

# ДИАГНОСТИКА ПАТОЛОГИЙ ПО ДАННЫМ ВИДЕОЭНДОСКОПИИ С ИСПОЛЬЗОВАНИЕМ АНСАМБЛЯ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

DOI: 10.17691/stm2018.10.2.01

УДК 004.932.72–07

Поступила 24.11.2017 г.



**С.В. Аксенов**, к.т.н., доцент кафедры теоретических основ информатики<sup>1</sup>; доцент отделения информационных технологий<sup>2</sup>; доцент кафедры автоматизации обработки информации<sup>3</sup>;

**К.А. Костин**, магистрант Научно-образовательного центра компьютерных наук и технологий<sup>1</sup>; аспирант отделения информационных технологий<sup>2</sup>;

**А.В. Иванова**, магистрант Научно-образовательного центра компьютерных наук и технологий<sup>1</sup>;

**J. Liang**, PhD, Associate Professor, Biodesign Center for Biosignatures Discovery Automation<sup>4</sup>;

**А.В. Замятин**, д.т.н., доцент, зав. кафедрой теоретических основ информатики<sup>1</sup>;

директор Научно-образовательного центра компьютерных наук и технологий<sup>1</sup>

<sup>1</sup>Национальный исследовательский Томский государственный университет, Томск, 634050, пр. Ленина, 36;

<sup>2</sup>Национальный исследовательский Томский политехнический университет, Томск, 634050, пр. Ленина, 30;

<sup>3</sup>Томский государственный университет систем управления и радиоэлектроники, Томск, 634050, пр. Ленина, 40;

<sup>4</sup>Arizona State University, University Center, 411 N Central Ave., Phoenix, AZ, 85004, USA

Предлагается технология построения классификатора образов патологических образований, полученных с помощью видеоэндоскопии с применением методов глубинного обучения (deep learning). Для обучения и тестирования нейросетевых моделей использованы выборки публичной базы данных CVC-ColonDB и 20 видеозаписей процесса колоноскопии Университета штата Аризона (Феникс, США). Для повышения надежности модели классификации данные зашумлены эффектами, которые присущи съемке современными камерами, используемыми при эндоскопии, а также проведено исследование построения модели при выборках небольшого размера.

При построении классификатора использованы результаты современных исследований моделей сверточных нейронных сетей в медицинской диагностике, что позволяет применять данный подход при проектировании архитектуры сверточных нейронных сетей в зависимости от особенностей задачи. Путем обобщения особенностей архитектур успешных моделей разработан новый подход к формированию избыточной сверточной нейронной сети. Согласно предлагаемому подходу, архитектура сети разделяется на блоки с определенными значениями характеристик, чередование которых позволяет сформировать наиболее эффективную структуру.

С использованием предложенного подхода к формированию оптимальной архитектуры сверточных нейронных сетей на основе рекомендаций по выбору значений характеристик сети и ранжирования наиболее значимых из них разработан второй подход к построению адаптивной модели классификатора. Он основывается на формировании ансамбля классификаторов типа «сверточная нейронная сеть». Для обеспечения устойчивости к изменению исходных данных и широкой применимости к различным классам задач классификации изображений в ансамбль входит набор сетей с отличными друг от друга наиболее значимыми факторами.

Экспериментальные исследования показали, что классификатор имеет потенциал улучшения качества распознавания путем разработки ансамбля сверточных нейронных сетей с учетом зависимостей, рассмотренных в предложенном подходе. Полученные результаты работы демонстрируют перспективность применения разработанного подхода для построения моделей классификации образов не только в ходе решения задач медицинской диагностики, но и для общих задач машинного зрения при малой выборке.

**Ключевые слова:** глубинное обучение; deep learning; сверточная нейронная сеть; классификатор патологий; медицинская диагностика.

**Для контактов:** Замятин Александр Владимирович, e-mail: avzamyatin@inbox.ru

**Как цитировать:** Aksenov S.V., Kostin K.A., Ivanova A.V., Liang J., Zamyatin A.V. An ensemble of convolutional neural networks for the use in video endoscopy. *Sovremennye tehnologii v medicine* 2018; 10(2): 7–19, <https://doi.org/10.17691/stm2018.10.2.01>

English

## An Ensemble of Convolutional Neural Networks for the Use in Video Endoscopy

**S.V. Aksenov**, PhD, Associate Professor, Department of Theoretical Foundations of Informatics<sup>1</sup>; Associate Professor, Department of Information Technologies<sup>2</sup>; Associate Professor, Department of Information Processing Automation<sup>3</sup>;

**K.A. Kostin**, MSc Student, Scientific and Educational Center of Computer Science and Technology<sup>1</sup>; PhD Student, Department of Information Technologies<sup>2</sup>;

**A.V. Ivanova**, MSc Student, Scientific and Educational Center of Computer Science and Technology<sup>1</sup>;

**J. Liang**, PhD, Associate Professor, Biodesign Center for Biosignatures Discovery Automation<sup>4</sup>;

**A.V. Zamyatin**, DSc, Associate Professor, Head of the Department of Theoretical Foundations of Informatics<sup>1</sup>; Director of the Scientific and Educational Center of Computer Science and Technology<sup>1</sup>

<sup>1</sup>National Research Tomsk State University, 36 Lenin Avenue, Tomsk, 634050, Russia;

<sup>2</sup>National Research Tomsk Polytechnic University, 30 Lenin Avenue, Tomsk, 634050, Russia;

<sup>3</sup>Tomsk State University of Control Systems and Radioelectronics, 40 Lenin Avenue, Tomsk, 634050, Russia;

<sup>4</sup>Arizona State University, University Center, 411 N Central Avenue, Phoenix, AZ, 85004, USA

In this study, a technology for creating a classifier able to identify pathological formations in images obtained with video endoscopy using the methods of deep learning is proposed. For the training and testing of neural network models, images from the CVC-ColonDB open database and 20 colonoscopy video records from the University of Arizona (Phoenix, USA) were used. To improve the performance of the proposed classification model, noise effects inherent to video cameras were considered. In addition, a study on building the model using small data samples was conducted.

In building the classifier, we utilized the results of recent studies on convolutional neural networks used in medical diagnostics, which allows us to apply the proposed approach to designing the architecture of a convolutional neural network adapted to a given task. By generalizing the features of the successful models, we developed an approach towards creating a non-excessive convolutional neural network. According to the proposed approach, the network architecture is divided into blocks, which alternate to enable composing the most efficient architecture.

Using the proposed approach based on the recommended selection strategy and then ranking the most significant parameters, a second approach towards building an adaptive model of classifier has been proposed. It is based on the formation of an ensemble of classifiers such as the “convolutional neural network”. To ensure the stability of the model and its insensitivity to changes in the input data as well as its applicability to different classification tasks, a set of networks with different major parameters are incorporated into the ensemble.

Our experimental studies have shown that the proposed classifier can be improved by developing an ensemble of convolutional neural networks, which considers the functions proposed in the present approach. The results imply the prospective application of the developed approach for building classification models not only for medical diagnostics but also for general problems of machine vision based on small samples.

**Key words:** deep learning; convolutional neural network; classifier of pathologies; medical diagnostics.

### Введение

Для многих отраслей экономики и научных исследований видеоналитика, использующая методы интеллектуального анализа, которые основаны на машинном обучении, стала одним из эффективных методов контроля работ, диагностики опасных ситуаций, детектирования особых объектов и событий [1, 2]. Однако для медицинских систем разработка таких приложений представлена небольшими проектами, связанными с локализацией определенных структурных особенностей органов и тканей, как, например,

физические размеры и окраска новообразований. По результатам анализа такие приложения могут сигнализировать о степени опасности детектированного объекта. К работам этого типа относятся исследования в области обработки данных от эндоскопа при оценке состояния внутренних поверхностей органов пищеварительного тракта, кишечника [3, 4], а также анализ МРТ-изображений различных органов [5, 6].

К первой группе следует отнести исследования в области эндоскопии и развития методов видеоналитики для пищеварительного тракта путем использования ансамбля классификаторов, обрабатывающих

результаты сравнения областей интереса и локальных бинарных паттернов на разных масштабах [7, 8]. Научная группа под руководством В. Li [9] предложила использовать дескриптор форм, применяемый в стандарте MPEG-7 ART, и улучшенные моменты Зернике, не чувствительные к повороту образа, как инструменты обучения многослойного перцептрона для классификатора областей патологий. В дальнейшей работе эта группа усилила классификатор включением в качестве признаков характеристик интенсивности, насыщенности и оттенка и сравнением результатов работы на обновленной выборке классификаторов на основе многослойного перцептрона и машины опорных векторов [10].

Этим же типом работ, но для решения задач локализации полипов в прямом кишечнике с помощью методов глубинного обучения, занимаются исследовательские группы N. Tajbakhsh [11] (в задаче используется классификатор типа «случайный лес», анализирующий компоненты дискретного косинусного преобразования фрагментов области интереса) и S.Y. Park [12], предлагающего использовать алгоритмы сегментации путем поиска активных контуров. Особенность этих подходов в том, что из каждого кадра видеопотока извлекаются области, визуально схожие с округлыми объектами.

Ко второй группе можно отнести проекты, использующие методы машинного обучения для анализа МРТ-изображений различных человеческих органов. Например, для рака молочной железы разными инструментами (УЗИ, МРТ, КТ и ПЭТ) извлекались характеристики аномальных областей: размеры патологии, наличие кальцификатов, асимметрия, повышенное поглощение трассировки радиофармпрепарата с увеличением массы опухоли (для ПЭТ) [13]. Далее характеристики, полученные с помощью каждой технологии, анализировались алгоритмом «случайный лес» и затем сравнивались результаты ROC-анализа по отдельности для каждого метода лучевой диагностики. При поиске новообразований в головном мозге для извлечения характеристик использовались главные компоненты коэффициентов вейвлета дискретного вейвлетового преобразования сегментированного МРТ-изображения и многослойный перцептрон в качестве классификатора [14]. В исследовании [15], посвященном локализации области рака простаты, использовалась трехмерная сверточная нейронная сеть (СНС), на вход которой поступали диффузионно-взвешенные карты, карты внешнего коэффициента диффузии и динамические контрастные МРТ-карты. Предложенный обзор иллюстрирует разные, существенно различающиеся подходы к формированию признакового пространства и классификаторов для анализа визуальных данных.

Однако в качестве первичного и наиболее доступного исследования тканей, поддающихся визуальному осмотру, используется эндоскопия. Данная процедура ежедневно применяется огромным числом

медицинских специалистов для выявления различных патологий и способствует достаточно точной постановке диагноза. Это предполагает визуальное исследование состояния внутренних тканей человека специальным инструментом, снабженным миниатюрной камерой. В процессе обследования с камеры производится видеозапись [16]. Для повышения качества диагностики, а также для уменьшения количества медицинских ошибок исследователи классифицируют области с характеристиками, отличающимися от установленной современной медициной нормы, что является существенным для постановки диагноза и определения состояния здоровья пациентов [17–19].

**Цель исследования** — разработка подхода к построению модели классификации патологических образований на основе данных видеозендоскопии с использованием сверточных нейронных сетей. Модель должна обладать высокими обобщающими способностями, а также нечувствительностью к разнообразным зрительным эффектам, вызванным как характеристиками камеры, так и особенностями процесса съема данных.

### Технология построения классификатора

Технология глубинного обучения, в частности алгоритмы настройки и использования СНС, является на текущий момент одним из самых совершенных инструментов, используемых для классификации зрительных образов [20]. К достоинству этой технологии можно отнести обучаемость, адаптивность, интуитивное понимание принципов работы и универсальность при решении различных типов задач анализа образов.

Канонически архитектура СНС (рис. 1) описывается следующими параметрами:

$N$  — размер плоскости в слое; для плоскости во входном слое представляет собой произведение высоты  $H$  на ширину изображения  $W$ ;

$D$  — глубина входного слоя сети; в нашем случае — количество цветовых каналов изображения;

$P$  — количество строк и столбцов, добавляемых к границам слоя, который предшествует сверточному слою, и заполняемых нулями;

$S$  — смещение между фильтрами при формировании сигналов нейронов в слоях свертки и субдискретизации;

$F$  — размер квадратных фильтров сверточного слоя;

$Filters$  — глубина сверточного слоя (количество фильтров);

$U$  — размер квадратного окна для слоя объединения;

$Subf$  — тип функции слоя субдискретизации ( $\max$  — максимум либо  $\text{avg}$  — расчет среднего);

$K$  — количество нейронов в полносвязном слое;

$C$  — количество классов в задаче, принадлежность к которым определяется классификатором (для нашей классификации  $C=2$  — т.е. полип/нет полипа);

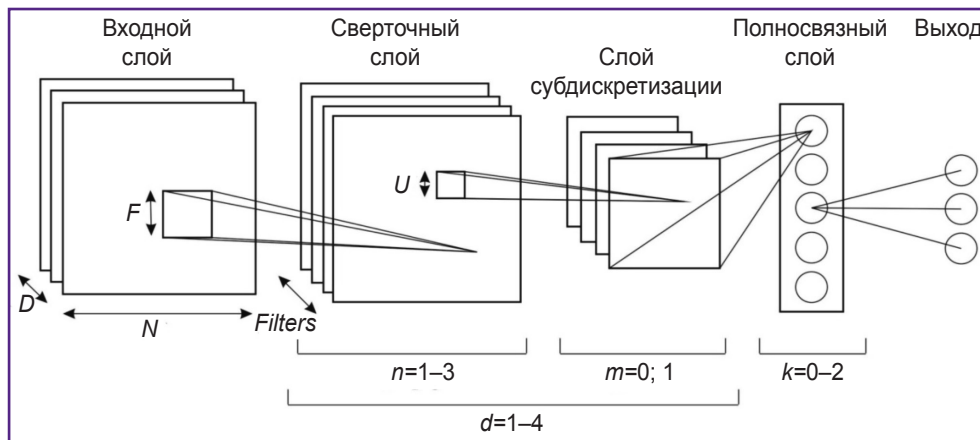


Рис. 1. Каноническая архитектура сверточной нейронной сети, обозначения см. в тексте

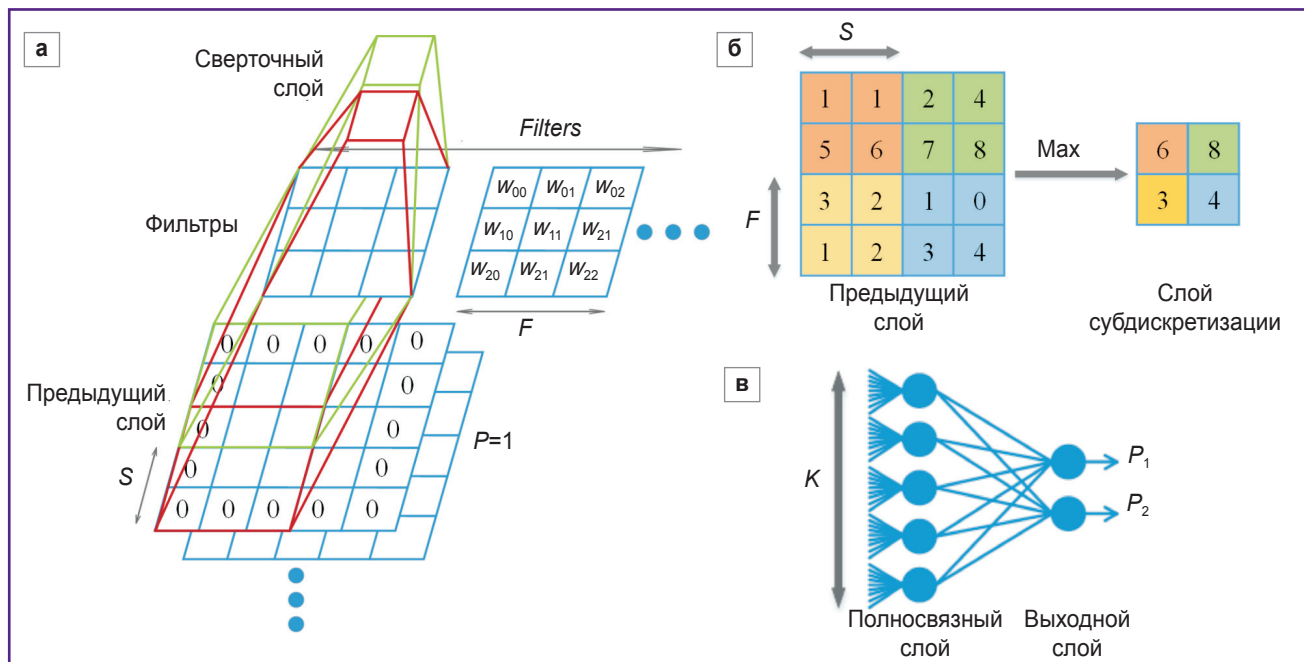


Рис. 2. Фрагменты слоев сверточной нейронной сети с указанными обозначениями а — сверточный слой; б — слой субдискретизации; в — полносвязный слой

$AF$  — функция активации нейронов — пороговая (1), сигмоидная (2) или гиперболический тангенс (3):

$$y = \max(0, x); \quad (1)$$

$$y = 1 / [1 + e^{(-x)}]; \quad (2)$$

$$y = \text{th}(x). \quad (3)$$

Уровень сети представляет собой последовательность из сверточного слоя и слоя субдискретизации. Для решения прикладных задач классификации двумерных образов зачастую используются те или иные модификации СНС, состоящей из двух уровней.

Современные результаты исследований работы таких сетей дали возможность установить следующие значения параметров, которые позволяют получать

безызбыточные модели высокого качества: количество слоев свертки в уровне —  $n \in [1; 3]$ ; количество слоев субдискретизации в уровне —  $m \in [0; 1]$ ; количество уровней —  $d \in [1; 4]$ ; количество полносвязных слоев в сети —  $k \in [0; 2]$  [21].

На рис. 2 приведены обозначения параметров сети на фрагментах сверточного слоя, слоя субдискретизации и полносвязного слоя СНС.

Однако процедура построения классификатора, использующего упомянутые алгоритмы, не формализована. Для решения большинства прикладных задач классификации с помощью СНС используются достаточно простые подходы к формированию архитектуры классификаторов [21–23]:

Таблица 1

Пример модифицированной архитектуры LeNet сверточной нейронной сети [23]

№ слоя	Тип слоя	Размерность плоскости нейронов	$P$	$S$	$F/U$	$K$	Filters	AF/Subf
1	Входной	32×32×3	—	—	—	—	—	—
2	Сверточный	14×14×6	0	2	5	—	6	Пороговая
3	Объединение	4×4×6	—	2	2	—	—	Max
4	Сверточный	2×2×6	0	1	3	—	6	Пороговая
5	Объединение	1×1×6	—	2	2	—	—	Max
6	Выходной	1×1×2	—	—	—	—	—	—

1. Перебор случайных конфигураций сети. Установка случайных параметров архитектуры и проверка качества конфигурации каждой модели на определенном наборе данных. Этот способ является самым неэффективным и реже всего может использоваться для решения прикладных задач.

2. Экспертный подход. Применение архитектуры сети, показавшей достойные результаты на похожих задачах. Таким путем идет большинство исследователей, использующих в качестве модели классификации СНС.

3. Автоматизированный подбор параметров с оптимизацией. Подбор значений параметров сети с использованием алгоритмов оптимизации, например байесовской. Подход используется очень редко в связи с большими предварительными затратами на его выполнение.

4. Построение собственной архитектуры СНС эмпирически для решения определенного класса задач. Этот пункт выбирают исследователи в области машинного обучения проектирования и моделей СНС.

Как отмечено выше, второй подход является самым распространенным, однако при его использовании одна и та же модель на разных задачах может иметь различную эффективность.

В данной статье сравнение эффективности работы моделей классификаторов на основе СНС будет производиться с одной из архитектур, представленной в работе [23]. Пример конфигурации такой архитектуры приведен в табл. 1.

### Подход к формированию эффективной архитектуры сверточной нейронной сети (СНС с эффективной архитектурой)

Задача построения архитектуры СНС сводится к «сворачиванию» размеров входного изображения к трехмерным слоям с размерами 2×2 или 1×1 таким образом, чтобы получить выходные сигналы в виде вероятности принадлежности входного изображения к одному из представленных классов. Согласно нашему подходу, архитектура сети разделяется на несколько блоков с определенными значениями характеристик, чередование которых позволяет сформировать наиболее эффективную архитектуру СНС.

Как и в случае с многослойными нейронными сетями прямого распространения, при проектировании архитектуры СНС увеличение количества слоев и связей внутри сети дает возможность строить более сложные модели, позволяющие оперировать более сложными образами. В данном случае предлагается подход к процедуре автоматического построения эффективной архитектуры сети в зависимости от размерности входного вектора и требований к качеству работы классификатора.

1. Формирование параметров входного слоя сети. В задаче классификации изображений одним из правил выступает приведение исходного изображения к виду с высотой и шириной, равными  $N$ , при этом  $N$  должно многократно делиться на 2 вплоть до однозначных чисел. Как правило, в задачах классификации изображений входной слой имеет параметр глубины  $D$ , равный количеству цветовых каналов изображения.

2. Формирование значений характеристик сверточного слоя сети.

Если предыдущий слой не является сверточным слоем, то значения  $P$  и  $S$  выбираются так, чтобы размер сверточного слоя равнялся размеру предыдущего слоя согласно формулам (4) и (5) для расчета ширины  $W$  и высоты  $H$  сверточного слоя, основываясь на размерах предыдущего слоя  $W_p, H_p$ :

$$W = (W_p - F + 2P) / S + 1; \quad (4)$$

$$H = (H_p - F + 2P) / S + 1. \quad (5)$$

Исключения:

Если  $N \leq 2$ , значение  $P=1$  или  $P=3$ .

Если предыдущий слой является сверточным слоем, то  $P=0$  или ближайшему к нулю положительному значению, чтобы размеры сверточного слоя являлись целыми числами согласно (4) и (5).

$S=1$ , если  $F \geq 5$ , или значению, близкому к 1, но таким образом, чтобы размер сверточного слоя являлся целым числом согласно формулам (4) и (5).

$F=5$  или  $F=7$  при больших размерах предыдущего слоя ( $N \geq 64$ ) для первого сверточного слоя сети. При  $N \leq 32$  для первого сверточного слоя сети —  $F=5$  или  $F=3$ , последнее чаще более эффективно.

Количество фильтров *Filters* зависит от размер-

Таблица 2

Пример эффективной архитектуры сверточной нейронной сети, построенной согласно предложенному подходу для решения задачи видеоаналитики медицинских данных

№ слоя	Тип слоя	Размерность плоскости нейронов	$P$	$S$	$F/U$	$K$	Filters	AF/Subf
1	Входной	32×32×3	—	—	—	—	—	—
2	Сверточный	32×32×24	1	1	3	—	24	Пороговая
3	Объединение	16×16×24	—	2	2	—	—	Max
4	Сверточный	16×16×32	1	1	3	—	32	Пороговая
5	Объединение	8×8×32	—	2	2	—	—	Max
6	Сверточный	8×8×32	1	1	3	—	32	Пороговая
7	Объединение	4×4×32	—	2	2	—	—	Max
8	Сверточный	4×4×32	1	1	3	—	32	Пороговая
9	Объединение	2×2×32	—	2	2	—	—	Max
10	Полносвязный	1×1×2	—	—	—	2	—	Сигмоидная
11	Выходной	1×1×2	—	—	—	—	—	—

ности предыдущего слоя, глубины расположения текущего сверточного слоя и выбранного значения  $F$ . Следуя рекомендациям выбора значения  $F$ , количество фильтров увеличивается с возрастанием глубины расположения сверточных слоев в рекомендуемом диапазоне от 16 до 96 при использовании фильтров из работы [6]. При этом, по мнению авторов [20], следует брать четное число фильтров, многократно делящееся на 2 вплоть до однозначных чисел.

3. Для выбора функции активации нейронов сверточного слоя, как показывает практика, использование пороговой функции позволяет получать лучшие результаты.

4. Выбор значений характеристик слоя субдискретизации:

$U=2$ ,  $S=U$ ,  $Subf=max$ . При  $N \leq 1$  слой субдискретизации не используется.

5. Принятие решения о формировании нового сверточного слоя основывается на следующем: если  $N \geq 3$ , то формируется другой сверточный слой — переход к пункту 2; иначе производится переход к следующему пункту.

6. Для полносвязного слоя сети следует установить  $K=C$  и сигмоидную функцию активации.

При неудовлетворительных показателях качества работы сети производится усложнение ее архитектуры по следующим правилам:

добавляется сверточный слой к одному из уровней сети начиная с последнего, наиболее глубокого уровня;

увеличивается число фильтров (увеличение производится на величину, равную 10–25% от начального количества) начиная с последнего, наиболее глубокого уровня;

добавляется новый уровень сети.

Примером СНС с эффективной архитектурой, сформированной с использованием предложенного подхода, является сеть, характеристики которой представлены в табл. 2.

Анализ использования данной архитектуры в задачах классификации медицинских изображений в сравнении с модифицированной архитектурой LeNet [23] показан в разделе экспериментов (см. табл. 1). Представленный подход к формированию эффективной архитектуры СНС учитывает опыт разработки успешных конфигураций и позволяет установить проверенные значения характеристик сети исходя из основных особенностей входных данных.

#### Подход к формированию адаптивного ансамбля сверточных нейронных сетей

На основании работ [18, 21–23], описывающих эксперименты исследователей в сфере машинного обучения с различными конфигурациями архитектур СНС, определены характеристики архитектуры, которые оказывают значительное влияние на результаты классификации. Ниже приведен список выявленных характеристик, ранжированный по степени влияния на точность решения задачи классификации.

Наиболее значимые параметры:

1) размер фильтра  $F$  определяет эффективность выделения признаков на изображении, количество хранимых параметров сети и поэтому является одной из наиболее значимых характеристик архитектуры;

2) глобальная характеристика сети (конфигурация уровня сети) определяется типами слоев, входящих в нее. В исследовании рассматриваются уровни трех типов: уровень, состоящий из последовательно расположенных сверточного слоя и слоя субдискретизации; усеченный уровень, состоящий только из сверточного слоя; расширенный уровень, состоящий

из последовательно расположенных двух сверточных слоев и одного слоя субдискретизации.

Менее значимые параметры:

1) количество используемых фильтров *Filters* в сверточном слое сети — определяет разнообразие выделяемых признаков, но также может способствовать быстрому переобучению сети. Этот параметр является важным, но диапазон его возможных значений может значительно ограничивать другие параметры сети;

2) количество уровней сети, или глубина сети — в значительной мере определяется другими параметрами сети, но оказывает существенное влияние на точность классификации для сложных образов. Усложнение архитектуры и добавление нового уровня сети позволяют улучшить качество результатов.

Остальные характеристики слоев сети являются либо однозначно заданными согласно предложенному нами подходу к формированию оптимальной архитектуры сети, как, например, характеристики слоя субдискретизации, либо вспомогательными и оказывают ограниченное влияние на получаемые результаты [24–30].

Разработка нового подхода к формированию эффективной архитектуры СНС на основе рекомендаций к выбору значений характеристик сети и ранжирования наиболее значимых из них позволяет предложить более сложный подход к построению адаптивной модели классификатора на базе СНС. Он основывается на формировании ансамбля классификаторов СНС. При этом для обеспечения устойчивости к изменению исходных данных и широкой применимости к различным классам задач классификации изображений в ансамбль входит набор сетей с отличными друг от друга наиболее значимыми факторами. Причем каждая сеть формируется с использованием таких значений параметров, при которых выбранные архитектуры будут находиться в как можно более удаленных точках пространства параметров архитектур сети согласно предложенному нами подходу к формированию единичного экземпляра эффективной архитектуры.

Формирование архитектур СНС для ансамбля производится на основе выбора отличающихся значимых характеристик сети (размер фильтров сверточного слоя, конфигурация уровня сети). Это включает определение архитектур СНС для ансамбля:

с отличающимися значениями па-

раметра  $F$  сверточного слоя сети согласно предлагаемому подходу;

с отличающимися уровнями сети (расширенным, полным и усеченным).

Из полученного набора СНС формируется ансамбль моделей, в котором результат выводится путем усреднения результатов, полученных от каждой модели в отдельности:

$$Y_{ensemble} = \frac{1}{T} \sum_{i=1}^T y_i(x). \quad (6)$$

Для задачи классификации патологий предложена модель, составленная из адаптивного ансамбля СНС, чьи конфигурации представлены в табл. 3–5.

В следующем разделе будут рассчитаны характеристики эффективности предлагаемых архитектур для решения задачи классификации патологий по данным видеокolonоскопии, а также исследованы их адаптивность и устойчивость к изменению исходных данных.

Таблица 3

Архитектура первой сверточной нейронной сети, входящей в адаптивный ансамбль

№ слоя	Тип слоя	Размерность плоскости нейронов	$P$	$S$	$F/U$	$K$	<i>Filters</i>	<i>AF/Subf</i>
1	Входной	32×32×3	—	—	—	—	—	—
2	Сверточный	32×32×16	2	1	5	—	16	Пороговая
3	Объединение	16×16×16	—	2	2	—	—	Max
4	Сверточный	16×16×20	2	1	5	—	20	Пороговая
5	Объединение	8×8×20	—	2	2	—	—	Max
6	Сверточный	8×8×20	2	1	5	—	20	Пороговая
7	Объединение	4×4×20	—	2	2	—	—	Max
8	Полносвязный	1×1×2	—	—	—	2	—	Сигмоидная
9	Выходной	1×1×2	—	—	—	—	—	—

Таблица 4

Архитектура второй сверточной нейронной сети, входящей в адаптивный ансамбль

№ слоя	Тип слоя	Размерность плоскости нейронов	$P$	$S$	$F/U$	$K$	<i>Filters</i>	<i>AF/Subf</i>
1	Входной	32×32×3	—	—	—	—	—	—
2	Сверточный	32×32×24	1	1	3	—	24	Пороговая
3	Объединение	16×16×24	—	2	2	—	—	Max
4	Сверточный	16×16×32	1	1	3	—	32	Пороговая
5	Объединение	8×8×32	—	2	2	—	—	Max
6	Сверточный	8×8×32	1	1	3	—	32	Пороговая
7	Объединение	4×4×32	—	2	2	—	—	Max
8	Сверточный	4×4×32	1	1	3	—	32	Пороговая
9	Объединение	2×2×32	—	2	2	—	—	Max
10	Полносвязный	1×1×2	—	—	—	2	—	Сигмоидная
11	Выходной	1×1×2	—	—	—	—	—	—

Таблица 5

## Архитектура третьей сверточной нейронной сети, входящей в адаптивный ансамбль

№ слоя	Тип слоя	Размерность плоскости нейронов	<i>P</i>	<i>S</i>	<i>F/U</i>	<i>K</i>	<i>Filters</i>	<i>AF/Subf</i>
1	Входной	32×32×3	—	—	—	—	—	—
2	Сверточный	32×32×24	2	1	5	—	24	Пороговая
3	Объединение	16×16×24	—	2	2	—	—	Max
4	Сверточный	14×14×32	0	1	3	—	32	Пороговая
5	Сверточный	12×12×32	0	1	3	—	32	Пороговая
6	Объединение	6×6×32	—	2	2	—	—	Max
7	Сверточный	4×4×48	0	1	3	—	48	Пороговая
8	Сверточный	2×2×48	0	1	3	—	48	Пороговая
9	Объединение	1×1×48	—	2	2	—	—	Max
10	Полносвязный	1×1×2	—	—	—	2	—	Сигмоидная
11	Выходной	1×1×2	—	—	—	—	—	—

## Эксперименты

Для оценки эффективности работы как одиночной СНС с эффективной архитектурой, так и адаптивного ансамбля сетей, а также сравнения их возможностей с существующей и активно используемой специалистами в области медицины модифицированной архитектурой LeNet [23] решена задача построения классификатора патологий по данным видеокolonоскопии. В процедуре колоноскопии используется визуальный анализ внутренней поверхности органа для локализации полипов, являющихся аномальными разрастаниями тканей над слизистой оболочкой, проведения биопсии и удаления поражений.

На рис. 3 приведены примеры регионов видео, содержащих полипы и не обладающих особенностями, однако визуально похожих на новообразования (к таким объектам относятся анатомические особенности прямой кишки, фекальные массы, световые пятна и т.д.).

При исследовании серии изображений видеоко-

лоноскопии, полученных разными эндоскопами для отличающихся групп пациентов, отмечено, что качество кадра может сильно варьироваться. В частности, на качество съема данных влияют следующие характеристики эндоскопа: угол зрения оптической системы, разрешающая способность камеры, фокусное расстояние камерной головки, величина освещенности операционного поля, тип источника освещения [11]. Существенное влияние на анализ оказывают особенности организма пациента, в частности цвет внутренних стенок прямой кишки, сосудистый рисунок, подготовка пациента к процедуре (наличие остатков фекалий, их цвет и количество), а также вероятность расфокусировки, недоэкспонирование или переэкспонирование некоторых участков, наличие жидкости на линзе эндоскопа, появление бликов на стенках прямой кишки и других событий, связанных с оборудованием [20].

Для получения моделей сетей и ансамбля сетей, обладающих большей способностью к обобщению, 40% образцов обучающей и тестирующей выборки подвергнуты модификации для деформации кадра, вызванной оптическими эффектами. Недоэкспонированные и переэкспонированные изображения получены путем гистограммных преобразований, в которых вид функции преобразования определяется адаптивно, на основании первоначальной гистограммы исходного изображения. В нашем случае дискретных цифровых изображений функция отображения яркости принимает вид так называемой таблицы отображения яркости LUT [24].

Гистограммная обработка изображения при помощи таблицы отображения яркости осуществляется с помощью преобразования:

$$Im[i, j] = LUT(Im[i, j]), \quad (7)$$

где  $Im[i, j]$  — интенсивность пикселя изображения с

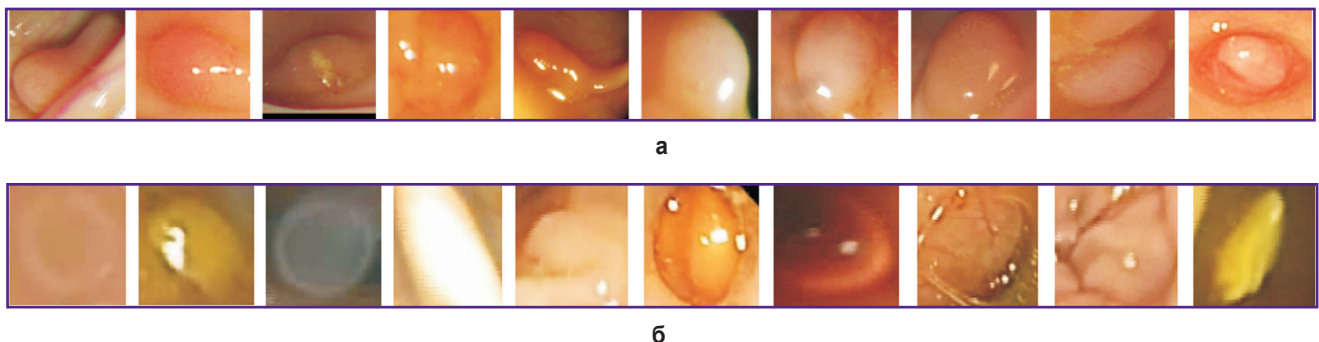


Рис. 3. Примеры областей прямой кишки, полученные с помощью видеокolonоскопии: а — образцы, содержащие полипы; б — образцы случайных структур и артефактов прямой кишки (норма)



координатами  $i, j$ , а LUT — целочисленный массив размером 256 элементов для 8-битного изображения или 65 536 элементов — для 16-битного изображения, получаемый согласно процедуре адаптивной настройки LUT.

Недоэкспонирование возникает при приближении гистограммы изображения к сглаженному виду в области ярких цветов («светов»). Переэкспонирование, соответственно, наоборот, в области темных цветов («теней»).

На рис. 4 показаны примеры преобразования исходного изображения в недоэкспонированное и пере-

экспонированное, а также гистограммы распределения интенсивности пикселей для них.

Для имитации попадания жидкости на линзу видеокамеры используется гауссово размытие (рис. 5).

Для проведения тестирования разработанного подхода к формированию СНС с эффективной архитектурой и адаптивного ансамбля СНС использован набор данных, полученных из базы SVC-ColonDB [25], и 20 коротких видеозаписей процедуры колоноскопии от клиник штата Аризона (Феникс, США).

Исследование включало в себя обучение и

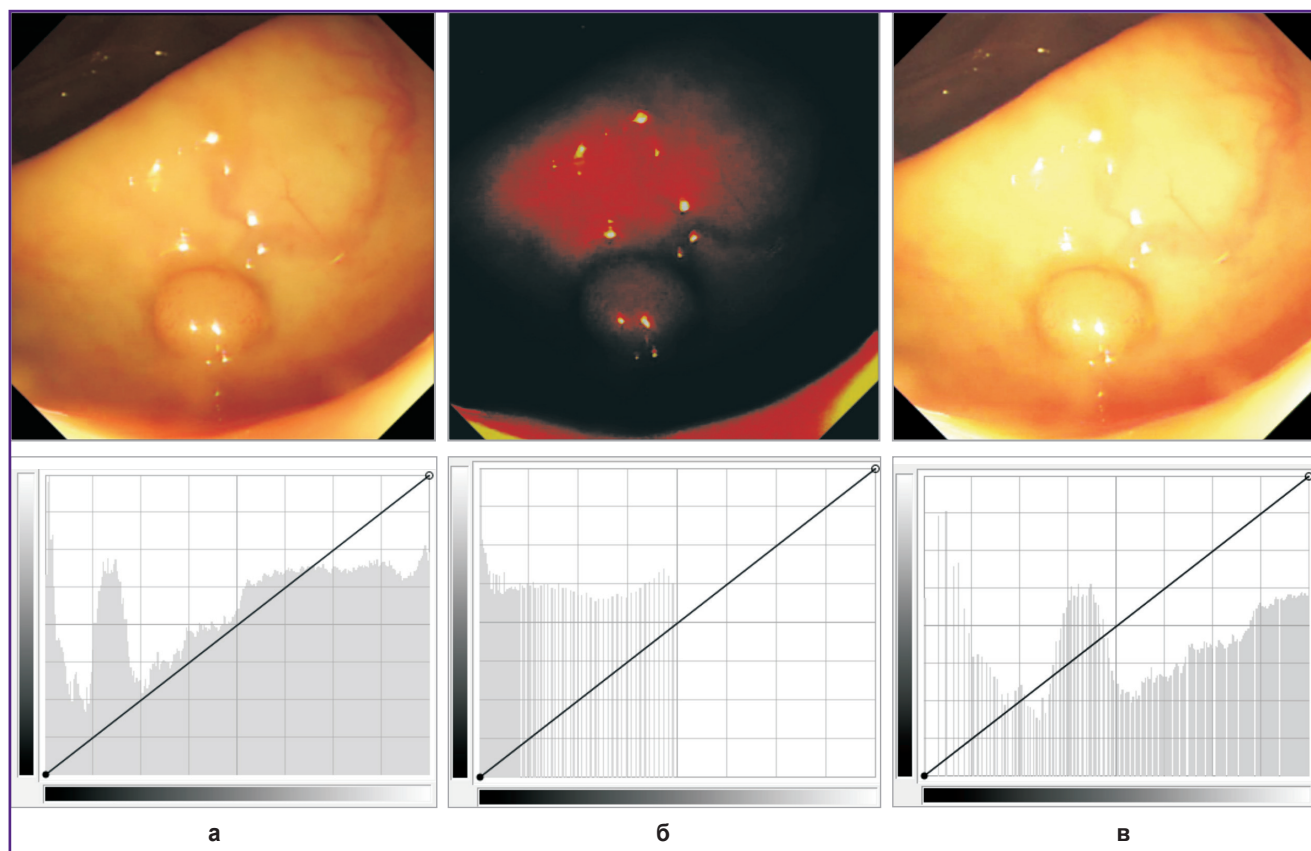


Рис. 4. Примеры преобразования исходного изображения (а) в недоэкспонированное (б) и переэкспонированное (в)

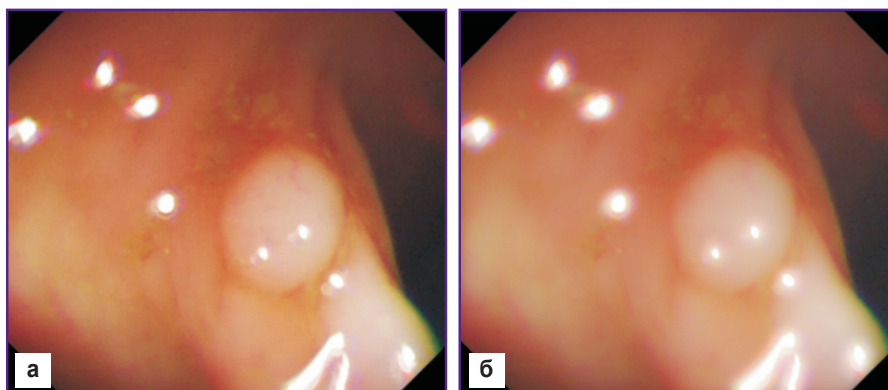


Рис. 5. Пример гауссова размытия видеокadra колоноскопии: а — исходное изображение; б — размытое изображение

сравнение работы на одних и тех же обучающих и тестирующих выборках трех классификаторов: модифицированная архитектура LeNet [23], предлагаемые нами СНС с эффективной архитектурой и адаптивный ансамбль СНС, характеристики архитектур которых представлены в табл. 1, 2 и 3–5 соответственно. Оценка характеристик чувствительности и специфичности для каждой модели позволяет ранжировать анализируемые модели по качеству решения и сделать вывод о перспективности использования предложенных алгоритмов на практике.

Весь набор исследуемых данных включал в себя выборку из 7470 образцов изображений 3735 полипов и 3735 образцов изображений случайных структур прямой кишки человека. Дополнительно подготовлена тестовая выборка из 830 образцов, из которых 415 — образцы видеозаписей полипов и 415 — случайных структур кишечника. Тестовая выборка подготовлена для оценки обобщающих свойств моделей. Использовались изображения в формате RGB размером 32×32 пикселя.

При проведении экспериментов из исходного набора данных случайным образом сформированы поднаборы обучающих выборок — из 100 и 1000 образцов из большой выборки в 7470 образцов, где половину примеров составляли изображения полипов, а другую половину — фрагменты кадров, не содержащие новообразований. Данные поднаборы необходимы для того, чтобы протестировать изменение качества работы классификатора при изменении объема обучающей выборки, а также определить устойчивость архитектур, полученных с помощью предложенных подходов, путем

тестирования сетей, обученных на поднаборах, на тестовой выборке, состоящей из 830 образцов.

### Тестирование одиночных сверточных нейронных сетей

На рис. 6 приведены ROC-кривые построенных на основе модифицированной LeNet классификаторов и предлагаемой нами СНС с эффективной архитектурой на тестовой выборке из 830 образцов, при их обучении на поднаборах обучающих выборок из 100, 1000 примеров и на полной обучающей выборке из 7470 примеров.

На представленных результатах сравнения точности классификации видно, что при тестировании сетей, обученных на выборке, состоящей из 100 образцов, предлагаемая нами архитектура имеет значительно более высокое качество. Однако при увеличении размеров обучающей выборки качество распознавания является сравнимым и только при больших объемах обучающей выборки становится снова однозначно выше, чем качество модифицированной архитектуры LeNet. Полученные результаты говорят о том, что СНС с эффективной архитектурой работает лучше на больших обучающих выборках, но обладает чувствительностью к изменению исходных данных.

### Тестирование ансамбля и одиночных сверточных нейронных сетей

Рис. 7 демонстрирует кривые ROC-анализа при тестировании результатов работы классификаторов, основанных на модифицированной архитектуре LeNet и предлагаемой нами адаптивном ансамбле СНС, на тестовой выборке из 830 образцов, при их обучении на поднаборах обучающих выборок из 100,

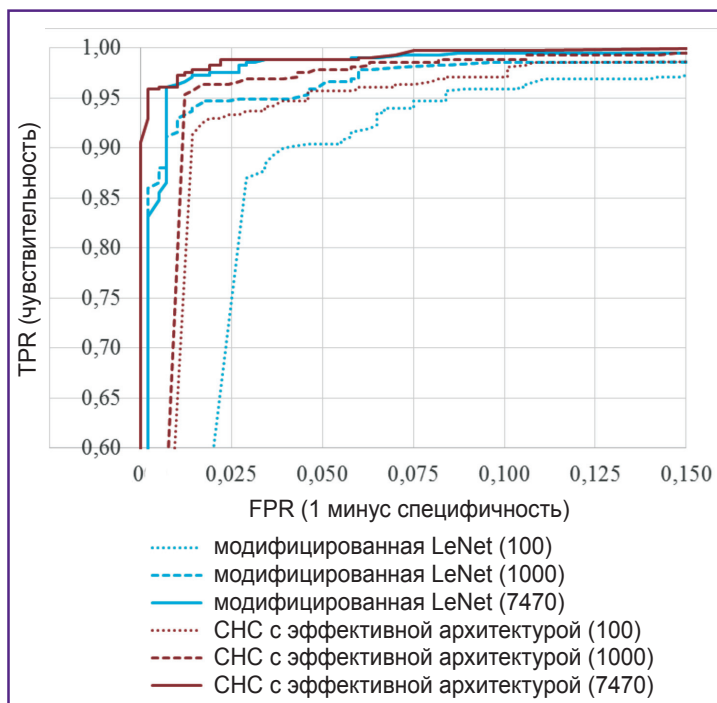
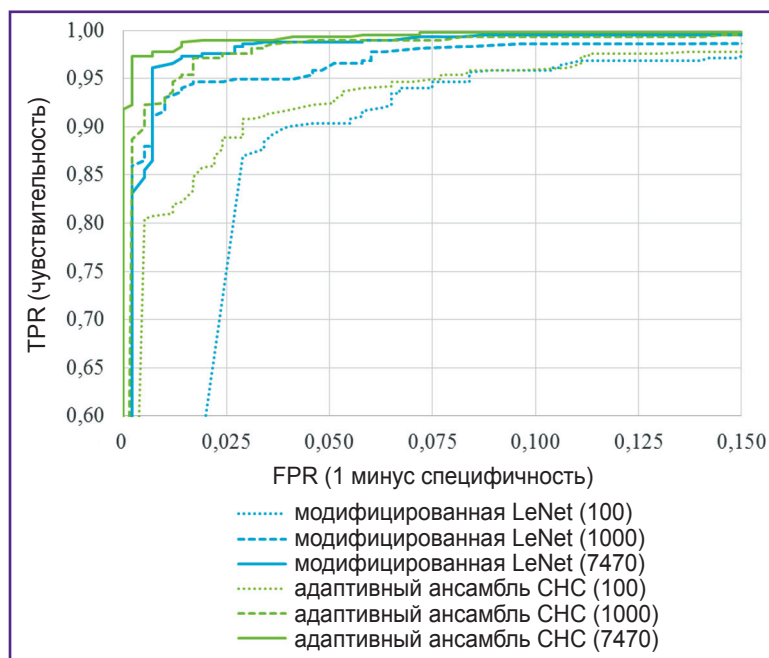


Рис. 6. ROC-кривые сравнения изменения точности классификации при изменении объемов обучающей выборки для модифицированной LeNet и предлагаемой сверточной нейронной сети (СНС) с эффективной архитектурой. Значение в скобках после названия модели обозначает количество использованных обучающих образцов

**Рис. 7. ROC-кривая сравнения изменения точности классификации при изменении объемов обучающей выборки для модифицированной LeNet и адаптивного ансамбля сверточных нейронных сетей (СНС)**

Значение в скобках после названия модели обозначает количество использованных обучающих образцов



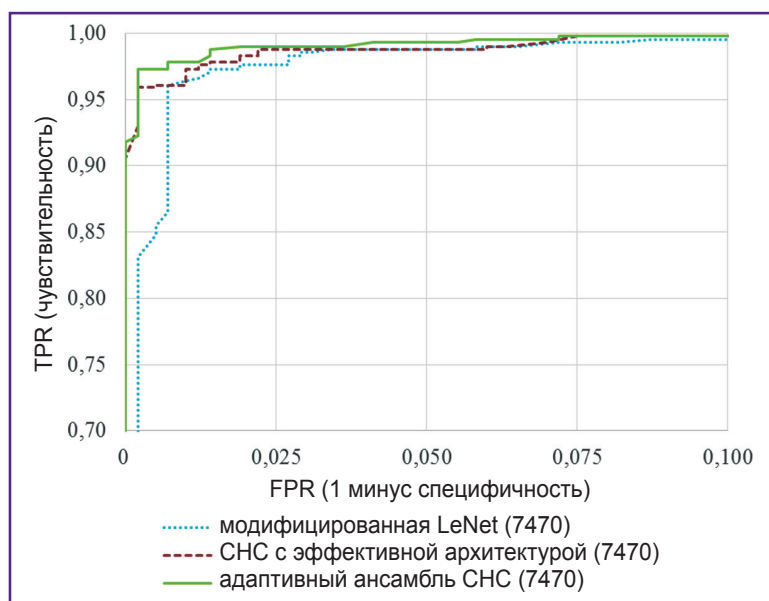
1000 примеров и на полной обучающей выборке из 7470 примеров.

Представленные результаты сравнения точности классификации, достигаемой модифицированной архитектурой LeNet и ансамблем СНС, показывают, что последний обеспечивает более высокое качество классификации на любом наборе обучающей выборки. Если сравнивать все три модели, обученные на полной выборке из 7470 примеров, то ансамбль также дает наиболее высокую точность классификации по сравнению с любой моделью одиночной нейронной сети — как модифицированной LeNet, так и СНС с эффективной архитектурой (рис. 8).

Таким образом, полученные экспериментальные данные убедительно свидетельствуют об эффективности предлагаемых подходов к построению надежных классификаторов образов как одиночных нейронных сетей, так и ансамбля сетей и перспективности их использования даже на ограниченной выборке.

## Заключение

Технология предлагаемых классификаторов патологических образований основана на анализе архитектур СНС, хорошо зарекомендовавших себя в решении практических задач медицинской диагностики. Путем анализа и обобщения особенностей архитектур успешных моделей разработан подход к формированию безызыточной сверточной нейронной сети — сверточной нейронной сети с эффектив-



**Рис. 8. Сравнение точности классификации трех моделей сверточных нейронных сетей (СНС), обученных на полной выборке из 7470 примеров**

ной архитектурой, которая позволяет эффективно решать вопросы классификации образов патологических образований по данным видеокolonоскопии. Комплексные исследования показали, что улучшить качество классификации возможно путем применения ансамбля сверточной нейронной сети с учетом зависимостей, описанных в подходе к формированию сверточной нейронной сети с эффективной архитектурой. Эксперименты демонстрируют высокое качество работы предлагаемого ансамбля даже на выборках малого размера.

**Финансирование исследования.** Научное исследование проведено в рамках программы по повышению конкурентоспособности Национального исследовательского Томского государственного университета и выполнено при поддержке гранта РФФИ 16-47-700289.

**Конфликт интересов.** Авторы декларируют отсутствие конфликта интересов.

## Литература/References

1. Varol G, Salah A.A. Efficient large-scale action recognition in videos using extreme learning machines. *Expert Syst Appl* 2015; 42(21): 8274–8282, <https://doi.org/10.1016/j.eswa.2015.06.013>.
2. Taylor F.R. *Evaluation of supervised machine learning for classifying video traffic*. Doctoral dissertation. Nova Southeastern University; 2016.
3. Li B., Meng M.Q.-H., Lau J.Y.W. Computer-aided small bowel tumor detection for capsule endoscopy. *Artif Intell Med* 2011; 52(1): 11–16, <https://doi.org/10.1016/j.artmed.2011.01.003>.
4. Wang Y., Tavanapong W., Wong J., Oh J., de Groen P.C. Part-based multiderivative edge cross-sectional profiles for polyp detection in colonoscopy. *IEEE J Biomed Health Inform* 2014; 18(4): 1379–1389, <https://doi.org/10.1109/jbhi.2013.2285230>.
5. Moon W.K., Shen Y.W., Bae M.S., Huang C.S., Chen J.H., Chang R.F. Computer-aided tumor detection based on multi-scale blob detection algorithm in automated breast ultrasound images. *IEEE Trans Med Imaging* 2013; 32(7): 1191–1200, <https://doi.org/10.1109/tmi.2012.2230403>.
6. Sudharani K., Sarma T.C., Prasad K.S. Advanced morphological technique for automatic brain tumor detection and evaluation of statistical parameters. *Procedia Technology* 2016; 24: 1374–1387, <https://doi.org/10.1016/j.protcy.2016.05.153>.
7. Goel R.M., Patel K.V., Borrow D., Anderson S. Video capsule endoscopy for the investigation of the small bowel: primary care diagnostic technology update. *Br J Gen Pract* 2014; 64(620): 154–156, <https://doi.org/10.3399/bjgp14x677680>.
8. Silva F.B., Dinis-Ribeiro M., Vieth M., Rabenstein T., Goda K., Kiesslich R., Haringsma J., Edebo A., Toth E., Soares J., Areia M., Lundell L., Marschall H.U. Endoscopic assessment and grading of Barrett's esophagus using magnification endoscopy and narrow-band imaging: accuracy and interobserver agreement of different classification systems (with videos). *Gastrointest Endosc* 2011; 73(1): 7–14, <https://doi.org/10.1016/j.gie.2010.09.023>.
9. Li B., Meng M.Q., Xu L. A comparative study of shape features for polyp detection in wireless capsule endoscopy images. *Conf Proc IEEE Eng Med Biol Soc* 2009, <https://doi.org/10.1109/iembs.2009.5334875>.
10. Li B., Fan Y., Meng M.Q.-H., Qi L. Intestinal polyp recognition in capsule endoscopy images using color and shape features. *International Conference on Robotics and Biomimetics (ROBIO)* 2009, <https://doi.org/10.1109/robio.2009.5420969>.
11. Tajbakhsh N., Gurudu S.R., Liang J. A classification-enhanced vote accumulation scheme for detecting colonic polyps. *Lecture Notes in Computer Science* 2013; p. 53–62, [https://doi.org/10.1007/978-3-642-41083-3\\_7](https://doi.org/10.1007/978-3-642-41083-3_7).
12. Park S.Y., Sargent D., Spofford I., Vosburgh K.G., A-Rahim Y. A colon video analysis framework for polyp detection. *IEEE Trans Biomed Eng* 2012; 59(5): 1408–1418, <https://doi.org/10.1109/TBME.2012.2188397>.
13. Song S.E., Seo B.K., Cho K.R., Woo O.H., Son G.S., Kim C., Cho S.B., Kwon S.S. Computer-aided detection (CAD) system for breast MRI in assessment of local tumor extent, nodal status, and multifocality of invasive breast cancers: preliminary study. *Cancer Imaging* 2015; 15(1): 1, <https://doi.org/10.1186/s40644-015-0036-2>.
14. El-Dahshan El-S.A., Mohsen H.M., Revett K., Salem A.-B.M. Computer-aided diagnosis of human brain tumor through MRI: a survey and a new algorithm. *Expert Syst Appl* 2014; 41(11): 5526–5545, <https://doi.org/10.1016/j.eswa.2014.01.021>.
15. Liu L., Tian Z., Zhang Z., Fei B. Computer-aided detection of prostate cancer with MRI: technology and applications. *Acad Radiol* 2016; 23(8): 1024–1046, <https://doi.org/10.1016/j.acra.2016.03.010>.
16. ГИЦ колопроктологии Росздрава. Отделение эндоскопической хирургии. Как происходит колоноскопия? State Scientific Center of Coloproctology of the Federal Health Service. Department of endoscopic surgery. *How is a colonoscopy?* URL: <http://www.colonoscopy.ru/patient/procedure2.htm>.
17. Углов А.С., Замятин А.В. Информационно-программный комплекс для решения задач персонализированной медицины с применением интеллектуального анализа данных. В кн.: Информационные технологии и математическое моделирование. Томск; 2017; с. 126–134. Uglov A.S., Zamyatin A.V. Informatsionno-programmnyy kompleks dlya resheniya zadach personalizirovannoy meditsiny s primeneniem intellektual'nogo analiza dannykh. V kn.: *Informatsionnye tekhnologii i matematicheskoe modelirovanie* [Information and software complex for solving problems of personalized medicine with the use of data mining. In: Information technologies and mathematical modeling]. Tomsk; 2017; p. 126–134.
18. Axyonov S., Zamyatin A., Liang J., Kostin K. Advanced pattern recognition and deep learning for colon polyp detection. In: *Distributed computer and communication networks: control, computation, communications*. Moscow; 2016; p. 27–34.
19. Аксенов С.В., Костин К.А., Джанминг Л., Замятин А.В. Использование методов Deep Learning для детектирования полипов при колоноскопии. В кн.: Информационные технологии и математическое моделирование. Томск; 2016; с. 75–79. Aksenov S.V., Kostin K.A., Jianming L., Zamyatin A.V. Ispol'zovanie metodov Deep Learning dlya detektirovaniya polipov pri kolonoskopii. V kn.: *Informatsionnye tekhnologii i matematicheskoe modelirovanie* [The use of Deep Learning methods for polyp detection during colonoscopy. In: Information technologies and mathematical modeling]. Tomsk; 2016; p. 75–79.
20. Bernal J., Sánchez J., Vilariño F. Towards automatic polyp detection with a polyp appearance model. *Pattern Recognition* 2012; 45(9): 3166–3182, <https://doi.org/10.1016/j.patcog.2012.03.002>.
21. Nibali A., He Z., Wollersheim D. Pulmonary nodule classification with deep residual networks. *Int J Comput Assist Radiol Surg* 2017; 12(10): 1799–1808, <https://doi.org/10.1007/s11548-017-1605-6>.

22. Tajbakhsh N., Gurudu S.R., Liang J. Automatic polyp detection in colonoscopy videos using an ensemble of convolutional neural networks. *IEEE 12th International Symposium on Biomedical Imaging (ISBI) 2015*, <https://doi.org/10.1109/isbi.2015.7163821>.
23. LeCun Y., Kavukcuoglu K., Farabet C. Convolutional networks and applications in vision. *Proceedings of 2010 IEEE International Symposium on Circuits and Systems 2010*, <https://doi.org/10.1109/iscas.2010.5537907>.
24. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных. Москва: ДМК Пресс; 2015. Flach P. *Mashinnoe obuchenie. Nauka i iskusstvo postroeniya algoritmov, kotorye izvlekayut znaniya iz dannykh* [Machine learning: the art and science of algorithms that make sense of data]. Moscow: DMK Press; 2015.
25. CVC colon DB. URL: <http://mv.cvc.uab.es/projects/colon-qa/cvccolondb>.
26. Park S.Y., Sargent D., Spofford I., Vosburgh K.G., A-Rahim Y. A colon video analysis framework for polyp detection. *IEEE Transactions on Biomedical Engineering* 2012; 59(5): 1408–1418, <https://doi.org/10.1109/tbme.2012.2188397>.
27. Tajbakhsh N., Gotway M.B., Liang J. Computer-aided pulmonary embolism detection using a novel vessel-aligned multi-planar image representation and convolutional neural networks. *Medical Image Computing and Computer-Assisted Intervention 2015*; p. 62–69, [https://doi.org/10.1007/978-3-319-24571-3\\_8](https://doi.org/10.1007/978-3-319-24571-3_8).
28. Zhu R., Zhang R., Xue D. Lesion detection of endoscopy images based on convolutional neural network features. *8th International Congress on Image and Signal Processing (CISP) 2015*, <https://doi.org/10.1109/cisp.2015.7407907>.
29. Kooi T., Litjens G., van Ginneken B., Gubern-Mérida A., Sánchez C.I., Mann R., den Heeten A., Karssemeijer N. Large scale deep learning for computer aided detection of mammographic lesions. *Med Image Anal* 2017; 35: 303–312, <https://doi.org/10.1016/j.media.2016.07.007>.
30. Tajbakhsh N., Gurudu S.R., Liang J. A comprehensive computer-aided polyp detection system for colonoscopy videos. *Lecture Notes in Computer Science 2015*; p. 327–338, [https://doi.org/10.1007/978-3-319-19992-4\\_25](https://doi.org/10.1007/978-3-319-19992-4_25).