

## Раздел V. Техническое зрение

УДК 004.932.2

DOI 10.23683/2311-3103-2019-1-271-280

**В.А. Бондаренко, Г.Э. Каплинский, В.А. Павлова, В.А. Тупиков, Н.Г. Холод**

### **НЕЙРОСЕТЕВЫЕ АЛГОРИТМЫ В ЗАДАЧАХ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ И СОПРОВОЖДЕНИЯ ОБЪЕКТОВ ДЛЯ ВСТРАИВАЕМЫХ ВЫЧИСЛИТЕЛЕЙ ОПТИКО-ЭЛЕКТРОННЫХ СИСТЕМ**

*Целью данной работы является повышение эффективности систем автоматического сопровождения и распознавания локальных детерминированных объектов в сложных условиях наблюдения. В задачах автоматического распознавания объектов зачастую невозможно предсказать ракурс наблюдения объекта и, соответственно, невозможно заранее подготовить эталоны для текущего ракурса. В подобных условиях наилучшим образом показали себя алгоритмы обработки изображений, основанные на применении глубоких искусственных нейронных сетей (ИНС) сверточного типа. Новизна данной работы состоит в разработке оригинального алгоритма распознавания объектов с применением ИНС. Разработанный алгоритм относится к группе алгоритмов машинного обучения с учителем. Особенностью алгоритмов данной группы является их хорошая масштабируемость на другие образцы изображений объекта распознавания, отсутствовавшие в обучающей выборке. Типовой процесс разработки ИНС можно разделить на три основных этапа: исследование задачи и выбор архитектуры, обучение и применение ИНС. На первом этапе осуществляется анализ предметной области, особенностей изображений типовых объектов применения, определяется количество основных групп (классов) к которым необходимо относить распознаваемые объекты, выполняется подготовка обучающей выборки (набора эталонных изображений типовых объектов каждого класса) определенного размера и выбор подходящей архитектуры ИНС. Результатом выполнения первого этапа разработки ИНС является набор данных для выполнения процесса обучения. На следующем этапе осуществляется итерационный процесс обучения ИНС с возможной корректировкой её архитектуры в зависимости от анализа оценочных метрик, определяемых в процессе обучения. Обучение выполняется в лабораторных условиях на вычислителях, обладающих вычислительной мощностью. На этапе применения ИНС выполняется процесс распознавания обученной ИНС, также называемый прямым проходом. Данный этап выполняется на встраиваемом вычислителе. Ключевым отличием разработанной ИНС от широкого семейства известных аналогичных алгоритмов является его изначальная ориентированность на применение в составе встраиваемых вычислительных комплексов, имеющих серьезные ограничения на массу, габариты, энергопотребление и другие характеристики. Это достигается путем применения целочисленной бинарной арифметики, что позволяет осуществить ее эффективную реализацию с использованием программируемых логических интегральных схем (ПЛИС). Оценка эффективности разработанной ИНС осуществлялась с применением комплексной математической модели программно-аппаратных комплексов оптико-электронных систем (ОЭС). После чего были проведены испытания в составе специального программного обеспечения ОЭС беспилотного летательного аппарата, показавшие высокую эффективность разработанной ИНС.*

*Обработка изображений; распознавание образов; глубокая нейронная сеть; оптико-электронная система.*

V.A. Bondarenko, G.E. Kaplinskiy, V.A. Pavlova, V.A. Tupikov, N.G. Holod

## NEURAL NETWORK ALGORITHMS IN THE PROBLEMS OF AUTOMATIC RECOGNITION AND TRACKING OF OBJECTS FOR OPTICAL-ELECTRONIC SYSTEMS EMBEDDED PLATFORMS

*The goal of this paper is to improve the efficiency of automatic tracking and recognition systems of local objects in difficult observation conditions. In automatic object recognition tasks, it is often difficult to predict object's angle of observation, and therefore impossible to prepare required etalons. In such conditions the best performance is shown by the algorithms based on deep convolutional neural networks (CNN). The novelty of this work is in the development of original object recognition algorithm using CNN. This algorithm belongs to the group of supervised machine learning algorithm. The main feature of algorithms of this group is their good scalability on new object image samples that not represented in a training set. Typical CNN development process can be divided in three main stages: researching and architecture selection, training and CNN application. On first stage the research of application field is performed, typical objects images analysis is conducted, typical object class amount is determined, training set preparation, and CNN architecture selection. The result of the first stage of development is a defined CNN architecture and prepared training set. On the next stage iterative CNN training process is performed CNN architecture correction possibility depending on training metrics analysis. The training is conducted in laboratory conditions on powerful workstations. On CNN application stage the recognition procedure by train CNN is performed, also known as forward propagation. This stage is being performed on embedded systems. The key difference of developed CNN from wide group of similar known algorithms is its fundamental orientation usage in embedded systems, which are have strict restrictions on their mass, dimensions, energy consumption, etc. This is achieved by using of integer binary arithmetic, which allows its efficient FPGA implementation. The efficiency evaluation of developed CNN has been conducted with the help of complex mathematical model of hardware and software complex of electro-optical systems (EOS). The developed algorithm was tested as part of special purpose software of unmanned aerial vehicle EOS and demonstrated high practical efficiency.*

*Image processing; object recognition; deep neural network; electro-optical system.*

**Введение.** Глубокие нейронные сети показывают хорошие результаты во многих областях машинного обучения. В задачах компьютерного зрения и, в частности, распознавания образов, сверточные нейронные сети показывают лучшие результаты, по сравнению с другими моделями. При этом, с развитием данного направления сложность нейросетевых моделей постоянно увеличивается, поэтому, для обучения и применения обычно используется высокопроизводительные вычислительные системы.

Однако, для применения во встраиваемых вычислительных системах существуют жесткие требования по вычислительной сложности алгоритмов при сохранении высоких требований к их эффективности. Графические процессоры, используемые как вычислительные устройства, имеют свои недостатки, такие как большое потребление энергии, требование охлаждения. Всё это затрудняет использование графических процессоров в качестве вычислителей в автономных беспилотных и безэкипажных системах.

Одной из возможных альтернатив графическим процессорам является использование программируемых логических интегральных схем (ПЛИС). Однако вещественная арифметика на таких устройствах работает существенно медленнее, чем целочисленная, поскольку логические блоки, реализующие сложение и умножение вещественных чисел, состоят из значительно большего количества элементарных операций. В связи с этим возникает задача обучения нейронной сети с целочисленными весами и соответствующей арифметикой.

В рамках данной работы рассмотрены некоторые существующие способы бинаризации глубоких нейронных сетей, предложен способ дискретизации весов нейронной сети, основанный на максимизации логарифма правдоподобия весов,

произведено сравнение и анализ результатов работы различных методов бинаризации нейронных сетей для задачи классификации изображений, полученных с применением комплексной математической модели программно-аппаратных комплексов оптико-электронных систем, а также с действующего опытного образца оптико-электронной системы БЛА.

Для демонстрации подхода была сформулирована следующая задача классификации: имеется выборка из 70 тысяч изображений размером 48x48 пикселей в градациях серого, которые были получены полуавтоматической обработкой более 30 видеозаписей. В качестве типовых принимаются объекты следующих четырех классов: грузовики, строения, дороги и фон. Классы являются несбалансированными, фоновые изображения около 2/3 объема всей выборки. Примеры изображений из обучающей выборки представлены на рис. 1.

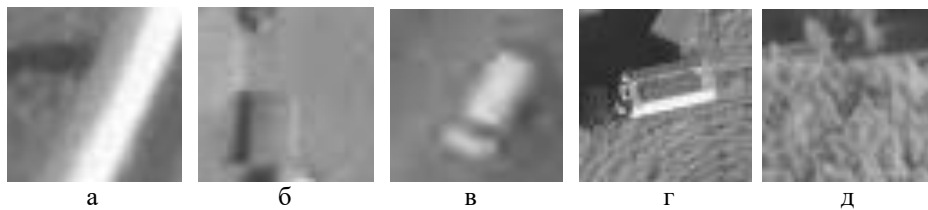


Рис. 1. Примеры изображений из обучающей выборки: а – помеха; б, д – фон; в, г – объект интереса

**Постановка задачи.** Задача заключается в обучении алгоритма предсказания класса объекта по его изображению. При этом, множество рассматриваемых алгоритмов ограничено нейронными сетями, у которых веса и активации в слое могут принимать только два возможных значения. На активации последнего слоя данные ограничения не накладываются, так как нейронная сеть предсказывает вероятность принадлежности изображения разным классам. Смысл данных ограничений для нейронной сети заключается в том, что они позволяют значительно оптимизировать прямой проход сети с помощью логических операций. Обучающая выборка была разделена на две части в пропорциях 6:1, большая часть использовалась для обучения сети, меньшая – для валидации.

Существует два основных способа бинаризации, которые могут использоваться в рамках данной задачи: детерминистический и стохастический. Детерминистическая бинаризация осуществляется простым взятием знака весов [3]. Стохастическая бинаризация предполагает конвертацию вещественных значений весов в вероятность равенства веса определенному значению и последующее приведение весов согласно полученным вероятностям. Для преобразования вещественных весов в вероятность можно использовать любую функцию  $f: \mathbb{R} \rightarrow [0; 1]$ , например логистическую функцию  $f(x) = \frac{1}{1+e^{-x}}$ . Для того, чтобы избежать неограниченного роста весов по модулю, веса обрезаются так, чтобы они лежали в отрезке  $[-1; 1]$ .

Так как функция знака не является дифференцируемой, то для вычисления градиента используется straight-through estimator [6], суть которого состоит в замене градиента знака на градиент некоторой дифференцируемой функции. Бинаризованное значение весов используется при прямом проходе сети и при вычислении градиентов вместо вещественных значений весов, однако обновляются именно вещественные веса. Таким образом, дискретизация создает случайный шум, который добавляется к значениям весов и обеспечивает также регуляризирующий эффект. Описанный подход можно расширить и применить бинаризацию не только к весам нейронной сети, но и к активациям.

Такая сеть получила название BinaryNet [5]. Важной частью обучения данной сети является так называемая пакетная нормализация [4], которая заключается в вычитании среднего значения активаций внутри пакета и делении на корень из дисперсии. Данная процедура позволяет контролировать долю вещественных активаций, принимающих положительные и отрицательные значения, что в свою очередь дает возможность корректно использовать операцию взятия знака для бинаризации.

Существует еще один способ детерминистической бинаризации сети, который позволяет обучать как сети с бинарными весами, так и сети с бинаризованными весами и активациями: XNOR-Net [7]. Для бинаризованных весов и активаций сети используется представление в виде произведения бинарной величины и нормировочной константы. Как и в вышеописанных алгоритмах, хранятся и обновляются вещественные веса, на которые не накладывается никаких ограничений, а в прямом проходе используются бинаризованные веса, которые вычисляются путем минимизации суммы квадратов отклонения от вещественных весов. Процесс вычисления бинаризованных весов не представим в виде дифференцируемой функции, поэтому, как и ранее градиент вычисляется с помощью *straight-through estimator*.

Задача бинаризации нейронной сети перекликается с задачей компрессии нейронных сетей с целью уменьшения памяти, необходимой для хранения весов. Долгое время лучшие результаты по компрессии нейронных сетей показывали методы, основанные на кластеризации весов [2]. Они как правило требовали многократных процедур дообучения сети после каждой кластеризации, что значительно увеличивает требуемое для обучения сети время. В связи с этим, можно отметить подход *soft weight sharing* [1], который производит сжатие весов сети за один цикл обучения, в связи с чем, процесс обучения сети существенно ускоряется.

В рамках данного подхода предполагается, что веса нейронной сети априорно распределены согласно смеси гауссовских распределений, параметры этой смеси (математические ожидания, дисперсии и весовые коэффициенты компонентов отдельных Гауссиан) заранее неизвестны. Так как предполагается, что веса примут одно из набора возможных значений, то в качестве апостериорного распределения используется дельта-функция Дирака. Для достижения дискретизации весов, к функции потерь добавляется регуляризирующее слагаемое в виде дивергенции Кульбака-Лейбнера между априорным и апостериорными распределениями. Итоговый функционал оптимизируется не только по весам нейронной сети, но также и по неизвестным параметрам смеси Гауссиан априорного распределения. Для достижения разреженности весов сети, в смесь добавляют Гауссиану с фиксированным нулевым математическим ожиданием и фиксированным весовым коэффициентом, который определяет степень разреженности сети.

После обучения сети происходит этап постобработки весов: сильно пересекающиеся Гауссианы объединяются в одну и все значения весов заменяются на математическое ожидание соответствующей Гауссианы.

Для описания предлагаемого способа дискретизации, рассмотрим возможность применения подхода *soft weight sharing* для бинаризации весов нейронной сети. В данном случае априорным распределением будет являться смесь из двух Гауссиан, симметричных относительно нуля. Тогда априорное распределение для весов сети можно записать в виде:

$$p(w) = \frac{1}{2\pi\sigma} \left( \exp\left(\frac{-(w - \gamma)^2}{2\sigma^2}\right) + \exp\left(\frac{-(w + \gamma)^2}{2\sigma^2}\right) \right),$$

где  $\gamma$  – нормирующий коэффициент,  $\sigma$  – СКО Гауссиан.

В процессе обучения используется следующая функция потерь:

$$L(w, d, \sigma_i, \pi_i) = -\log p(y | X, w) - \tau \log p(w).$$

Обучение начинается с  $\tau = 0$ , и с каждой эпохой  $\tau$  линейно увеличивается, пока не достигнет значения  $\tau = 1$ . В процессе оптимизации параметры  $\gamma$  и  $\sigma$  могут принимать такие значения, что Гауссианы будут сильно пересекаться, а это недопустимо в рамках данной задачи. Для недопущения подобных значений параметров, применяются априорные ограничения на эти параметры в виде плотности гамма-распределения. В ходе постобработки веса заменяются на  $+\gamma$  или на  $-\gamma$ , в зависимости от того, в какой кластер попал вес.

Описанный выше способ позволяет обучать сеть с бинарными весами, однако для получения бинарных активаций данное решение не подходит по нескольким причинам. Во-первых, требуется значительный объем дополнительной памяти для обучения сети. Во-вторых, не удастся достичь приемлемой точности распознавания. Поэтому было принято решение использовать следующий подход для сжатия активаций: а именно пересчитать активации  $c_i$  по формуле

$$c_i = \sum_j \frac{|c_j|}{n} \text{sign}(c_i),$$

что позволяет ограничить возможные значения активаций множеством  $\{-1, 1\}$ .

В рамках данной работы реализован следующий эксперимент: проведено обучение бинаризованной сети разными способами с архитектурой, подобранной с помощью Grid Search [1–10]. Архитектура разработанной искусственной нейронной сети представлена на рис. 2.

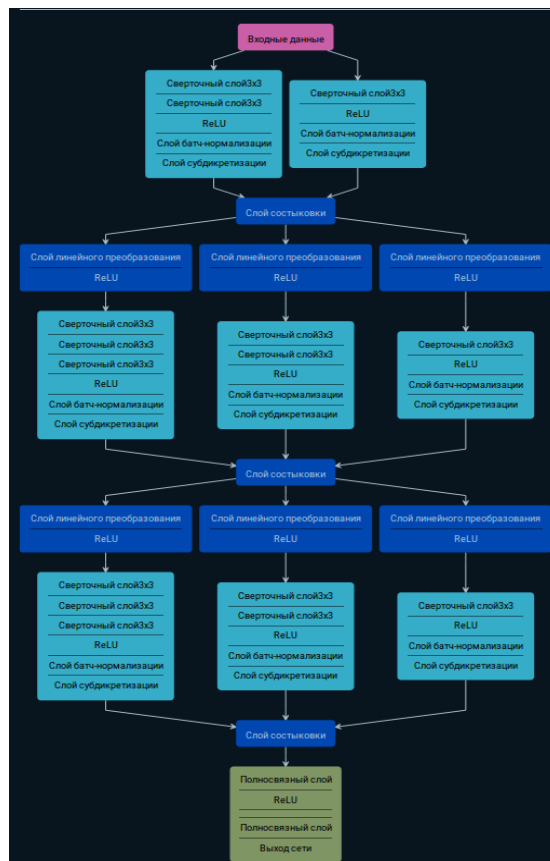


Рис. 2. Архитектура применяемой модели искусственной нейронной сети

Для того, чтобы результаты были стабильными, для каждой архитектуры обучение было проведено несколько раз, в качестве результатов применялись среднее значение и дисперсия функции потерь. Исследовались следующие способы бинаризации: BinaryNet, BinaryConnect, XnorNet и предложенный подход с использованием soft weight sharing с применением сжатия активаций и без сжатия. В качестве базовой модели для сравнения использовалась сеть без бинаризации с  $l_2$  регуляризацией [9].

В качестве функции потерь в части классификации использовалась комбинированная функция потерь, которая включает в себя бинарную перекрестную энтропию (binary cross-entropy) по классам объект/не объект и многоклассовую кросс-энтропию (categorical cross-entropy) по 4 классам. За вероятностью того, что на изображении находится объект, принимается максимум из вероятностей того, что изображение принадлежит классу из группы объектов. Так как обучающая выборка является несбалансированной, то для балансировки классов в функции потерь используется весовые коэффициенты для классов-объектов.

Результаты экспериментов представлены в табл. 1. Как видно, предложенный подход помогает достичь лучших результатов, по сравнению с другими методами. Кроме того, можно заметить, что большинство сетей с бинаризованными весами или активациями достигают лучших результатов, чем без ограничений, что подтверждает, наличие регуляризационного эффекта от бинаризации.

Таблица 1

**Результаты экспериментов с различными способами бинаризации весов**

Способ бинаризации	Mean $L(w, d, \sigma_i, \pi_i)$	Std $L(w, d, \sigma_i, \pi_i)$
Baseline	1.15	0.05
Binaryconnect	1.18	0.07
BinaryNet	1.07	0.05
Xnornet	1.04	0.04
Soft Weight Sharing based	1.01	0.03
Soft Weight Sharing based + activ compress	1.03	0.05

В качестве программный пакет для обучения сетей был выбран keras, так как он имеет более гибкую структуру чем caffe [8] и позволяет разработчику легко задавать свою функцию потерь для сети. В качестве метода оптимизации был выбран алгоритм ADAM [9].

Имея сеть с бинарными весами и активациями, можно применить некоторый набор оптимизаций, которые позволят значительно ускорить прямой проход сети. Наиболее затратными по времени слоями сверточной нейронной сети являются сверточные и полносвязные слои. Суть этих слоев сводится к множественным операциям вида multiply-accumulate. Именно на выполнение таких операций тратится больше всего ресурсов. Для бинарной сети данную операцию можно заменить комбинацией логического оператора «отрицание исключающего или» (XNOR) и операции подсчета количества единичных бит числа (bitpopcount). Данные операции выполняются намного быстрее, чем операции сложения и умножения вещественных чисел.

Для первого слоя сети такая схема не работает, так как входные данные сети не являются бинарными, а представляют собой восьмибитные целые числа. Для этого слоя ускорение достигается за счет использование только операций сложения

ния и вычитания и устранения операций умножения. Возможность такой замены операции обуславливается схожестью таблицы истинности оператора XNOR с таблицей умножения чисел  $\{+1, -1\}$ , достаточно кодировать число  $-1$  как  $0$ . Слои активации и нормализации в данной сети нацелены на то, чтобы промежуточные активации тоже оставались бинарными, поэтому главной операцией, которая используется в данных слоях, является операция сравнения с заданным пороговым значением.

Слой субдискретизации содержит только операции поиска максимума нескольких чисел, поэтому в случае, когда данные числа принимают только два возможных значения, операцию взятия максимума двух чисел можно привести к оператору логического ИЛИ.

**Заключение.** Разработанная ИНС была использована в составе аппаратно-программного комплекса экспериментального образца опико-электронной системы беспилотного летательного аппарата. Результаты работы алгоритма представлены на рис. 3–5. Зелёным прямоугольником выделены объекты, распознанные представленным алгоритмом как объекты интереса, чёрным прямоугольником выделены объекты, распознанные как помехи.



Рис. 3. Результат работы алгоритма



Рис. 4. Результат работы алгоритма

На этапе разработки данного экспериментального образца опико-электронной системы беспилотного летательного аппарата был применен ряд программных моделей, в совокупности описывающих подобную опико-электронную систему, и представляющих собой комплексную математическую модель (КММ). Данная комплексная математическая модель позволяет генерировать тестовые ви-

деопоследовательности с оптико-электронной системы летательного аппарата в реальном времени, задавать траекторию носителя ОЭС, создавать искусственные и естественные помехи на моделируемых изображениях. Использование КММ позволило опробовать множество вариантов архитектур искусственных нейронных сетей и выбрать наиболее подходящую по совокупности признаков (высокая точность распознавания, малое количество ложноположительных и ложноотрицательных срабатываний, устойчивость к шумовым воздействиям и помехам на изображениях в сочетании с умеренной вычислительной сложностью получаемого в результате решения) для применения в разрабатываемом изделии.



Рис. 5. Результат работы алгоритма

В целом, применение компьютерного моделирования позволило существенно повысить качество разрабатываемого комплекса, а также сократить временные и финансовые затраты на разработку.

Искусственная нейронная сеть, описанная в данной работе, доказала свою эффективность и превосходство над другими рассмотренными архитектурами ИНС как на этапе компьютерного моделирования, так и на этапе натурных испытаний в составе ОЭС беспилотного летательного аппарата. Разработанная архитектура, благодаря своей изначальной ориентированности на использование в составе встраиваемых систем, позволила выполнить высокоэффективную аппаратную реализацию ИНС с использованием ПЛИС, входящей в состав вычислителя программно-аппаратного комплекса оптико-электронной системы. Производительность полученного решения позволила осуществлять автоматическое распознавание объектов интереса в реальном масштабе времени.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Ullrich K., Meeds E., Welling M.* Soft weight-sharing for neural network compression.
2. *Han S., Mao H., Dally W.J.* Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding.
3. *Courbariaux M., Bengio Y., David J.P.* Binaryconnect: Training deep neural networks with binaryweights during propagations // *Advances in neural information processing systems.* – 2015. – P. 3123-3131.
4. *Ioffe S., Szegedy C.* Batch normalization: Accelerating deep network training by reducing internal covariate shift // *arXiv preprint arXiv:1502.03167.* 2015.
5. *Courbariaux M. et al.* Binarized neural networks: Training deep neural networks with weights and activations constrained to+ 1 or-1 // *arXiv preprint arXiv:1602.02830.* 2016.



6. *Hinton Geoffrey*. Neural networks for machine learning. Coursera. Video lectures. – 2012.
7. *Rastegari M. et al*. Xnor-net: Imagenet classification using binary convolutional neural networks // *European Conference on Computer Vision*. – Springer, Cham, 2016. – P. 525-542.
8. *Jia Y. et al*. Caffe: convolutional architecture for fast feature embedding // *Proceedings of the 22nd ACM International Conference on Multimedia*. – 2014. – P. 675-678.
9. *Kingma D., Ba J. Adam*. A method for stochastic optimization // *International Conference on Multimedia*. – 2014. – P. 675-678.
10. *Тупиков В.А., Павлова В.А., Бондаренко В.А., Александров В.А.* Способ автоматического обнаружения объектов на морской поверхности в видимом диапазоне // *Известия ТулГУ. Технические науки*. – 2016. – № 11-3. – С. 105-121.
11. *Gonzalez R.C., Woods R.E.* Digital Image Processing. – 2nd ed. – Prentice-Hall, 2002. – 793 p.
12. *Sauvola J., Pietikainen M.* Adaptive document image binarization // *Pattern Recognition*. – 2000. – No. 33. – P. 225-236.
13. *Shafait F., Keysers D., Breuel T.M.* Efficient implementation of local adaptive thresholding techniques using integral images // *Proc. SPIE vol. 6815, Document Recognition and Retrieval XV*, 2008.
14. *Viola P., Jones M.J.* Robust real-time face detection // *Int. Journal of Computer Vision*. – 2004. – Vol. 57 (2). – P. 137-154.
15. *Crow F.* Summed-area tables for texture mapping // *In Proceedings of SIGGRAPH*. – 1984. – Vol. 18 (3). – P. 207-212.
16. *Баклицкий В.К.* Корреляционно-экстремальные методы навигации и наведения. – Тверь: ТО «Книжный клуб», 2009. – 360 с.
17. *LeCun Y., Boser B., Denker J.S., Henderson D., Howard R.E., Hubbard W. and Jackel L.D.* Backpropagation Applied to Handwritten Zip Code Recognition // *Neural Computation*. – 1989. – No. 1 (4). – P. 541-551.
18. *Goodfellow I., Bengio Y., Courville A.* Deep Learning. – MIT Press, 2016. – 781 p.
19. *Deng L. and Deep D.Yu.* Learning: Methods and Applications // *Foundations and Trends in Signal Processing*. – 2013. – Vol. 7, No. 3-4. – P. 197-387.
20. *Ioffe, Sergey, and Christian Szegedy*. Batch normalization: Accelerating deep network training by reducing internal covariate shift // *arXiv preprint*. – 2015. arXiv:1502.03167.

## REFERENCES

1. *Ullrich K., Meeds E., Welling M.* Soft weight-sharing for neural network compression.
2. *Han S., Mao H., Dally W.J.* Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding.
3. *Courbariaux M., Bengio Y., David J.P.* Binaryconnect: Training deep neural networks with binaryweights during propagations, *Advances in neural information processing systems*, 2015, pp. 3123-3131.
4. *Ioffe S., Szegedy C.* Batch normalization: Accelerating deep network training by reducing internal covariate shift, *arXiv preprint arXiv:1502.03167*. 2015.
5. *Courbariaux M. et al*. Binarized neural networks: Training deep neural networks with weights and activations constrained to+ 1 or-1, *arXiv preprint arXiv:1602.02830*. 2016.
6. *Hinton Geoffrey*. Neural networks for machine learning. Coursera. Video lectures. 2012.
7. *Rastegari M. et al*. Xnor-net: Imagenet classification using binary convolutional neural networks, *European Conference on Computer Vision*. Springer, Cham, 2016, pp. 525-542.
8. *Jia Y. et al*. Caffe: convolutional architecture for fast feature embedding, *Proceedings of the 22nd ACM International Conference on Multimedia*, 2014, pp. 675-678.
9. *Kingma D., Ba J. Adam*. A method for stochastic optimization, *International Conference on Multimedia*, 2014, pp. 675-678.
10. *Tupikov V.A., Pavlova V.A., Bondarenko V.A., Aleksandrov V.A.* Sposob avtomaticheskogo obnaruzheniya ob"ektov na morskoy poverkhnosti v vidimom diapazone [Algorithm of automatic helicopter type drone landing using onboard optical-electronic system], *Izvestiya TulGU. Tekhnicheskie nauki [Izvestiya TulSU, Engineering Sciences]*, 2016, No. 11-3, pp. 105-121.
11. *Gonzalez R.C., Woods R.E.* Digital Image Processing. 2nd ed. Prentice-Hall, 2002, 793 p.
12. *Sauvola J., Pietikainen M.* Adaptive document image binarization, *Pattern Recognition*, 2000, No. 33, pp. 225-236.

13. *Shafait F., Keysers D., Breuel T.M.* Efficient implementation of local adaptive thresholding techniques using integral images, *Proc. SPIE vol. 6815, Document Recognition and Retrieval XV, 2008.*
14. *Viola P., Jones M.J.* Robust real-time face detection, *Int. Journal of Computer Vision*, 2004, Vol. 57 (2), pp. 137-154.
15. *Crow F.* Summed-area tables for texture mapping, *In Proceedings of SIGGRAPH*, 1984, Vol. 18 (3), pp. 207-212.
16. *Baklitskiy V.K.* Korrelyatsionno-ekstremal'nye metody navigatsii i navedeniya [Correlation-extreme methods of navigation and guidance]. Tver': TO «Knizhnyy klub», 2009, 360 p.
17. *LeCun Y., Boser B., Denker J.S., Henderson D., Howard R.E., Hubbard W. and Jackel L.D.* Backpropagation Applied to Handwritten Zip Code Recognition, *Neural Computation*, 1989, No. 1 (4), pp. 541-551.
18. *Goodfellow I., Bengio Y., Courville A.* Deep Learning. MIT Press, 2016, 781 p.
19. *Deng L. and Deep D.Yu.* Learning: Methods and Applications, *Foundations and Trends in Signal Processing*, 2013, Vol. 7, No. 3–4, pp. 197-387.
20. *Ioffe, Sergey, and Christian Szegedy.* Batch normalization: Accelerating deep network training by reducing internal covariate shift, *arXiv preprint. – 2015. arXiv:1502.03167.*

Статью рекомендовал к опубликованию к.т.н. С.Н. Крюков.

**Тупиков Владимир Алексеевич** – АО «Научно-производственное предприятие «Авиационная и морская электроника»; e-mail: tupikov@nppame.ru; 198097, г. Санкт-Петербург, ул. Маршала Говорова, д. 29, лит. «О», а/я 51; тел.: 88123274667; д.т.н.; профессор; зам. ген. директора; директор научно-производственного комплекса робототехнических систем специального назначения (НПК РТС СН).

**Павлова Валерия Анатольевна** – e-mail: pavlova@nppame.ru, к.т.н.; зам. директора НПК РТС СН по НИОКР.

**Бондаренко Владимир Александрович** – e-mail: shockus@gmail.com; зам. начальника центра средств интеллектуальной обработки изображений.

**Каплинский Глеб Эдуардович** – e-mail: glebqq@gmail.com; начальник отдела прикладного программирования.

**Холод Николай Геннадьевич** – e-mail: nikkholod@mail.ru; младший научный сотрудник Центра средств интеллектуальной обработки изображений в РТС СН.

**Tupikov Vladimir Alekseevich** – SPE "Research and Production Enterprise "Air and Marine Electronics"; e-mail: tupikov@nppame.ru; 198097, Saint Petersburg, str. Marshal Govorov, 29, lit. "O", P.O.B 51; phone: 88123274667; dr. of eng. sc.; professor; Deputy Director General; Director of Research and production complex of special purpose robotic systems.

**Pavlova Valeria Anatolyevna** – e-mail: pavlova@nppame.ru; cand. of eng. sc.; Deputy Director of research at Research and production complex of special purpose robotic systems.

**Bondarenko Vladimir Alexandrovich** – e-mail: shockus@gmail.com; Deputy Chief of intellectual image processing center.

**Kaplinskiy Gleb Eduardovich** – e-mail: glebqq@gmail.com; chief of applied programming department.

**Holod Nikolay Gennadievich** – e-mail: nikkholod@mail.ru; junior researcher Center Intellectual image processing tools in RTS SP.