

ОБЗОРЫ

УДК 004.032.26:004.932

**ГЛУБОКИЕ НЕЙРОННЫЕ СЕТИ ДЛЯ РЕШЕНИЯ ЗАДАЧИ  
РАСПОЗНАВАНИЯ ЛИЦ ПО ФОТОИЗОБРАЖЕНИЮ**

**Габдиев Ф.Ф., Сметанина О.Н., Сазонова Е.Ю.**

*ФГБОУ ВО «Уфимский государственный авиационный технический университет», Уфа,  
e-mail: faritgabdiev@mail.ru, smoljushka@mail.ru, rassadnikova\_ekaterina@mail.ru*

В статье определено место нейронных сетей в задаче распознавания образов. Приведены архитектуры нейронных сетей, такие как неокогнитрон, автокодировщик, кресцептрон, очень глубокий обучатель, глубокие сети доверия, сеть долго-краткосрочной памяти, сети Inception, описана история их развития. Авторами также представлен анализ методов обучения глубоких нейронных сетей, а именно, метод на основе правил Хэбба, метод группового учета аргументов, алгоритм обратного распространения ошибки, метод стохастического градиентного спуска. Приведены результаты анализа существующих библиотек, во многом благодаря которым в последнее время отмечен высокий рост применения описываемых технологий. Широкий спектр архитектур нейронных сетей, методов их обучения и готовых библиотек позволяет выявить наиболее приспособленные для конкретной задачи и типа данных. Представлены результаты решения задачи распознавания лиц по фотоизображению с применением глубоких нейронных сетей. Методика решения задачи распознавания лиц по фотоизображению включает пять этапов. Первые два этапа направлены на подготовку данных с использованием методов выравнивания гистограммы, сегментации цветного изображения в цветовом пространстве RGB на основе цветопередачи. Два следующих этапа – поиск «точек интереса», дальнейшее нахождение расстояний между ними и их отношениями – выполняются с целью обнаружить черты лица с использованием алгоритма ближайшего соседа. Заключительным этапом в распознавании лиц в методике на основе фотографий является процесс идентификации, когда классифицируются точки интереса, изображенные на фотографии, полученной на предыдущих этапах с использованием нейросетей. Предложенная методика для решения задачи обеспечивает положительный результат.

**Ключевые слова:** глубокие нейронные сети, распознавание образов, обучение нейронной сети, архитектура нейронной сети, машинное обучение

**DEEP NEURAL NETWORKS FOR SOLVING PERSONAL RECOGNITION TASK**

**Gabdiev F.F., Smetanina O.N., Sazonova E.Yu.**

*Federal State Budgetary Educational Institution of Higher Education Ufa State Aviation Technical University, Ufa, e-mail: faritgabdiev@mail.ru, smoljushka@mail.ru, rassadnikova\_ekaterina@mail.ru*

This article defines the place of neural networks in pattern recognition task. Neural network architectures (Neocognitron, Autoencoder, Cresceptron, History Compressor, Deep Belief Network, Long Short-Term Memory, Inception) are given. Also, the history of these architectures' development is described. Authors presented the analysis of the learning methods of deep neural networks, namely, Hebb Rule Method, Group method of data handling, The backpropagation algorithm, Stochastic gradient descent algorithm. The results of existing libraries analysis are given, in many respects thanks to which a high growth in the use of the described technologies has been noted recently. A wide range of neural network architectures, learning methods and ready-made libraries allow to identify the most suitable for a particular task and data type. The results of the face recognition problem solving by photographs using deep neural networks. The technique for solving the face recognition problem by photographs includes five steps. The first two stages are aimed at preparing data using histogram equalization methods, color image segmentation in RGB color space based on color saliency. The next two stages are the search for «points of interest», that is finding the distances between them and their relationships are performed in order to detect facial features and using the nearest neighbor algorithm. The final step in face recognition is identification process. That is, points of interest are classified as depicted in a photograph obtained in previous steps using neural networks. The proposed technique for solving the problem assures a positive result.

**Keywords:** deep neural networks, pattern recognition, neural network training, neural network architecture, machine learning

Широкое распространение информационных технологий, усовершенствование архитектур нейронных сетей (НС), методов их обучения и разнообразие их программных реализаций [1–3], рост производительности компьютеров, в том числе GPU, способствуют использованию НС для решения задач в области распознавания образов. При наличии других инструментов для решения задач распознавания [4–6], в том числе с использованием машинного

обучения, перечисленные факторы позволили глубоким нейронным сетям (ГНС) занять одну из значимых позиций за счет большей точности распознавания. В частности, ГНС хорошо себя зарекомендовали в компьютерном зрении, распознавании речи, биоинформатике и пр.

Вопросами разработки и использования ГНС для решения задач распознавания образов занимались многие российские и зарубежные специалисты [7–9].

В статье представлено современное состояние проблемы распознавания образов, определено место НС в решении задачи, описаны история возникновения и виды архитектур НС, методы обучения сетей, а также их программные реализации. Предложено решение задачи распознавания лиц по фотоизображению на основе ГНС.

*Современное состояние проблемы распознавания образов с использованием нейронных сетей*

Для этапа принятия решений в системах распознавания образов характерны такие задачи, как распознавание и классификация, математические постановки которых приведены на рис. 1.

Согласно [10] алгоритм распознавания – есть некоторая абстрактная система  $R: R = \langle A, S, P \rangle$ , где  $A$  – множество классов,  $S$  – множество признаков для классификации,  $P$  – правила принятия решения при отнесении образа к классу.

Для рассматриваемых задач в настоящее время разработано множество методов и их реализация. Лопатина А.Д. [4] рассматривает решение задачи с использованием

метода алгоритмических композиций. В публикациях Зенина А.В. [11], Туркова П.А., Красоткиной О.В. [12] приведен байесовский подход к распознаванию образов, Фомина А.Я., Каркищенко А.Н. [13] – методы статистической классификации, Симакова В.С., Луценко Е.В. [5] – методы классификации по сходству / делимости и др. Как уже отмечено ранее, одним из средств решения задачи распознавания образов выступают НС (рис. 2) [7–9; 14].

Модель искусственного нейрона Маккаллока-Питтса, предложенная в 1943 г. [15], имеет  $N$  входных бинарных величин  $x_1, \dots, x_n$ , трактуемых как импульсы, и веса  $\omega_1, \dots, \omega_n$  [8]. Связи могут быть возбуждающими (вес положительный), тормозящими (вес отрицательный). Выходной сигнал нейрона определяется как  $a = \varphi \left( \sum_{i=1}^N \omega_i x_i \right)$ , где  $\varphi$  – функция активации

преобразует суммарный импульс в выходное значение нейрона [16]. В НС нейроны соединяются между собой определенным образом, сети имеют входной и выходной слои. При наличии нескольких слоев между ними можно говорить о многослойной НС.

Распознавание	Классификация
<ul style="list-style-type: none"> <li>• Отнесение предъявляемых объектов к определенным классам с помощью применения известных правил классификации.</li> <li>• Обучение системы на множестве примеров - обучающей выборке объектов распознавания.</li> <li>• Процедура "обучение с учителем".</li> </ul>	<ul style="list-style-type: none"> <li>• Разбиение множества объектов на непересекающиеся классы по их формализованным описаниям.</li> <li>• От системы требуется способность различать объекты каким-либо способом по определенным признакам.</li> <li>• Процедура "обучение без учителя".</li> </ul>
<p>Дано: множество объектов <math>W = \{w_1, w_2, \dots, w_N\}</math>. Объекты задаются значениями признаков <math>x_i</math>, <math>i=1, \dots, N</math>.</p> <p>На всем множестве <math>W</math> существует разбиение на подмножества (классы объектов). Разбиение на классы может быть задано полностью или определяться некоторой априорной информацией о классах.</p> <p>Найти: для каждого данного объекта <math>w</math> по его описанию и априорной информации вычислить значения предикатов, каждый из которых принимает истинное значение в том и только в том случае, когда объект <math>w</math> принадлежит к классу <math>i</math>.</p>	<p>Дано: множество объектов <math>W = \{w_1, w_2, \dots, w_N\}</math>. Объекты задаются значениями признаков <math>x_i</math>, <math>i=1, \dots, N</math>.</p> <p>Задан критерий <math>K(I(w))</math>, позволяющий отличать объекты друг от друга в соответствии с определенным условием. Например, критерий может выдавать целое число, определяющее положение в некоторой ранговой системе.</p> <p>Построить: разбиение множества на классы в соответствии с заданным критерием.</p> <p>Число классов определяется по результатам выполнения классификации.</p>

Рис. 1. Задачи системы распознавания образов и их математическая постановка

Система распознавания образов (используемые методы)					
Классификации по сходству/разделимости [5]	Байесовский вывод [11, 12]	Кластеризация/ регрессия [6]	<b>НЕЙРОННЫЕ СЕТИ</b> [7-9, 14]	Статистическая классификация [10, 13]	Алгоритмические композиции [4]

Рис. 2. Место НС в системе распознавания образов

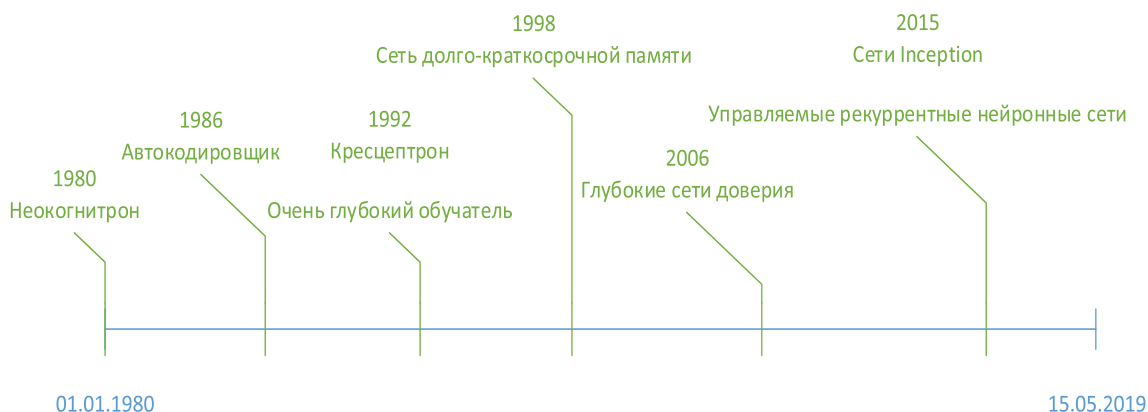


Рис. 3. Развитие архитектур НС

Среди трендов нынешнего времени в области нейронных сетей следует отметить сложность их архитектур («глубина», «ветвление слоев», «блочность», «структура»). Другими характеристиками как решаемой задачи, так и НС являются число описываемых нейронной сетью параметров, а также количество данных для обучения. С обучением также связана проблема «затухающего градиента».

Начало создания архитектур НС относят к 1980 г. Была предложена архитектура НС – неокогнитрон (рис. 3, 4). Во второй половине 80-х годов прошлого столетия Д. Баллард предложил подход к обучению НС без учителя на основе автокодировщика [17].

В 1992 г. появилась архитектура НС кресцептрон, в основу которого положен неокогнитрон. Наличие проблемы исчезающего градиента при обучении привело к новым архитектурам: очень глубокий обучатель (1992 г., сотни слоев) [18] направлен на применение предварительного обучения без учителя иерархии рекуррентных НС; НС – сеть долго-краткосрочной памяти. В 2006 г. Дж. Хинтоном и Р. Салахутдиновым была создана архитектура НС – глубокие сети доверия. Компанией Google

в 2014 г. предложена архитектура сверточных НС Inception, которые начали активно развиваться [19].

Также предложена архитектура рекуррентных НС – управляемые рекуррентные НС. К 2017 г. НС имеют от нескольких тысяч до нескольких миллионов единиц и миллионы соединений. В 2017 г. Дж. Хинтоном описаны капсульные НС, позволяющие снижать ошибку распознавания объекта в другом ракурсе на 45%, для обучения предложен алгоритм динамической маршрутизации между капсулами [7].

Процесс обучения НС направлен на настройку ее параметров посредством моделирования среды, в которую эта сеть встроена. Ниже представлены несколько методов обучения ГНС (рис. 5) [20–22]. Для указанных методов часто используются модификации, которые описаны далее при рассмотрении конкретной задачи.

#### Библиотеки глубокого машинного обучения

За последнее время создано множество библиотек для обучения НС, среди них можно отметить ApacheMXNet, Caffe, Keras, Theano и др. (рис. 6). Сравним наиболее популярные библиотеки (табл. 1) Theano, Caffe, Keras.

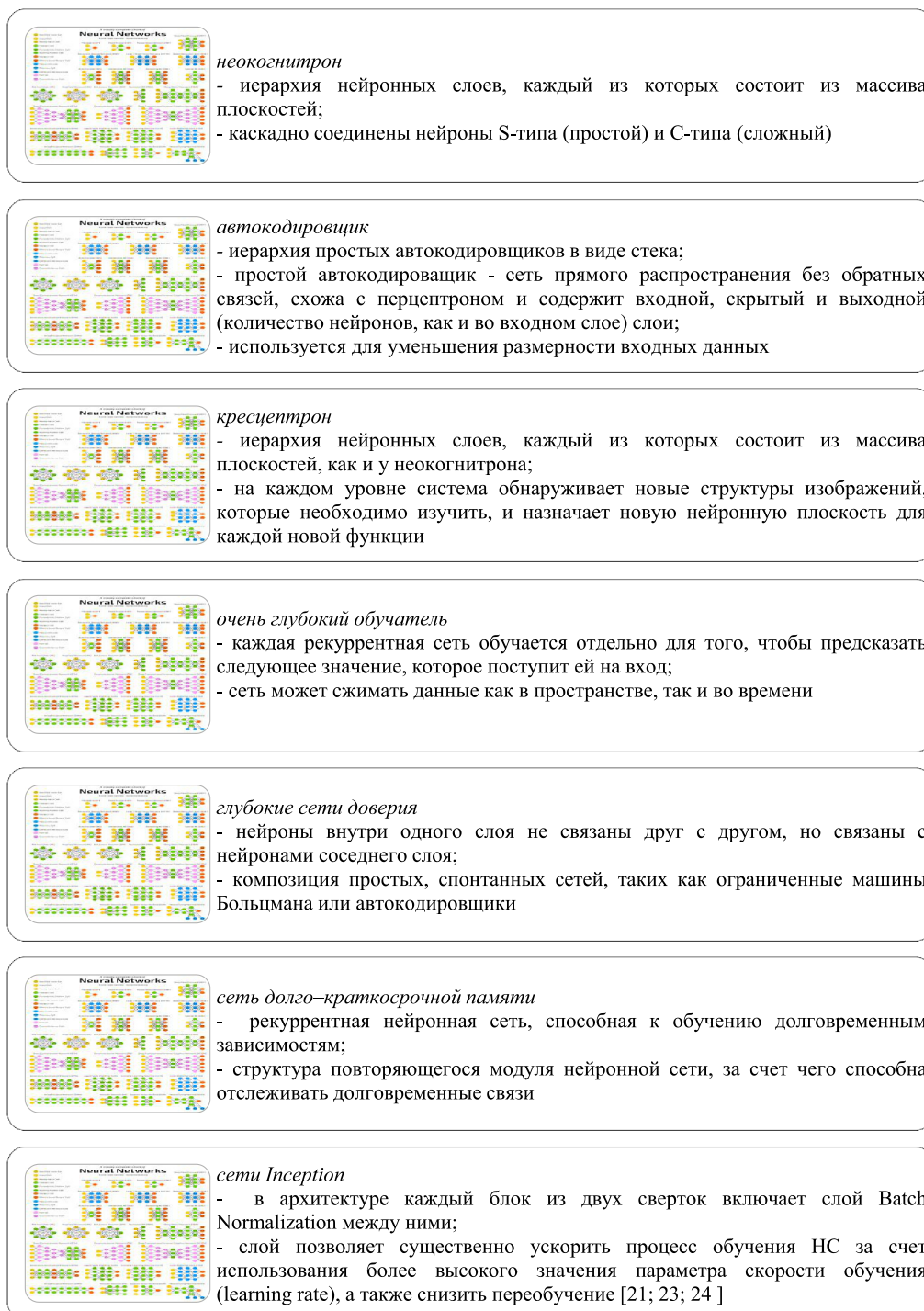


Рис. 4. Архитектуры НС

Библиотека Theano написана на Python. Вычисления в Theano выражаются синтаксисом NumPy и компилируются для эффективных параллельных вычислений. Keras, надстройка над Theano, позволяет настраивать гиперпараметры, предоставляет богатый выбор таких не-

обходимых средств, как функции потерь, алгоритмы оптимизации, инструменты визуализации и загрузки популярных датасетов.

Caffe обеспечивает сверточные нейронные сети, долго-краткосрочную память и полностью соединенные НС.

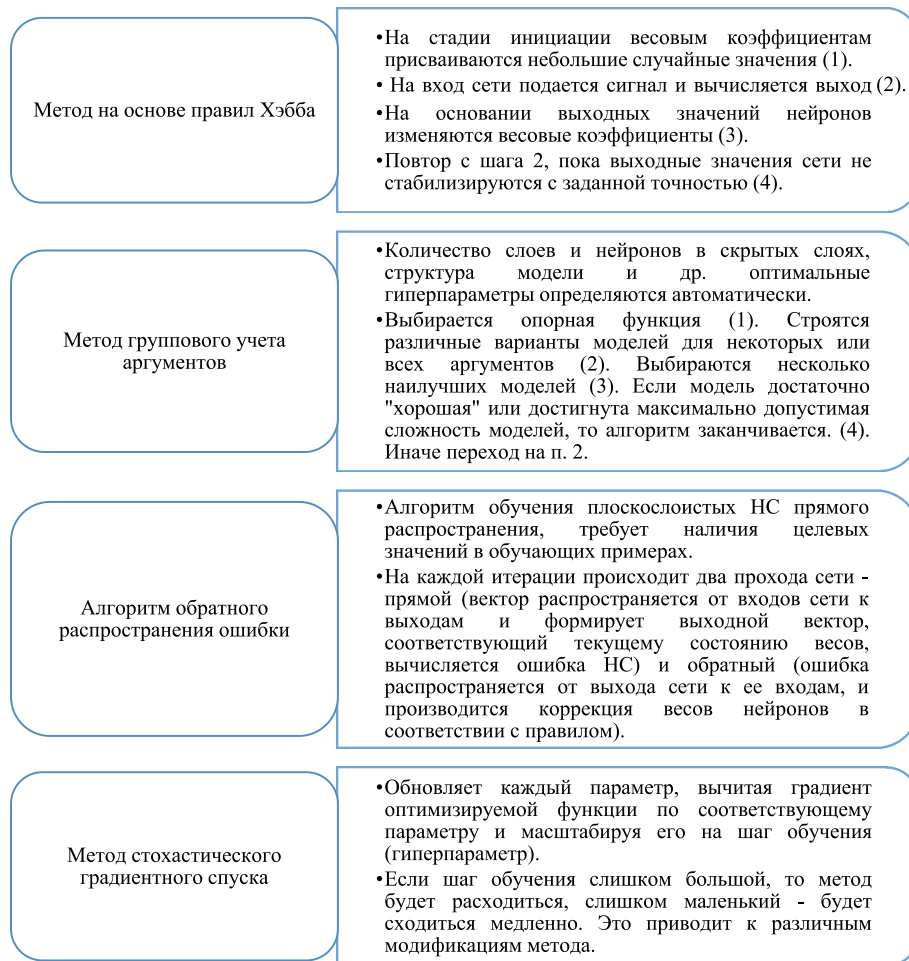


Рис. 5. Методы обучения ГНС

Таблица 1

Сравнение наиболее популярных библиотек

Критерии \ Наименование библиотеки	Theano	Caffe	Keras
Язык	Python	C++	Python
Интерфейс	Python	Python, MATLAB	Python
API	+ (Python)	+ (C++, Python)	+ (Python)
Открытый исходный код	+	+	+
ОС	межплатформенность	Linux, Mac OS X, Windows	Linux, macOS, Windows
Точность на многослойном перцептроне Румельхарта, %	97.42	98.26	Нет данных
сверточной НС, %	99.16	99.1	Нет данных
Возможность создания сверточные НС	+	+	+
глубокие сети доверия	+	-	+
ограниченные машины Больцмана	+	-	+

*Решение задачи распознавания лиц по фотоизображению*

Методика решения задачи распознавания лиц по фотоизображению включает

пять этапов. Первые два этапа направлены на подготовку данных с использованием методов выравнивания гистограммы, сегментации цветного изображения в цветовом пространстве RGB на осно-

ве цветопередачи [8; 25]. Два следующих этапа – поиск «точек интереса», дальнейшее нахождение расстояний между ними и их отношениями – выполняются с целью обнаружить черты лица с использованием алгоритма ближайшего соседа. Заключительным этапом в распознавании лиц в методике на основе фотографий является процесс идентификации, когда классифицируются точки интереса, изображенные на фотографии, полученной на предыдущих этапах с использованием НС.

Используются в эксперименте библиотеки Theano, Keras. Взаимодействие с Theano осуществлялось посредством библиотеки Keras, включающей ряд алгоритмов, в частности RMSProp, Adagrad, Adam и др. Обучение сверточной нейронной сети проведено с их использованием на базе 3 наборов по 1000 изображений различных размеров формата JPEG (табл. 2). Предложенный подход к решению проблемы в виде набора методов, реализованных на разных этапах, дает хороший результат.

Apache MXNet	<ul style="list-style-type: none"> <li>гибкая и масштабируемая среда глубокого обучения,</li> <li>поддерживает модели глубокого обучения (сверточные НС и НС с кратковременной памятью).</li> </ul>
Caffe	<ul style="list-style-type: none"> <li>поддерживает типы машинного обучения (задачи классификации и сегментации изображений),</li> <li>обеспечивает сверточные НС, долговую краткосрочную память и полносвязные НС.</li> </ul>
CNTK	<ul style="list-style-type: none"> <li>среда для проектирования и развития НС разнообразных видов,</li> <li>позволяет совместить типы моделей - глубокие, сверточные и рекуррентные НС.</li> </ul>
Fastai	<ul style="list-style-type: none"> <li>упрощает обучение быстрым и точным нейронным сетям,</li> <li>используются современные методы.</li> </ul>
Elephas	<ul style="list-style-type: none"> <li>расширение Keras,</li> <li>позволяет запускать распределенные модели Deep Learning в масштабе с помощью Spark.</li> </ul>
Keras	<ul style="list-style-type: none"> <li>позволяет описывать множество современных типов сетей в терминах предметной области,</li> <li>упрощает и ускоряет эксперименты и разработку моделей.</li> </ul>
Lasagne	<ul style="list-style-type: none"> <li>библиотека для НС, работающая поверх Theano,</li> <li>набор готовых компонентов: слоев, алгоритмов оптимизации, функций потерь, инициализаций параметров.</li> </ul>
Nolearn	<ul style="list-style-type: none"> <li>библиотека для создания и обучения НС в Theano,</li> <li>поддерживает сверточные нейронные сети (CNN).</li> </ul>
PyTorch	<ul style="list-style-type: none"> <li>библиотека машинного обучения для языка Python,</li> <li>глубокие НС на базе системы autograd.</li> </ul>
Theano	<ul style="list-style-type: none"> <li>прозрачное использование GPU, эффективное дифференцирование переменных, быстрая и стабильная оптимизация.</li> </ul>

Рис. 6. Библиотеки глубокого машинного обучения

Таблица 2

Анализ характеристик методов

Наименование метода	Точность, %	Характеристики метода
RMSProp	83	+* решение нестационарных задач
Метод адаптивного градиента (Adagrad)	81	+ отсутствие необходимости точно подбирать скорость обучения, -** сумма квадратов обновлений может увеличиваться сколько угодно, что может привести к параличу алгоритма
Adadelta	73	+* сумма градиентов заменяется на экспоненциально затухающее среднее квадратов градиентов, полученных на предыдущих шагах
Метод адаптивной оценки моментов (Adam)	76	+* способность справляться с разреженными градиентами, +* способность решать нестационарные задачи
Метод стохастического градиентного спуска (SGD)	78	+* легко реализуется, легко добавить регуляризацию, +* функция потерь и семейство алгоритмов могут быть любыми, +* возможно потоковое обучение, +* подходит для задач с большими данными, -** нет универсального набора эвристик, их нужно выбирать для конкретной задачи отдельно
*достоинства, ** недостатки методов		

Добавление этапа предварительной обработки, в частности использование метода выравнивания гистограммы, может значительно улучшить качество распознавания. В сложных наборах данных могут быть некоторые ошибки, которые могут быть обнаружены только в экспериментальных условиях.

Программная реализация подхода демонстрирует устойчивость к горизонтальному и вертикальному повороту изображения на углы 10 и 5 градусов соответственно. Некоторые проблемы возникают при нанесении различных посторонних предметов на лицо (очки, усы, синяки).

**Выводы**

Результаты анализа современного состояния проблемы распознавания образов с использованием нейронных сетей показывают, что на этапе принятия решений характерны задачи непосредственно распознавания и кластеризации. Широкий спектр архитектур нейронных сетей, методов их обучения и готовых библиотек позволяет выявить наиболее приспособленные для конкретной задачи и типа данных.

Предложенный подход к решению проблемы в виде набора методов (выравнивание гистограммы; сегментация цветного изображения в цветовом пространстве RGB на основе цветопередачи; измерение расстояний между маркерами, прикрепленными к ориентирам на лице; алгоритм ближайшего соседа), реализованных на разных этапах, дает хороший результат.

*Результаты исследований, приведенные в статье, получены в рамках выполнения грантов РФФИ 18-07-00193,*

*19-07-00709 и государственного задания № FEUE-2020-0007.*

**Список литературы**

1. Базанов П.В., Джосан О.В. Выделение информативных признаков на изображении лица в задаче идентификации человека // Сибирский журнал вычислительной математики. 2006. Т. 9. № 3. С. 207–214.
2. Чабан Л.Н. Методы и алгоритмы распознавания образов в автоматизированном дешифрировании данных дистанционного зондирования: учебное пособие. М.: МИИГА-ИК, 2016. 94 с.
3. Seide F., Agarwal A. CNTK: Microsoft's Open-Source Deep-Learning Toolkit. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16) (San Francisco, California, USA, August 13–17, 2016). 2016. P. 2135–2135.
4. Лопатина А.Д. Выделение области лица с помощью комбинации методов цветовой и яркостной сегментации // Вестник УГАТУ. 2009. Т. 13. № 2 (35). С. 180–187.
5. Симаков В.С., Луценко Е.В. Адаптивное управление сложными системами на основе теории распознавания образов. Краснодар: Техн. ун-т Кубан. гос. технол. ун-та, 1999. 318 с.
6. Таскин А.С., Миркес Е.М. Линейная регрессия с кластеризацией по признаку на данных с действительными величинами // Вестник Сибирского государственного аэрокосмического университета имени академика М.Ф. Решетнева. Математика, механика, информатика. 2012. № 3 (43). С. 71–76.
7. Головинов А.О. Анализ существующих решений в области распознавания образов (на примере капсульных сетей) // Молодой исследователь Дона. 2018. № 3 (12). С. 37–39.
8. Gabdiev F.F., Yusupova N.I., Smetanina O.N., Sazonova E.Yu. Models and Methods for Solving Face Recognition Problem by Photos. Computer Science and Information Technologies (CSIT 2019): proceedings of the 21st International Workshop. 2019. [Electronic resource]. URL: <https://www.atlantis-press.com/proceedings/csit-19/125927871> (date of access: 12.04.2020).
9. Ghorpade S., Ghorpade J., Mantri Sh. Pattern recognition using neural networks. International Journal of Computer Science & Information Technology (IJCSIT). 2010. Vol. 2. No. 6. P. 92–98. [Electronic resource]. URL: <https://www.researchgate.net/publication/312511111>

searchgate.net/publication/49595577\_Pattern\_Recognition\_Using\_Neural\_Networks (date of access: 14.04.2020).

10. Фомин Я.А. Распознавание образов: теория и приложения. М.: ФАЗИС, 2012. 429 с.

11. Зенин А.В. Анализ методов распознавания образов // Молодой ученый. 2017. № 16. С. 125–130.

12. Турков П.А., Красоткина О.В. Байесовский подход к задаче обучения распознаванию образов в нестационарной генеральной совокупности // Приборы и методы экспериментальной ядерной физики. Электроника и автоматика экспериментальных установок: труды 2-й Международной молодежной научной школы. 2011. [Электронный ресурс]. URL: <https://docplayer.ru/45679251-Bayesovskiy-podhod-k-zadache-obucheniya-raspoznavaniyu-obrazov-v-nestacionarnoy-generalnoy-sovokupnosti.html> (дата обращения: 12.04.2020).

13. Каркищенко А.Н., Гречухин И.А. Статистическое распознавание лиц по геометрии характерных точек для систем транспортной безопасности // Информационные технологии в управлении. Управление большими системами. 2020. № 38. С. 65–77.

14. Еремеев Е.А. Распознавание образов в экспертных системах принятия решений // Научно-технический вестник информационных технологий, механики и оптики. 2019. Т. 19. № 4. С. 704–713.

15. Bergstra J., Breuleux O., Bastien F., et al. Theano: a CPU and GPU Math Expression Compiler. Proceedings of the Python for Scientific Computing Conference (SciPy) (Austin, TX, USA, June 28 – July 3, 2010). 2010. P. 3–10.

16. Созыкин А.В. Обзор методов обучения глубоких нейронных сетей // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2017. Т. 6. № 3. С. 28–59.

17. Ballard D.H. Modular Learning in Neural Networks. Proceedings of the Sixth National Conference on Artificial In-

telligence (Seattle, Washington, USA, July 13–17, 1987). 1987. Vol. 1. P. 279–284.

18. Schmidhuber J. Learning Complex, Extended Sequences Using the Principle of History Compression. Neural Computation. 1992. Vol. 4. No. 2. P. 234–242.

19. Jia Y., Shelhamer E., Donahue J., et al. Caffe: Convolutional Architecture for Fast Feature Embedding. Proceedings of the 22nd ACM International Conference on Multimedia (Orlando, FL, USA, November 03–07, 2014). 2014. P. 675–678. DOI: 10.1145/2647868.2654889.

20. Chollet. F., et al. Keras. 2015. [Electronic resource]. URL: <https://github.com/fchollet/keras> (date of access: 12.04.2020).

21. PaddlePaddle: PArallel Distributed Deep LEarning. [Electronic resource]. URL: <http://www.paddlepaddle.org/> (date of access: 12.04.2020).

22. Chen T., Li M., Li Y. MXNet: A Flexible and Efficient Machine Learning Library for Heterogeneous Distributed Systems. [Electronic resource]. URL: <https://arxiv.org/abs/1512.01274> (date of access: 12.04.2020).

23. Abadi M., Agarwal A., Barham P. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16) (Savannah, GA, USA, November 2–4, 2016). 2016. P. 265–283.

24. Collobert R., Kavukcuoglu K., Farabet C. Torch7: a Matlab-like Environment for Machine Learning. BigLearn, NIPS Workshop (Granada, Spain, December 12–17, 2011).

25. Друки А.А. Алгоритмы выделения лиц на статических RGB изображениях и в видеопотоке // Известия Томского политехнического университета. Управление, вычислительная техника и информатика. 2012. Т. 320. № 5. С. 65–69.