18} + N_4^{5/4} + N_4^{9/8} \right). \end{aligned}$$

Наконец, можно построить коммутатор  $K_8$  с топологией квазиполного графа на основе коммутатора  $K_4$  и мультиплекторов  $M_4 \times 1$  и демультиплекторов  $1 \times M_4$  с  $M_4 = N_8 = m^8$  со сложностью  $s_2 = M_2$  Он имеет  $N_8 = N_4^2 = m^{16}$  каналов, сложность  $S_8 = N_8(S_4 + 2s_4)$  точек коммутации и сохраняет неблокируемость и самомаршрутизируемость коммутатора  $K_4$ .

Если демультиплекторы  $1 \times M_4$  и мультиплекторы  $M_4 \times 1$  строить в виде деревьев из демультиплекторов  $1 \times m$  и мультиплекторов  $m \times 1$ , то их сложность составит  $s_4 = m^4 s_2 + s_2 = m^8 + m^7 + m^6 + m^5 + m^4 + m^3 + m^2 + m$  точек коммутации. Тогда сложность коммутатора  $K_8$  по построению задается формулой

$$(4) \quad \begin{aligned} S_8 &= N_8(S_4 + 2s_4) = \\ &= 2m^{16}(m^{16} + 3m^{12} + m^{11} + m^{10} + m^9 + m^8 + m^7 + m^6 + m^5 \\ &\quad + m^4 + m^3 + m^2 + m) = \\ &= 2 \left( N_8^2 + 3N_8^{7/4} + N_8^{27/16} + N_8^{26/16} + N_8^{25/16} + N_8^{3/2} + N_8^{23/16} + N_8^{11/8} + N_8^{21/16} + \right. \\ &\quad \left. + N_8^{5/14} + N_8^{19/16} + N_8^{9/8} + N_8^{17/16} \right) = \\ &= 2 \left( N_8^2 + 3N_4^{7/4} + \frac{N_8^{7/4} - N_8^{17/16}}{m - 1} \right). \end{aligned}$$

В таблице 3 представлены числа каналов и коммутационные сложности неблокируемых самомаршрутизируемых коммутаторов  $K_0$ – $K_8$ .

Таблица 3. Параметры наиболее быстрых неблокируемых самомаршрутизируемых коммутаторов

m	$N_0$	$S_0$	$N_1$	$S_1$	$N_2$	$S_2$	$N_4$	$S_4$	$N_8$	$S_8$
2	2	$N_0^{3,00}$	4	$N_1^{2,79}$	16	$N_2^{2,48}$	256	$N_4^{2,25}$	65 536	$N_8^{2,12}$
3	3	$N_0^{2,63}$	9	$N_1^{2,45}$	81	$N_2^{2,25}$	6 561	$N_4^{2,12}$	43 046 721	$N_8^{2,06}$
4	4	$N_0^{2,50}$	16	$N_1^{2,33}$	256	$N_2^{2,18}$	65 536	$N_4^{2,09}$	$4,29E + 09$	$N_8^{2,04}$
5	5	$N_0^{2,43}$	25	$N_1^{2,27}$	625	$N_2^{2,14}$	390 625	$N_4^{2,07}$	$1,53E + 11$	$N_8^{2,04}$
6	6	$N_0^{2,39}$	36	$N_1^{2,24}$	1 296	$N_2^{2,12}$	1 679 616	$N_4^{2,06}$	$2,82E + 12$	$N_8^{2,03}$
7	7	$N_0^{2,36}$	49	$N_1^{2,21}$	2 401	$N_2^{2,11}$	5 764 801	$N_4^{2,05}$	$3,32E + 13$	$N_8^{2,03}$
8	8	$N_0^{2,33}$	64	$N_1^{2,19}$	4 096	$N_2^{2,10}$	16 777 216	$N_4^{2,05}$	$2,81E + 14$	$N_8^{2,02}$
9	9	$N_0^{2,32}$	81	$N_1^{2,18}$	6 561	$N_2^{2,09}$	43 046 721	$N_4^{2,05}$	$1,85E + 15$	$N_8^{2,02}$
10	10	$N_0^{2,30}$	100	$N_1^{2,17}$	10 000	$N_2^{2,09}$	$1E + 08$	$N_4^{2,04}$	$1E + 16$	$N_8^{2,02}$
11	11	$N_0^{2,29}$	121	$N_1^{2,16}$	14 641	$N_2^{2,08}$	$2,14E + 08$	$N_4^{2,04}$	$4,59E + 16$	$N_8^{2,02}$
12	12	$N_0^{2,28}$	144	$N_1^{2,16}$	20 736	$N_2^{2,08}$	$4,3E + 08$	$N_4^{2,04}$	$1,85E + 17$	$N_8^{2,02}$
13	13	$N_0^{2,27}$	169	$N_1^{2,15}$	28 561	$N_2^{2,08}$	$8,16E + 08$	$N_4^{2,04}$	$6,65E + 17$	$N_8^{2,02}$
14	14	$N_0^{2,26}$	196	$N_1^{2,14}$	38 416	$N_2^{2,07}$	$1,48E + 09$	$N_4^{2,04}$	$2,18E + 18$	$N_8^{2,02}$
15	15	$N_0^{2,26}$	225	$N_1^{2,14}$	50 625	$N_2^{2,07}$	$2,56E + 09$	$N_4^{2,04}$	$6,57E + 18$	$N_8^{2,02}$
16	16	$N_0^{2,25}$	256	$N_1^{2,14}$	65 536	$N_2^{2,07}$	$4,29E + 09$	$N_4^{2,03}$	$1,84E + 19$	$N_8^{2,02}$

Видно, что сложность быстрых коммутаторов оказывается больше квадратичной, что делает их мало приемлемым для практических применений, кроме как для внутрикристалльных реализаций, например, как в коммутаторе YARK [9], имеющего структуру коммутатора  $K_1$  при  $m = 8$ .

## 2. Неблокируемые сети на квазиперестановочном графике

В данном разделе в коммутаторы вводятся цепи обратной связи (каналы и схемы), чтобы сделать их неблокируемыми в широком смысле (без потерь пакетов в результате конфликтов). Раздел строится на основе  $m$ -канального дуального пакетного коммутатора с цепями обратной связи. Впервые он был использован в работе [1] для понижения сложности неблокируемого коммутатора посредством понижения его быстродействия (увеличения задержек из-за повторных передач конфликтных пакетов). Общий вид такого коммутатора представлен рисунок 4 для  $m = 4$  в двух вариантах – в исходном и общем видах. Исходный вариант – это дуальный пакетный коммутатор (ДПК $m$ )  $m \times m$ , составленный из каскадов по  $t$  входных демультиплекторов  $1 \times t$  (Д $t$ ) и мультиплекторов  $t \times 1$  (М $t$ ), соединенных обменными связями. В общем варианте (ОПК $m$ ) к ДПК $m$  добавлены цепи (каналы и схемы) обратной связи (ОС), обозначенные пунктиром. В ОПК $m$  демультиплекторы Д $t$  из ДПК $m$  заменяются на пары Д $t$  и М $t$  ОС, обозначенные как ДМ $t$ . Аналогично, мультиплекторы М $t$  из ДПК $m$  заменяются на пары М $t$  и Д $t$  ОС, обозначенные как МД $t$ .

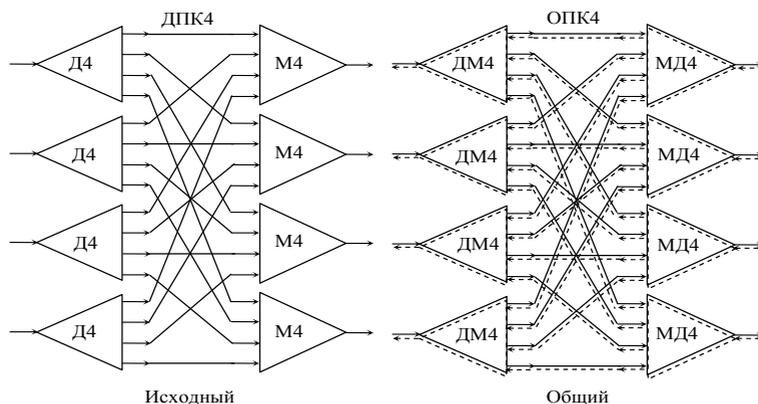


РИСУНОК 4. Дуальные пакетные коммутаторы при  $m = 4$

По линиям и схемам ОС в ОПК $m$  передаются сигналы «свободно» или «занято». В ОПК $m$  параллельно каждому выходному М $t$  включается входной Д $t$ ОС, и они образуют пару МД $t$ . В этой паре

демультиплексор  $DmOS$  первоначально передает сигнал «свободно». Аналогично, параллельно входному  $Dm$  включается выходной  $MmOS$ , и они образуют пару  $DMm$ . В этой паре мультиплексор  $MmOS$  передает сигнал «свободно» в том и только в том случае, если он имеется на всех его входах. В противном случае он распространяет сигнал «занято» с любого входа на свой выход.

Каждый источник при наличии сигнала «свободно» от своей пары  $DMm$  может вести передачу пакета данных.

При возникновении конфликта в выходном  $Mm$  один пакет пропускается на выход, а остальные каналы отключаются от выхода. Одновременно этот  $Mm$  порождает передачу сигнала «занято» в  $DmOS$  своей пары  $MDm$  по каналам ОС, одноименных конфликтным каналам. При этом сигнал «занято», который поступает на вход входного  $DmOS$  извне, распространяется на все его выходы. Каждый получивший сигнал «занято» источник прекращает передачу пакета. Каждый выходной  $Mm$  порождает передачу сигнала «свободно» по всем каналам входного  $DmOS$  сразу после того, как передаваемый пакет полностью пройдет на выход  $Mm$ . По сигналу «свободно» каждый источник сразу повторно начинают передачу конфликтных пакетов с минимальными задержками. В результате любой конфликт разрешается максимум за время  $T_0$  последовательной передачи  $m$  пакетов, т.е. при  $T_0 \leq m$ .

По построению схемная сложность ОПК $m$  оценивается как  $S_0 = 4m^2 = 4N_0^2$  точек коммутации, и он является неблокируемым коммутатором на перестановочном трафике, в котором цепи ОС не используются, т.е.  $T_0 = 1$ .

Будем называть перестановочный трафик квазиперестановочным, если его образуют пакеты перестановки, но адресованные с равной вероятностью на любой адрес. В этом трафике конфликтные пакеты, адресованные разным приемникам, между собой не конфликтуют. Конфликты таких пакетов разрешаются посредством их повторных передач, осуществляемых параллельно. В результате все конфликты разрешаются за время разрешения конфликта наибольшей кратности.

Для квазиперестановочного трафика можно оценить среднее значение  $T_0$ , т.е. значение  $T_0^*$  на множестве произвольных квазиперестановок. Делается это следующим образом. Порождается случайная квазиперестановка между входами и выходами СПК $_1M_1$ . В ней определяется распределение конфликтов по выходам и среди них выделяется конфликт наибольшей кратности  $\mathfrak{Z}$ , который разрешается за время последовательной передачи  $\mathfrak{Z}$  пакетов. Это время и задает случайную задержку передачи пакетов в данной квазиперестановке. Затем осуществляется усреднение этой величины на множестве произвольных квазиперестановок. Полученное среднее значение и задает задержку  $T_0^*$  (таблица 4). Это было сделано посредством численного моделирования для малых  $N_0 = m$ .

На основе ОПК $m$  можно построить однокаскадный общий коммутатор с топологией квазиполного орграфа – ОПК $_1N_1$  (рисунок 5 при  $m = 2$  и  $N_1 = 4$ ).

ОПК $_1N_1$  имеет  $N_1 = N_0^2$  каналов, состоит из  $N_1$  ОПК $m$  и  $N_1$  внешних пар ДМ $m$  и МД $m$  и является неблокируемым самомаршрутизируемым коммутатором на произвольной перестановке пакетов. В него добавлены линии и выходные МД ОС от входов ОПК $m$  до входов ОПК $_1N_1$  и входные ДМ от выходов ОПК $m$  до выходов ОПК $_1N_1$  ОС (пунктир на рисунке 4). Сигнал «занято», который поступают на вход Д $_m$  ОС извне, распространяется на все его выходы. Схемная сложность ОПК $_1N_1$  оценивается как  $S_1 = N_1(S_0 + 4m) = 4N_1^2 + 4N_1^{3/2}$  точек коммутации, и, как показано в разделе 1, он является неблокируемым коммутатором на перестановочном трафике, т.е.  $T_1 = 1$ .

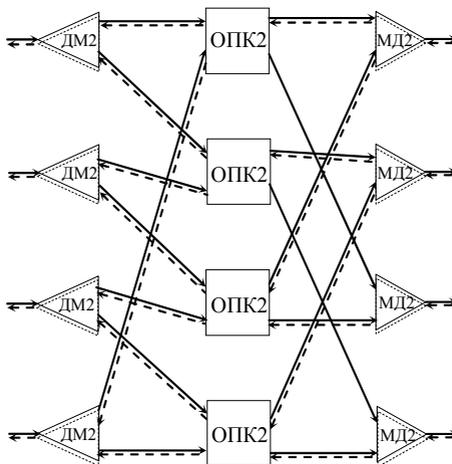


Рисунок 5. Однокаскадный 4-канальный общий пакетный коммутатор (ОПК $_{14}$ )

Для квазиперестановочного трафика было оценено среднее значение  $T_1$ , т.е. значение  $T_1^*$  в предположении равной вероятности любых маршрутов. Это было сделано опять же посредством численного моделирования для малых  $m$  и  $N_1 = m^2$ . Полученные значения  $T_1^*$  представлены в таблице 5.

Аналогичные оценки были сделаны для ОПК $_2N_2$  и ОПК $_4N_4$  с числом каналов  $N_2 = m^4$  и  $N_4 = m^8$ . Значения средних задержек  $T_2^*$  и  $T_4^*$  для них представлены в таблице 6. Схемные сложности ОПК $_2N_2$  и ОПК $_4N_4$  оцениваются по формулам (2) и (3) для  $S_2$  и  $S_4$  из раздела 1 с удвоенными коэффициентами. Коммутаторы ОПК $_2N_2$  и ОПК $_4N_4$  являются неблокируемыми коммутаторами на перестановочном трафике, т.е.  $T_2 = T_4 = 1$ .

Таблица 4. Средние задержки (в пакетах) разрешения конфликтов для ОПК $m$  на квазиперестановочном трафике

$N_0$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$T_0^*$	1,49	1,89	2,13	2,29	2,41	2,51	2,60	2,72	2,75	2,82	2,88	2,93	2,99	3,03	3,08

Таблица 5. Средние задержки (в пакетах) разрешения конфликтов для ОПК $1N_1$  на квазиперестановочном трафике

m	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$N_1$	4	9	16	25	36	49	64	81	100	121	144	169	196	225	256
$T_1^*$	2,12	2,68	3,08	3,38	3,61	3,79	3,96	4,10	4,23	4,34	4,45	4,54	4,62	4,70	4,77

Таблица 6. Средние задержки (в пакетах) разрешения конфликтов для ОПК $2N_2$  и ОПК $4N_4$  на квазиперестановочном трафике

m	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$N_2$	2 <sup>4</sup>	3 <sup>4</sup>	4 <sup>4</sup>	5 <sup>4</sup>	6 <sup>4</sup>	7 <sup>4</sup>	8 <sup>4</sup>	9 <sup>4</sup>	10 <sup>4</sup>	11 <sup>4</sup>	12 <sup>4</sup>	13 <sup>4</sup>	14 <sup>4</sup>	15 <sup>4</sup>	16 <sup>4</sup>
$T_2^*$	3,08	4,11	4,77	5,27	5,65	5,98	6,26	6,47	6,67	6,87	7,04	7,21	7,32	7,44	7,56
$N_4$	2 <sup>8</sup>	3 <sup>8</sup>	4 <sup>8</sup>	5 <sup>8</sup>	6 <sup>8</sup>	7 <sup>8</sup>	8 <sup>8</sup>								
$T_4^*$	4,77	6,64	7,55	8,37	9,01	9,65	10,10								

Метод получения значений  $T_i^*$  ( $i = 1, 2, 4$ ) аналогичен методу получения значений  $T_0^*$ . Он основывается на том, что конфликтные пакеты, адресованные разным приемникам, между собой не конфликтуют. Это свойство является следствием того факта, что рассматриваемые неблокируемые коммутаторы имеют структуру встречных деревьев демультимплексоров  $1 \times m$  и мультимплексоров  $m \times 1$ , соединенных через слой коммутаторов  $m \times m$ . Поэтому конфликтные пакеты не могут иметь разные адреса. В результате все конфликты на любой квазиперестановке разрешаются за время разрешения конфликта наибольшей кратности.

В данном разделе рассмотрены неблокируемые в широком смысле многокаскадные коммутаторы со сквозными цепями обратной связи, для которых оценены их схемные сложности и средние задержки передачи пакетов на квазиперестановочном трафике.

### 3. Непрокируемые сети промежуточной сложности

В основе этого раздела лежат неблокируемые на произвольном трафике пакетные коммутаторы ОПК $m$  и многокаскадные коммутаторы ОПК $rN_r$ , ( $r = 1, 24, 8$ ) из предыдущего раздела.

На основе ОПК $m$ , имеющих  $N_0 = m$  каналов и сложность  $S_0 = 4m^2 = 4N_0^2$ , можно построить новый 2-каскадный составной коммутатор СПК $_1M_1$  в виде сети с обменными связями и с  $M_1$  каналами, представленной на рисунке 6 для  $m = 2$  и  $N_0 = 2$ . Он состоит из  $N_0$  ОПК $m$  в каждом

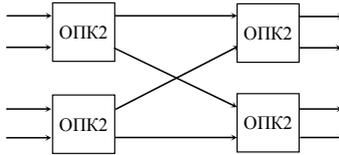


Рисунок 6. Двухкаскадный составной 4-канальный пакетный коммутатор СПК $_{14}$

каскаде и имеет  $M_1 = N_0^2 = 2m^2$  каналов. Коммутатор СПК $_1M_1$  является самомаршрутизируемым коммутатором на любом трафике. По построению схемная сложность СПК $_1M_1$  задается формулой

$$(5) \quad \Sigma_1 = 2N_0S_0 = 8N_0^3 = 8M_1^{3/2}.$$

Задержки перестановочного трафика оцениваются следующим образом. Порождается множество случайных перестановок между входами и выходами СПК $_1M_1$ . По каждой случайной перестановке определяется распределение маршрутов на выходах 1-го каскада и среди них выделяется значение наибольшей кратности, порождающей наибольшую задержку. Затем осуществляется усреднение этих значений по множеству произвольных перестановок.

Главной особенностью коммутатора  $СПК_1M_1$  является то, что распределения маршрутов на выходах 1-го каскада образуют последовательный набор разреженных перестановок. Эти перестановки передаются через 2-й каскад без дополнительных задержек, т.к. по построению он состоит из неблокируемых коммутаторов  $ОПК_m$ .

Задержки разрешения конфликтов для перестановочного трафика на  $ОПК_1M_1$  представлены в таблице 7, а для квазиперестановочного трафика – в таблице 5, считая  $M_1 = N_1$ .

Аналогично на основе  $ОПК_1N_1$  можно построить новый 2-каскадный составной коммутатор  $СПК_2M_2$  в виде сети с обменными связями и с  $M_2$  каналами, представленной на рисунке 7 для  $m=2$  и  $N_1 = 4$ . Он имеет  $M_2 = M_1^2 = 2(m^4)$  каналов, состоит из  $M_1$   $ОПК_1M_1$  в первом каскаде и из  $M_1$  блоков коммутаторов  $ОПК_m$ , каждый из которых состоит из  $N_0$  коммутаторов  $ОПК_m$ , во втором каскаде. Коммутаторы  $ОПК_1N_1$  и блоки коммутаторов  $ОПК_m$  имеют равное количество выходов и входов, по которым и осуществляются обменные связи.

Коммутатор  $СПК_2M_2$  является самомаршрутизируемым коммутатором на любом трафике. По построению схемная сложность  $СПК_2M_2$  задается формулой

$$(6) \quad \Sigma_2 = M_1(\Sigma_1 + N_0S_0) = m^2(4(m^4 + m^3) + 4m^3) = 4M_2^{6/4} + 8M_2^{5/4}.$$

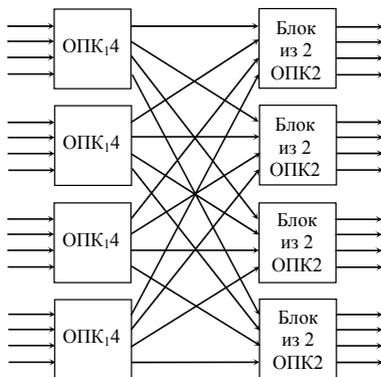


Рисунок 7. Двухкаскадный составной 16-канальный пакетный коммутатор  $СПК_216$

Задержки перестановочного трафика для  $СПК_2M_2$  оцениваются также как и для  $СПК_1M_1$ , а их значения представлены в таблице 8. При этом значения задержек для квазиперестановочного трафика на  $СПК_2M_2$  представлены в таблице 6, считая  $M_2 = N_2$ .

Таблица 7. Средние задержки (в пакетах) разрешения конфликтов для СПК<sub>2</sub>M<sub>2</sub> на перестановочном трафике

m	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$M_1$	2 <sup>2</sup>	3 <sup>2</sup>	4 <sup>2</sup>	5 <sup>2</sup>	6 <sup>2</sup>	7 <sup>2</sup>	8 <sup>2</sup>	9 <sup>2</sup>	10 <sup>2</sup>	11 <sup>2</sup>	12 <sup>2</sup>	13 <sup>2</sup>	14 <sup>2</sup>	15 <sup>2</sup>	16 <sup>2</sup>
$T_1^*$	1,30	1,97	2,35	2,67	2,94	3,16	3,33	3,49	3,64	3,77	3,91	4,02	4,12	4,22	4,30

Таблица 8. Средние задержки (в пакетах) разрешения конфликтов для СПК<sub>2</sub>M<sub>2</sub> на перестановочном трафике

m	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$M_2$	2 <sup>4</sup>	3 <sup>4</sup>	4 <sup>4</sup>	5 <sup>4</sup>	6 <sup>4</sup>	7 <sup>4</sup>	8 <sup>4</sup>	9 <sup>4</sup>	10 <sup>4</sup>	11 <sup>4</sup>	12 <sup>4</sup>	13 <sup>4</sup>	14 <sup>4</sup>	15 <sup>4</sup>	16 <sup>4</sup>
$T_2^*$	2,34	3,48	4,30	4,87	5,35	5,71	6,04	6,31	6,51	6,72	6,91	7,09	7,24	7,36	7,50

Теперь на основе  $ОПК_2N_2$  можно построить новый 2-каскадный составной коммутатор  $СПК_4M_4$  в виде сети с обменными связями и с  $M_4$  каналами. Он имеет  $M_4 = M_2^2 = m^8$  каналов и состоит из  $M_2$  коммутаторов  $ОПК_2N_2$  в первом каскаде и  $M_2$  блоков коммутаторов  $ОПКm$  во втором каскаде. При этом каждый блок состоит из  $M_1/N_0$  коммутаторов  $ОПКm$ . Коммутаторы  $ОПК_2N_2$  и блоки коммутаторов  $ОПКm$  имеют равное количество выходов и входов, по которым и осуществляются обменные связи.

Коммутатор  $СПК_4M_4$  является самомаршрутизируемым коммутатором на любом трафике. Схемная сложность  $СПК_4M_4$  оценивается как величина  $\Sigma_4 = M_2(2S_2 + M_1S_0)$  точек коммутации, где значение  $S_2$  берется из формулы (2) раздела ??.

Задержки перестановочного трафика для  $СПК_4M_4$  оцениваются также как и для  $СПК_2M_2$ , а их значения представлены в таблице 9. При этом

Таблица 9. Средние задержки (в пакетах) разрешения конфликтов для  $СПК_4M_4$  на перестановочном трафике

m	2	3	4	5	6	7	8
$M_4$	256	6 561	65 536	390 625	1 679 616	5 764 801	16 777 216
$T_4^*$	4,30	6,30	7,49	8,35	8,99	9,53	10,14

значения задержек для квазиперестановочного трафика представлены в таблице 6, считая  $M_4 = N_4$ .

Наконец, на основе  $ОПК_4N_4$  можно построить новый 2-каскадный составной коммутатор  $СПК_8M_8$  в виде сети с обменными связями и с  $M_8$  каналами. Он имеет  $M_8 = M_4^2 = m^{16}$  каналов, состоит из  $M_4$   $ОПК_4N_4$  в первом каскаде и из  $M_4$  блоков коммутаторов  $ОПКm$  во втором каскаде. При этом каждый блок состоит из  $M_2/N_0$  коммутаторов  $ОПКm$ . Коммутаторы  $ОПК_4N_4$  и блоки коммутаторов  $ОПКm$  имеют равное количество выходов и входов, по которым и осуществляются обменные связи.

Коммутатор  $СПК_8M_8$  является самомаршрутизируемым коммутатором на любом трафике. Схемная сложность  $СПК_8M_8$  оценивается как величина  $\Sigma_8 = M_4(2S_4 + M_2S_0)$  точек коммутации, где значение  $S_4$  берется из формулы (3) раздела 1.

Задержки перестановочного трафика для  $СПК_8M_8$  оцениваются также как и для  $СПК_2M_2$ , а их значения представлены в таблице 10.

Таблица 10. Средние задержки(в пакетах) разрешения конфликтов для СПК<sub>8</sub>M<sub>8</sub> на перестановочном трафике

m	2	3
M <sub>8</sub>	65 536	43 046 721
T <sub>8</sub> *	7,48	10,35

Сравнение задержек на перестановочном (таблицы 7– 9) и квази-перестановочном (таблицы 5– 6) трафиках показывает, что последний мажорирует сверху первый. Единственное нарушение для  $N_4 = 8^8$  укладывается в среднеквадратичное отклонение. Этот факт используется в следующем разделе.

В таблице 11 для произвольных перестановок приводятся результаты сравнения быстрого и промежуточного коммутаторов по критерию отношения сложности к средним задержкам. Сложности быстрых коммутаторов берутся как  $S_i$  из таблицы 3, а критерии как  $\aleph_i = \Sigma_i/T_i$  с  $T_i = 1$ . Сложности промежуточных коммутаторов – это рассмотренные выше  $\Sigma_i$ , которые приведены в таблице 11 в показательном виде, а критерии – это  $\aleph_i = \Sigma_i/T_i^*$  с  $T_i^*$  из таблиц 7– 10. Сравнение проводится по отношениям критериев, т.е. по значениям  $\mathfrak{p}_i = \Sigma_i/(T_i^* S_i)$ , большее значение которых задает превосходство по критерию. По таблице 11 видно, что промежуточные коммутаторы имеют преимущество только при большом числе каналов и тем большее, чем больше число каналов.

Таблица 11. Сравнение быстрого и промежуточного коммутаторов по критерию  $\aleph_i$ 

m	$\Sigma_1$	$\Sigma_2$	$\Sigma_4$	$\Sigma_8$	$\mathfrak{p}_1$	$\mathfrak{p}_2$	$\mathfrak{p}_4$	$\mathfrak{p}_8$
2	$M_1^{3,0}$	$M_2^{2,25}$	$M_4^{1,87}$	$M_8^{1,69}$	0,58	0,80	1,9	17,2
3	$M_1^{2,45}$	$M_2^{1,93}$	$M_4^{1,70}$	$M_8^{1,60}$	0,51	1,15	6,5	317
4	$M_1^{2,25}$	$M_2^{1,82}$	$M_4^{1,65}$	$M_8^{1,58}$	0,53	1,65	17,1	
5	$M_1^{2,15}$	$M_2^{1,77}$	$M_4^{1,62}$	$M_8^{1,56}$	0,56	2,29	37,4	
6	$M_1^{2,08}$	$M_2^{1,73}$	$M_4^{1,61}$	$M_8^{1,55}$	0,60	3,03	72,1	
7	$M_1^{2,03}$	$M_2^{1,71}$	$M_4^{1,60}$	$M_8^{1,55}$	0,63	3,89	126	
8	$M_1^{2,0}$	$M_2^{1,69}$	$M_4^{1,59}$	$M_8^{1,55}$	0,68	4,84	202	

В некотором смысле промежуточными свойствами обладают и неблокируемые сети Клоза [4, 5] – доказано их существование и известен способ построения, но нет никакой процедуры параллельной самомаршрутизации. Эти сети имеют элементную базу из квадратных коммутаторов  $N \times N$  и трапециевидных коммутаторов  $N \times 2N$  и  $2N \times N$ .

Для последующих сравнений оценим сложность 5-каскадного коммутатора Клоза, построенного на базе коммутатора YARK [9] с  $N_0 = 64$  каналами и коммутационной сложностью  $S_0 = 64^2$ . Для реализации трапецевидного коммутатора  $32 \times 64$  используется один коммутатор YARK, и он же используется для реализации 2-х квадратных коммутаторов  $32 \times 32$ .

В 3-каскадном варианте неблокируемого коммутатора входной и выходной каскады содержат  $N_1 = 32$  трапецевидных коммутаторов, хребет содержит  $2N_1$  квадратных коммутаторов, все они содержат по 32 коммутатора YARK. 3-каскадный вариант имеет  $N_3 = 1024$  каналов и коммутационную сложность  $S_3 = N_3^{1,86}$ .

В 5-каскадном варианте входной и выходной каскады содержит  $N_3$  трапецевидных коммутаторов, а хребет содержит  $2N_1$  3-каскадных коммутаторов Клоза. 5-каскадный вариант имеет  $N_5 = 32 \cdot 678$  каналов и коммутационную сложность  $S_5 = 2N_3S_0 + 2N_1S_3 = N_5^{1,67}$ .

Таким образом, теоретический неблокируемый коммутатор Клоза при сравнимом числе каналов (см. таблицу 11) имеет сопоставимую или большую сложность по сравнению с неблокируемыми коммутаторами СПК<sub>4</sub>N<sub>4</sub> и СПК<sub>8</sub>N<sub>8</sub> только при  $m \leq 3$  и  $m = 1$  соответственно. При этом он имеет большее быстродействие, но практически не реализуем из-за отсутствия процедуры параллельной самомаршрутизации.

В данном разделе рассмотрены промежуточные коммутаторы, полученные посредством расширения быстрых коммутаторов второго раздела расширенные сетями с обменными связями. Оценены схемные сложности и средние задержки передачи пакетов в таких коммутаторах на перестановочном трафике. Показано, что по сложности они сопоставимы с неблокируемыми коммутаторами Клоза, а по критерию сложность/задержка в несколько раз превосходят быстрые коммутаторы при большом числе каналов.

#### 4. Неблокируемые сети малой сложности

Коммутаторы этого раздела строятся на основе  $m$ -канального составного коммутатора СПК<sub>1</sub>M<sub>1</sub> (рисунок 3), рассмотренного в предыдущем разделе. Из него строится 2-каскадный коммутатор с обменными связями ОСПК<sub>2</sub>Λ<sub>2</sub> с  $\Lambda_2 = M_1^2 = m^4$  каналами. Каждый каскад состоит из  $M_1$  коммутаторов СПК<sub>1</sub>M<sub>1</sub>, связанных обменными связями. Сложность ОСПК<sub>2</sub>N<sub>2</sub> составляет величину  $Z_2 = 2M_1\Sigma_1 = 16M_1^{5/2} = 16\Lambda_2^{5/4}$  точек коммутации. Значения средних задержек  $T_2^*$  могут быть оценены только для квазиперестановочного трафика, поскольку СПК<sub>1</sub>M<sub>1</sub> во втором

каскаде имеет  $T_1^* > 1$ , что затрудняет оценку для перестановочного графика. Проведенные задержки представлены в таблице 6.

Из коммутатора ОСПК<sub>2</sub>Λ<sub>2</sub> строится 2-каскадный коммутатор с обменными связями ОСПК<sub>4</sub>Λ<sub>4</sub> с  $\Lambda_4 = \Lambda_2^2$  каналами. Каждый каскад состоит из Λ<sub>2</sub> коммутаторов ОСПК<sub>2</sub>Λ<sub>2</sub>, связанных обменными связями. Сложность ОСПК<sub>4</sub>Λ<sub>4</sub> составляет величину  $Z_4 = 2\Lambda_2 Z_2 = 1\ 32N_2^{9/4} = 1\ 32N_4^{9/8}$  точек коммутации. При этом значения средних задержек  $T_4^*$  для квазиперестановочного графика представлены также в таблице 6.

Аналогично из коммутатора ОСПК<sub>4</sub>Λ<sub>4</sub> строится 2-каскадная сеть с обменными связями ОСПК<sub>8</sub>Λ<sub>8</sub> с  $\Lambda_8 = \Lambda_4^2 = m^{16}$  каналами. Каждый каскад состоит из Λ<sub>4</sub> коммутаторов ОСПК<sub>4</sub>Λ<sub>4</sub>, связанных обменными связями. Сложность ОСПК<sub>8</sub>Λ<sub>8</sub> составляет величину  $Z_8 = 2\Lambda_4 Z_4 = 1\ 64\Lambda_4^{17/8} = 1\ 64\Lambda_8^{17/16}$  точек коммутации. При этом значения средних задержек  $T_8^*$  для квазиперестановочного графика представлены в таблице 6.

Таблица 12. равнение быстрого и малого коммутаторов по критерию  $\aleph_i$

m	$Z_1$	$Z_2$	$Z_4$	$Z_8$	$\mathfrak{p}_1$	$\mathfrak{p}_2$	$\mathfrak{p}_4$	$\mathfrak{p}_8$
2	$N_1^{3,0}$	$\Lambda_2^{2,25}$	$\Lambda_4^{1,75}$	$\Lambda_8^{1,44}$	0,58	0,61	3,3	274
3	$N_1^{2,45}$	$\Lambda_2^{1,88}$	$\Lambda_4^{1,52}$	$\Lambda_8^{1,30}$	0,51	1,2	30,9	64 986
4	$N_1^{2,25}$	$\Lambda_2^{1,75}$	$\Lambda_4^{1,44}$	$\Lambda_8^{1,25}$	0,53	2,2	181	
5	$N_1^{2,15}$	$\Lambda_2^{1,68}$	$\Lambda_4^{1,39}$	$\Lambda_8^{1,22}$	0,56	3,7	712	
6	$N_1^{2,08}$	$\Lambda_2^{1,64}$	$\Lambda_4^{1,37}$	$\Lambda_8^{1,21}$	0,60	5,7	2330	
7	$N_1^{2,03}$	$\Lambda_2^{1,61}$	$\Lambda_4^{1,35}$	$\Lambda_8^{1,20}$	0,63	8,4	6223	
8	$N_1^{2,0}$	$\Lambda_2^{1,58}$	$\Lambda_4^{1,33}$	$\Lambda_8^{1,19}$	0,68	11,7	14773	

Необходимо отметить, что сложности коммутаторов ОСПК<sub>4</sub>Λ<sub>4</sub> и ОСПК<sub>8</sub>Λ<sub>8</sub> приближаются к сложностям многокаскадных перестраиваемых сетей [5] с сопоставимым числом каналов и при вдвое меньшем числе каскадов.

Коммутаторы ОСПК<sub>4</sub>Λ<sub>4</sub> и ОСПК<sub>8</sub>Λ<sub>8</sub> демонстрируют преимущество по отношениям  $\mathfrak{p}_4$  и  $\mathfrak{p}_8$  в сравнении с быстрыми коммутаторами в вариантах с большим числом каналов. При этом их преимущество оказывается значительно больше, чем у промежуточных коммутаторов предыдущего раздела. Для перестановочного графика отмеченное преимущество не уменьшится.

## Заключение

В работе предложено коммутаторы с прямыми каналами снабжать сквозными цепями обратной связи. Последние позволяют обнаруживать любые конфликты пакетов и делать это в стартстопном режиме для пакетов с одинаковой адресацией и в параллельном режиме для пакетов с разной адресацией. Это позволяет разрешать конфликты пакетов посредством их повторной передачи с минимальными задержками в тех же режимах.

Для коммутаторов с прямыми каналами предложено расширить понятие неблокируемости до возможности передачи пакетов без потерь на канальном уровне. При таком расширении коммутаторы с прямыми каналами и с цепями обратной связи становятся неблокируемыми на любом трафике, но с разными задержками передачи исходного набора пакетов.

Были построены коммутаторы с прямыми каналами разной сложности и с разными задержками реализации произвольных перестановок. Коммутаторы промежуточной сложности при большом числе каналов сопоставимы по сложности с неблокируемым коммутатором Клоза, но имеют большие задержки. При этом они, в отличие от коммутатора Клоза, реализуемы не только теоретически, но и практически.

Построенные коммутаторы малой сложности сопоставимы по сложности с перестраиваемыми сетями с одинаковой элементной базой. При этом они имеют меньшие задержки вследствие используемых режимов обнаружения и разрешения конфликтов только на канальном уровне по сравнению с большими протокольными задержками обнаружения конфликтов в перестраиваемых сетях.

Для всех построенных коммутаторов задержки передачи были оценены для произвольного перестановочного и квазиперестановочного трафиков.

Разные коммутаторы были сопоставлены по критерию сложность/задержка. Было выявлено, что по этому критерию быстрые и сложные коммутаторы уступают более простым и медленным коммутаторам только в вариантах с большим числом каналов.

## Список литературы

- [1] Барабанова Е. А., Выговтов К. А., Подлазов В. С. *Неблокируемые отказоустойчивые двухкаскадные дуальные фотонные коммутаторы* // Проблемы управления. – 2021. – № 4. – С. 82–92. [doi](#) ↑116, 124
- [2] Подлазов В. С. *Самомаршрутизируемая неблокируемая системная сеть с прямыми каналами: сложность и быстродействие* // Программные системы: теория и приложения. – 2022. – Т. 13. – № 4(55). – С. 47–76. [doi](#) ↑116

- [3] Stepanenko S. *Structure and implementation principles of a photonic computer*, EPJ Web of Conferences.– Vol. **224**.– 2019.– id. 04002.– 7 pp.  [↑116](#)
- [4] Абрамов С. М., Степаненко С. А. *О подходах к разработке программного обеспечения для фотонной вычислительной машины*, Национальный Суперкомпьютерный Форум (НСКФ-2022) (Россия, Переславль-Залесский, ИПС имени А.К. Айламазяна РАН, 29 ноября–02 декабря 2022 года).– 2022.  [↑116, 132](#)
- [5] Pipenger N. *On rearrangeable and non-blocking switching networks* // J. Comput. Syst. Sci.– 1978.– Vol. **17**.– No. 2.– Pp. 145–162.  [↑116, 132, 134](#)
- [6] Clos C. *A study of non-locking switching networks* // Bell System Tech. J.– 1953.– Vol. **32**.– No. 2.– Pp. 406–424.  [↑116](#)
- [7] Бенеш В. Э. // *Математические основы теории телефонных сообщений*, М.: Связь.– 1968.– С. 83–150. [↑116](#)
- [8] Каравай М. Ф., Подлазов В. С. *Идеальная системная сеть на основе минимального квазитоположного орграфа* // *Системные сети с прямыми каналами для параллельных вычислительных систем – комбинаторный подход*, Глава 5.– С. 111–124.  [↑118](#)
- [9] Scott S., Abts D., Kim J., Dally W. *The Black Widow High-radix Clos network* // *Proc. 33rd Intern. Symp. Comp. Arch., ISCA'2006* (Boston, MA, USA, 17–21 June 2006).– 2006.– ISBN 0-7695-2608-X.– Pp. 16–28.   [↑124, 133](#)

Поступила в редакцию	02.02.2023;
одобрена после рецензирования	11.05.2023;
принята к публикации	11.05.2023;
опубликована онлайн	07.10.2023.

Рекомендовал к публикации

к.ф.-м.н. С. А. Романенко

## Информация об авторе:



Виктор Сергеевич Подлазов

Д. т. н., гл.н.с. Института проблем управления им. В.А. Трапезникова РАН, научные интересы: архитектуры интерконнекта и маршрутизация в суперкомпьютерных системах

 0000-0002-9175-1138  
 e-mail: [podlazov@ipu.ru](mailto:podlazov@ipu.ru)  
[podlazov@gmail.com](mailto:podlazov@gmail.com)

Автор заявляет об отсутствии конфликта интересов.



# Multichannel non-blocking system area network with direct channels

Viktor Sergeevich **Podlazov**

V. A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Russia

 [podlazov@ipu.ru](mailto:podlazov@ipu.ru)

**Abstract.** Three new versions of a non-blocking self-routed network with direct channels have been developed, in which packet collisions are detected and resolved at the entrance to the network. Collision-blocked packets are retransmitted by sources with minimal delay. End-to-end feedback loops have been added to the network to quickly detect packet collisions. The presence of feedback loops allows for arbitrary traffic without packet loss. Networks are developed in 1-, 2-, 4-, and 8-cascade variants with scaling the number of channels from several hundred to many millions with an assessment of the network speed achieved on permutation traffic. The practical orientation of the proposed networks is system networks with the transmission of routing information in packet headers using control routing information in each cascade. The proposed networks are made in an extended circuit basis, consisting of switches, multiplexers and demultiplexers. The paper presents the characteristics of the constructed networks and gives their comparative assessment. The novelty of the developed networks lies both in the very fact of their construction, and in the structure of the networks used and their characteristics. (*In Russian*).

**Key words and phrases:** direct links, dual switch, packet switch, multiplexers and demultiplexers, multistage switch, conflict-free routing, non-blocking switch, static self-routing, quasi-complete digraph, complexity and speed

2020 *Mathematics Subject Classification:* 65Y05; 68Q10

**For citation:** Viktor S. Podlazov. *Multichannel non-blocking system area network with direct channels*. Program Systems: Theory and Applications, 2023, 14:3(58), pp. 115–138. (*In Russ.*). [https://psta.psisras.ru/read/psta2023\\_3\\_115-138.pdf](https://psta.psisras.ru/read/psta2023_3_115-138.pdf)

## References

- [1] E. A. Barabanova, K. A. Vytovtov, V. S. Podlazov. “Neblokiruemye otkazoustojchivye dvuxkaskadnye dual’nye fotonnye kommutatory”, *Problemy upravleniya*, 2021, no. 4, pp. 82–92; A. Barabanova E., A. Vytovtov K., S. Podlazov V.. “Non-blocking fault-tolerant two-stage dual photon switches”, *Control Sciences*, 2021, no. 4, pp. 67–76. 
- [2] V. S. Podlazov. “Multichannel non-blocking system area network with direct channels”, *Program Systems: Theory and Applications*, **13**:4(55) (2022), pp. 47–76 (in Russian). 
- [3] S. Stepanenko. *Structure and implementation principles of a photonic computer*, EPJ Web of Conferences, vol. **224**, 2019, id. 04002, 7 pp. 
- [4] S. M. Abramov, S. A. Stepanenko. *On approaches to developing software for a photonic computer*, Nacional’nyj Superkomp’yuternyj Forum (NSKF-2022) (Rossiya, Pereslavl’-Zalesskij, IPS imeni A.K. Ajlamazyana RAN, 29 noyabrya–02 dekabrya 2022 goda), 2022 (in Russian). 
- [5] N. Pipenger. “On rearrangeable and non-blocking switching networks”, *J. Comput. Syst. Sci.*, **17**:2 (1978), pp. 145–162. 
- [6] C. Clos. “A study of non-locking switching networks”, *Bell System Tech. J.*, **32**:2 (1953), pp. 406–424. 
- [7] V. E. Benesh, *Matematicheskie osnovy teorii telefonnyx soobshhenij*, Svyaz’, M., 1968, pp. 83–150; Beneš V.E.. *Mathematical Theory of Connecting Networks and Telephone Traffic*, Academic Press, 1965, ISBN 9780080955230, 318 pp.
- [8] M. F. Karavaj, V. S. Podlazov. “An ideal system network based on a minimal quasicomplete digraph”, *System networks with direct channels for parallel computing systems - a combinatorial approach*, Glava 5, pp. 111–124 (in Russian). 
- [9] Scott S., Abts D., Kim J., Dally W.. “The Black Widow High-radix Clos network”, *Proc. 33rd Intern. Symp. Comp. Arch.*, ISCA’2006 (Boston, MA, USA, 17–21 June 2006), 2006, ISBN 0-7695-2608-X, pp. 16–28.  