

**К.А. Суминов, Н.А. Бочаров**

**ИССЛЕДОВАНИЕ ОГРАНИЧЕНИЙ ПРИМЕНИМОСТИ  
МИКРОПРОЦЕССОРОВ РЯДА ЭЛЬБРУС ДЛЯ РЕШЕНИЯ ЗАДАЧ  
ТЕХНИЧЕСКОГО ЗРЕНИЯ**

*Одним из ключевых направлений развития области искусственного интеллекта является направление технического зрения. Для ресурсоемких задач технического зрения создаются высокопроизводительные вычислительные комплексы (ВК), использующие специализированные ускорители. Использование таких ускорителей обусловлено неспособностью микропроцессоров (МП) общего назначения решить такие задачи за поставленное время вследствие большой вычислительной нагрузки. Однако, в настоящее время, МП ряда Эльбрус успешно используются для решения задач технического зрения как в серверном, так и в бортовом режимах, а появление МП Эльбрус шестого поколения должно еще больше повысить производительность на таких задачах. В силу дороговизны, большей сложности и ограничений в использовании систем со специализированными ускорителями встает вопрос об определении условий, в которых для решения рассматриваемых задач достаточно использовать МП общего назначения, например из ряда Эльбрус, без специальных ускорителей. Одними из самых ресурсоемких задач в области технического зрения являются задачи обнаружения и классификации объектов на изображении. Для обнаружения объектов одним из популярных методов является метод Виолы-Джонса. Для решения задачи классификации как правило используются сверточные нейронные сети. На примере сетей VGG16 и VGG19 разработаны математические модели вычислений применительно к актуальным микропроцессорам ряда Эльбрус. С использованием разработанных моделей обоснована теоретическая достаточность производительность микропроцессоров Эльбрус для задач технического зрения. Также, на основе данных методов разработаны программы моделирования поиска и объектов на фото и в видеопотоке. Программы написаны на языке C++ с использованием библиотеки OpenCV, ОПО Эльбрус, библиотеки Платформы-ГНС и базы данных соревнований ImageNet. С использованием реализованных программ было проведено сравнительное тестирование на ряде ВК с процессорами Эльбрус, а также ВК с процессорами intel и ускорителем в виде видеокарты Nvidia. На основе полученных результатов показано, что для решения задачи поиска объектов на изображении выбранным методом достаточно МП Эльбрус-8С для входных разрешений до 1920 x 1080, где обеспечивается скорость обработки видеопотока более 20 кадров в секунду.*

*Техническое зрение; микропроцессоры эльбрус; сверточные нейронные сети; OpenCV.*

**К.А. Suminov, N.A. Bocharov**

**RESEARCH OF APPLICABILITY LIMITATIONS FOR ELBRUS  
MICROPROCESSORS FOR SOLVING TASKS OF TECHNICAL VISION**

*One of the key areas in the artificial intelligence is technical vision. For resource-intensive tasks of technical vision high-performance, computing systems are created with use of specialized accelerators. The use of such accelerators is necessary due to the inability of general-purpose microprocessors (GPM) to solve such problems in a given time due to a high computational load. However, the microprocessors of Elbrus series are successfully used to solve technical vision problems in both server and on-board modes, and the appearance of the sixth-generation Elbrus microprocessors should further improve performance on such tasks. Due to the high cost, greater complexity and limitations in the use of systems with specialized accelerators, the question arises of determining the conditions under which, it is sufficient to use CPU's to solve the tasks of technical vision, for example, with the microprocessors of the Elbrus series without special accelerators. One of the most resource-intensive tasks in the field of technical vision are detection and classification of objects. For the detection of objects one of the popular methods is the Viola-Jones method. Convolutional neural networks are usually used to solve the classification problem.*

*Mathematical models of computations have been developed for VGG16 and VGG19 neural networks in relation to the actual microprocessors of the Elbrus series. Using the developed models, the theoretical sufficiency of the performance of Elbrus microprocessors for technical vision tasks is substantiated. Also, based on these methods, programs for modeling detection and classification objects in the image and video stream have been developed. The programs are written in C++ using the OpenCV library, OPO Elbrus, the GNS Platform library and the ImageNet competition database. Using the implemented programs, comparative testing was carried out on a number of high-performance computing systems with Elbrus and Intel CPU's and NVidia video card. Based on the results obtained, it is shown that the Elbrus-8S is sufficient to solve the problem of searching for objects in the image for input resolutions up to 1920 x 1080, where the processing speed of the video stream is more than 20 frames per second.*

*Technical vision; elbrus microprocessors; convolutional neural networks; OpenCV.*

**Введение.** Задачи технического зрения в настоящее время являются важным направлением развития области искусственного интеллекта [1]. Создание новых поколений автономных робототехнических комплексов, систем распознавания и систем технического зрения в целом невозможно без использования современной вычислительной техники. При этом для решения подобного рода задач в настоящее время активно применяются и разрабатываются [2–4] вычислительные комплексы с использованием специализированных ускорителей. Тем не менее, граница применимости таких спец. ускорителей при проектировании вычислительных комплексов для решения подобных задач часто определяется эмпирически, в особенности для ВК на основе МП серии «Эльбрус» [5]. Т.к. в силу особенности архитектуры МП «Эльбрус» – Very Long Instruction Word (VLIW) – сложно и не всегда возможно оценить сложность и ресурсоемкость решения на базе имеющихся решений подобных задач, реализованных для систем с МП других архитектур.

Одним из важных и актуальных применением бортовых систем с МП серии «Эльбрус» являются бортовые вычислители и системы технического зрения. В ходе проведенных в АО МЦСТ работ показано, что вычислители на основе МП серии «Эльбрус» могут и успешно используются для решения задач технического зрения как в серверном [6] так и в бортовом [7] режиме. В 2017 году АО МЦСТ разработан и внедрен бортовой вычислитель на базе МП Эльбрус-4С, обеспечивающий достаточную производительность для автономного движения робота на скоростях до 40 км\ч. В работе [8] было обосновано, что для обеспечения корректной работы на такой скорости достаточным условием является работа системы технического зрения с производительностью не менее 10 кадров в секунду.

Появление новых МП серии Эльбрус, таких как Эльбрус-8СВ, Эльбрус-2С3 и Эльбрус-16С [9], а также средств вычислительной техники на их основе [10] открывает новые перспективы перед разработчиками робототехнических комплексов (РТК). Высокая производительность, обеспечиваемая новыми МП серии Эльбрус, позволит создавать бортовые вычислительные комплексы, способные решать задачи технического зрения на РТК, с использованием отечественной программно-аппаратной платформы.

Целью данной работы является исследование ограничений применимости МП ряда Эльбрус для решения задач технического зрения для определения структуры перспективных бортовых вычислительных комплексов.

**Особенности аппаратно-программной платформы «Эльбрус» для задач технического зрения.** Ключевые особенности МП серии «Эльбрус» - в том числе архитектура VLIW, предполагающая содержание нескольких операций в одной инструкции процессора, когда распараллеливание между множеством вычислительных блоков процессора происходит на этапе компиляции программы, большой парк вычислительных устройств (АЛУ) в составе процессоров «Эльбрус», большое количество операций - 25 в скалярном и 41 и векторном режимах [11] за такт

на одно ядро (Для Эльбрус-8С), за счет явного параллелизма микроопераций дают возможность поставить процессоры «Эльбрус» в ряд между процессорами общего назначения и специализированными процессорами, использующими SIMD инструкции. В совокупности с большой степенью параллельности многих задач, обеспечивающих функционирование систем технического зрения, это дает основание предполагать, что задачи подобного класса могут эффективно решаться и использованием современных МП серии «Эльбрус».

В табл. 1 представлены основные характеристики актуальных МП «Эльбрус» а также эквивалентных процессоров Intel.

Таблица 1

Основные характеристики рассматриваемых МП

Процессор	Эльбрус 8С	Эльбрус 8СВ	Эльбрус 16С	Эльбрус 2С3	intel Core i7-2600	intel Core i7-8565U	intel Core i7-8700	Intel Xeon silver 4110	Intel Xeon e5 2620 v4	NVidia GeForce GTX 960
Микроархитектура	elbrus – v4	elbrus - v5	elbrus – v6	elbrus – v6	Sandy-Bridge	Whiskey Lake	Coffee Lake	SkyLake	Broadwell	Maxwell
Частота (МГц)	1300	1500	2000	2000	3400-3800	1800-4600	3200-4600	2100-3000	2100-3000	1126-1178
Ядра	8	8	16	2	4	4	6	8	8	1024
Техпроцесс (нм)	28	28	16	16	32	14	14	14	14	28
TDP(Вт) не более	70	80	110	15	95	25	65	85	85	120
Тип ОЗУ	DDR3-1600	DDR4 - 2400	DDR4 -3200	DDR4-3200	DDR3 -1333	DDR4-2400	DDR4 2666	DDR4-2400	DDR4-2133	GDDR5-1752
Год выпуска	2016	2020	2022	2022	2011	2018	2018	2017	2016	2015

**Математическая модель вычислений задач технического зрения на платформе Эльбрус.** Системы технического зрения предназначены для решения задач получая информацию из изображений различного рода, в том числе и не графических. Задачи, для решения которых применяются системы технического зрения, включают в себя классификацию, кластеризацию, отслеживание, обнаружение, сегментацию и др. [12].

Для решения задач классификации и кластеризации объектов успешно применяются алгоритмы, основанные на использовании нейронных сетей [13]. Одним из примеров такой сети является fully connected сеть, где каждый нейрон следующего слоя соединен с каждым нейроном предыдущего слоя. В таких сетях успешно используются различные виды признаков, в том числе: бинарный, номинальный, порядковый, количественный и др. В реальных случаях возможно очень большое количество признаков для каждого объекта, и если среди них есть коррелирующие между собой признаки находящиеся в непосредственной зависимости, если не все они имеют ценность в конкретном случае и такие признаки могут быть исключены без существенной потери информации, то применяются методы понижения размерности пространства признаков для уменьшения вычислительной нагрузки. При работе с изображениями с задачей лучше справляются сверточные нейронные сети [14, 15], которые позволяют сильно сократить объем необходимых для работы параметров.

Работа сверточной нейронной сети обычно интерпретируется как переход от конкретных особенностей изображения к более абстрактным деталям, и далее к ещё более абстрактным деталям вплоть до выделения понятий высокого уровня.

При этом сеть самоорганизуется и вырабатывает сама необходимую иерархию абстрактных признаков (последовательности карт признаков), фильтруя маловажные детали и выделяя существенные. Работа таких сетей основана на использовании операций свертки примененной к изображению. Используя различные ядра свертки формируется поиск различных признаков на изображении. Схема применения свертки приведена на рис. 1.

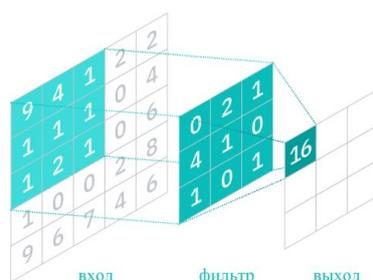


Рис. 1. Схема применения свертки к изображению

Из схемы видно, что для вычисления свертки основными операциями являются операции умножения и сложения, которые могут быть реализованы в виде совмещенных операций умножения-сложения. Также в архитектуре такой сети могут использоваться слои пуллинга, подвыборки, субдискретизации используемые для уменьшения размерности карты признаков. Пример работы max-pooling слоя представлен на рис. 2.

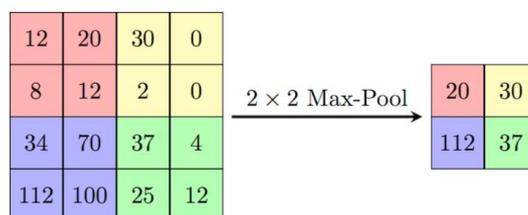


Рис. 2. Схема работы max-pooling слоя

Известными архитектурами сверточных нейронных сетей, дающими хороший результат top 5 accuracy (среди топ 5 результатов сети присутствует ожидаемый ответ) – более 92% на соревновании ImageNet [16] являются сети VGG16 и VGG19 [17]. Схематичное изображение слоев сети VGG19 представлено на рис. 3. Сеть использует изображения 224x224x3 на входе, оперирует с FP32 числами, имеет 19 слоев с весами и 5 pooling слоев. Первая часть сети (содержащая сверточные слои) используется для извлечения признаков с изображения, вторая часть, состоящая из последних 3 слоев представляет собой Fully Connected нейронную сеть для распознавания одного из классов, к которому принадлежит изображение. Архитектура VGG16 отличается от VGG19 отсутствием 8, 12 и 16 слоев.

Для теоретического обоснования времени выполнения вычислений на примере МП Эльбрус, разработана математическая модель вычислений, учитывающая количество операций, производимых при расчёте нейронной сети с архитектурой VGG16 и VGG19. В табл. 2 представлены необходимое количество параметров и их объем при выполнении вычислений сети указаны архитектуры, а также количество и тип необходимых операций для VGG19 и VGG16. Слои, помеченные (\*) не входят в VGG16.

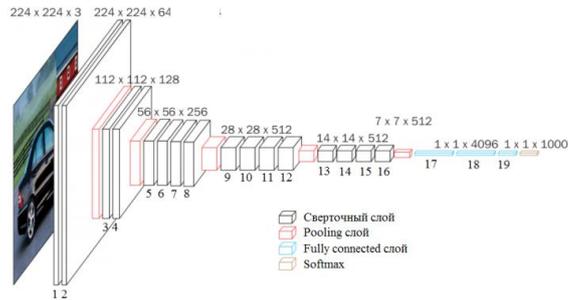


Рис. 3. Схема архитектуры VGG16

Таблица 2

Количество параметров, операций и их тип для слоев сети VGG19

номер слоя	слой	размерность входа			количество умножений в слое	количество сложений в слое	количество сравнений в слое	количество параметров в слое
		Ш	В	Г				
1	conv3-64	224	224	3	89915392	86704128	0	1792
2	conv3-64	224	224	64	1852899328	1849688064	0	36928
	pool2	224	224	64	0	0	2408448	0
3	conv3-128	112	112	64	926449664	924844032	0	73856
4	conv3-128	112	112	128	1851293696	1849688064	0	147584
	pool2	112	112	128	0	0	1204224	65664
5	conv3-256	56	56	128	925646848	924844032	0	295168
6	conv3-256	56	56	256	1850490880	1849688064	0	590080
7	conv3-256	56	56	256	1850490880	1849688064	0	590080
8(*)	conv3-256	56	56	256	1850490880	1849688064	0	590080
	pool2	56	56	256	0	0	602112	0
9	conv3-512	28	28	256	925245440	924844032	0	1180160
10	conv3-512	28	28	512	1850089472	1849688064	0	2359808
11	conv3-512	28	28	512	1850089472	1849688064	0	2359808
12(*)	conv3-512	28	28	512	1850089472	1849688064	0	2359808
	pool2	28	28	512	0	0	301056	0
13	conv3-512	14	14	512	462522368	462422016	0	2359808
14	conv3-512	14	14	512	462522368	462422016	0	2359808
15	conv3-512	14	14	512	462522368	462422016	0	2359808
16(*)	conv3-512	14	14	512	462522368	462422016	0	2359808
	pool2	14	14	512	0	0	75264	0
17	fc4096	1	1	25088	102760448	102760448	0	102764544
18	fc4096	1	1	4096	16777216	16777216	0	16781312
19	fc1000	1	1	4096	4096000	4096000	0	4097000
		Всего VGG19			19646914560	19632062464	4290048	143732904
		Всего VGG16			15483811840	15470264320	4591104	138423208

В табл. 3 представлены значения теоретического времени выполнения расчетов сетей (inference) при рассмотрении идеальной модели процессоров, в которой отсутствуют задержки по памяти, длительность вычислений ограничивается лишь скоростью работы и степенью конвейеризации (АЛУ) в составе. В действительности, с учетом не идеальности процессов вычисления и существующих потерь, учитывая размер кэш памяти, с использованием механизма предподкачки (prefetch) реализованного в «Эльбрусах» [18], механизма АРВ, а также при эффективной реализации умножения матриц блоками, использования схемы unroll and fuse и возможности АЛУ выполнять зацепленные операции умножения и сложения, скорость предподкачки достигает значения, обеспечивающего на ~90% эффективную загрузку АЛУ. Время выполнения  $T$  в этом случае можно посчитать как

$$T = \frac{(Nc+Nm+Ns)*R*k}{F*C*A*S},$$

где  $Nc$  – количество совмещенных операций,  $Nm$  и  $Ns$  – количество отдельных операций умножения и сложения соответственно,  $R$  – разрядность чисел,  $k$  – коэффициент эффективности,  $F$  – тактовая частота процессора,  $C$  – количество ядер,  $A$  – количество АЛУ включающих Float Point Unit (FPU),  $S$  – разрядность FPU. Теоретическое время выполнения представлено также в табл. 3.

Таблица 3

**Время выполнения вычислений для сетей VGG16 и VGG19 на МП Эльбрус**

	VGG16				VGG19			
количество совмещенных операций + *	15470264320				19632062464			
количество отдельных операций *	13547520				14852096			
количество отдельных операций +	0				0			
процессор	Эль-брус 8С	Эль-брус 8СВ	Эль-брус 16С	Эль-брус 2С3	Эль-брус 8С	Эль-брус 8СВ	Эль-брус 16С	Эль-брус 2С3
теоретическое время расчета прямого прохода сети (мс.)	137,9	59,7	22,4	179,2	174,9	75,8	28,4	227,4
теоретическое количество кадров в секунду	7,3	16,7	44,6	5,6	5,7	13,2	35,2	4,4

**Программная модель системы технического зрения на платформе «Эльбрус».** Т.к. большая часть задач технического зрения сводится к задачам обнаружения и классификации, то для исследования были выбраны именно эти задачи – для моделирования задачи обнаружения была разработана программа обнаружения объектов в видеопотоке методом Виолы-Джонса [19], для моделирования задачи классификации была разработана программа, реализующая вычисления нейронной сети с архитектурой VGG16. Также использованы модели других нейросетей, написанные с использованием «Платформы-ГНС» разработки ГосНИИАС, использующей оптимизированные для МП Эльбрус реализации некоторых нейронных сетей.

Для проведения эксперимента по решению задачи обнаружения объектов в видеопотоке была реализована программа на языке Си++ с использованием библиотеки OpenCV 3.2.0 оптимизированной для МП «Эльбрус-8С». Для проведения эксперимента по решению задачи классификации с помощью нейронной сети была реализована программа на языке программирования Си без внешних зависимостей для выполнения расчетов соответствующих нейронной сети с архитектурой VGG16. Использовались данные соревнования ImageNet, в том числе предобученные веса сети и база данных изображений. Для работы с реализациями нейронных сетей «Платформа-ГНС» также использовался язык программирования Си совместно с библиотекой разработанной ГосНИИАС. Результаты времени вычислений для реализации нейронной сети VGG16 представлены в табл. 4.

Таблица 4

**Время выполнения вычислений на Эльбрус 8С для различных нейронных сетей с использованием «платформы» ГосНИИАС**

Архитектура сети	AlexNet	LeNet	ResNet18	ResNet34	ResNet50	VGG19	MobileNetV1
Время выполнения (мс.)	25,1	12,3	40,6	259,2	316,1	266,7	40,6
Кадров в секунду	39,8	81,3	24,6	3,9	3,2	3,7	24,6

Для проведения эксперимента с обнаружением был выбран один видеоряд со следующими разрешениями: 424 x 240, 640 x 360, 854 x 480, 1280 x 720, 1920 x 1080, 2560 x 1440, 3840 x 2160. В качестве входных данных для нейронных сетей были взяты изображения из базы данных ImageNet с разрешением 224x224x3.

В качестве испытательных стендов были выбраны различные процессоры семейства Эльбрус, в том числе 2, 8 и 16 ядерный, а также МП Intel core i7 2600k, серверные МП Intel Xeon 4110 и Xeon e5 2620, выпуска 2016-2017 года, 8 ядерные, 16 поточные и мобильный процессор Intel core i7-8565U производства 2018 года. В качестве спец. ускорителя была использована видеокарта NVidia GeForce GTX 960 совместно с МП Intel core i7 2600k. Результаты с временем выполнения для задачи обнаружения в видеопотоке представлены на рис. 4.



Рис. 4. Количество кадров в секунду при решении задачи обнаружения

Из полученных данных видно, что все протестированные процессоры семейства Эльбрус стабильно превосходят Intel core i7 2600k, в том числе 2х ядерный Эльбрус-2С3. Из результатов видно, что на малых разрешениях использование совместно с процессором ускорителя в виде видеокарты не только не дает выиг-

рыша по времени, но и обеспечивает худшие результаты в отличие от использования только процессора, что связано с особенностью распределенных вычислений с использованием специализированных ускорителей, а именно с длительной и частой пересылкой маленьких для расчёта объемов данных между процессором и ускорителем. Однако, стоит заметить, что в ходе исследования использовалась относительно слабая видеокарта и при использовании современных специализированных ускорителей типа Модуль [20], Эльвис [21], Nvidia, разница в результатах будет отличаться. Также в ходе эксперимента установлено, что Эльбрус-8С превосходит вычислитель на основе Intel Core i7 2600k совместно с видеокартой GTX 960 на разрешениях вплоть до 854x480, а на разрешении 1280x720 Эльбрус-8С показал результат до 47 кадров в секунду, при этом Эльбрус-2С3 обеспечивает до 18, а Эльбрус-16С до 57 кадров в секунду в то время как Intel Core i7 2600 – всего около 11 кадров, а совместно с видеокартой – 62 кадра в секунду. Эльбрус-16С показал сравнимый с серверными процессорами Intel Xeon результат на больших разрешениях входного видеопотока, а Эльбрус-8С показал отставание от них на около 10-30% при вдвое меньшем количестве потоков.

**Выводы.** В ходе данного исследования разработаны математические модели вычислений сверточных нейронных сетей с архитектурой VGG16 и VGG19, получено теоретическое обоснование минимального времени выполнения расчетов на процессорах Эльбрус 2С3, 8С, 8СВ и 16С. Разработаны программные модели для решения задачи поиска объектов в видеопотоке с использованием OpenCV и для решения задач классификации с использованием ПО «Платформа-ГНС». Проведены эксперименты по обнаружению объектов в видеопотоке с различным разрешением, а также по классификации изображений с разрешением 224x224x3. Результаты, полученные в ходе экспериментов согласуются с теоретическими результатами, основанными на разработанной математической модели.

В результате проведенных экспериментов показано, что использование микропроцессоров Эльбрус 2С3, 8С, 8СВ и 16С без дополнительных ускорителей обеспечивает достаточную производительность для использования в бортовых вычислителях и системах технического зрения автономных роботов вплоть до разрешения 2560 x 1440, где обеспечивается частота обработки до 14–15 кадров в секунду.

Были продемонстрированы результаты и потенциал МП «Эльбрус» для решения задач технического зрения, что позволило согласовать требования к вычислителям перспективных автономных роботов.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Есенбаева Г.Б.* Техническое зрение для мобильных роботов // Передовые инновационные разработки. Перспективы и опыт использования, проблемы внедрения в производство. – 2019. – С. 73-75.
2. *Gondimalla A. [et al.]*. Sparten: A sparse tensor accelerator for convolutional neural networks // Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture. – 2019. – P. 151-165.
3. *Deng L. [et al.]*. Model compression and hardware acceleration for neural networks: A comprehensive survey // Proceedings of the IEEE. – 2020. – Vol. 108, No. 4. – P. 485-532.
4. *Auten A., Tomei M., Kumar R.* Hardware acceleration of graph neural networks // 2020 57th ACM/IEEE Design Automation Conference (DAC). – IEEE, 2020. – P. 1-6.
5. *Ким А.К. [и др.]*. Микропроцессоры и вычислительные комплексы семейства "Эльбрус". – СПб.: Издательский дом "Питер", 2013.
6. *Лимонова Е.Е. [и др.]*. Оценка быстродействия системы распознавания на VLIW архитектуре на примере платформы Эльбрус // Программирование. – 2019. – № 1. – С. 15-21.
7. *Бочаров Н.А., Зувев А.Г., Славин О.А.* Производительность микропроцессора Эльбрус-8СВ для решения задач технического зрения в условиях ограничений энергопотребления // Известия ЮФУ. Технические науки. – 2021. – № 1 (218). – С. 259-271.

8. Бочаров Н.А. и др. Производительность вычислительной техники с процессором «Эльбрус-8С» на задачах робототехнического комплекса // *Наноиндустрия*. – 2018. – № 82. – С. 79.
9. Кожин А.С. Основные проектные решения для процессора «Эльбрус-16С» // *Наноиндустрия*. – 2020. – Т. 13, № S4. – С. 74-75.
10. Бычков И.Н., Лобанов И.Н., Молчанов И.А. Вычислительная техника на основе аппаратно-программной платформы «Эльбрус» для перспективных информационных систем // *Приборы*. – 2018. – №. 8. – С. 14.
11. Кожин А.С., Нейман-заде М.И., Тихорский В.В. Влияние подсистемы памяти восьмиядерного микропроцессора «Эльбрус-8С» на его производительность // *Вопросы радиоэлектроники*. – 2019. – №. 3. – С. 13-21.
12. Визильтер Ю.В. и др. Обработка и анализ изображений в задачах машинного зрения. – 2010.
13. Визильтер Ю.В., Горбачевич В.С., Желтов С.Ю. Структурно-функциональный анализ и синтез глубоких конволюционных нейронных сетей // *Компьютерная оптика*. – 2019. – Т. 43, № 5. – С. 886-900.
14. Albawi S., Mohammed T.A., Al-Zawi S. Understanding of a convolutional neural network // 2017 international conference on engineering and technology (ICET). – Ieee, 2017. – P. 1-6.
15. O'Shea K., Nash R. An introduction to convolutional neural networks // *arXiv preprint arXiv:1511.08458*. – 2015.
16. Krizhevsky A., Sutskever I., Hinton G.E. Imagenet classification with deep convolutional neural networks // *Advances in neural information processing systems*. – 2012. – Vol. 25.
17. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition // *arXiv preprint arXiv: 1409.1556*. – 2014.
18. Нейман-заде М.И., Королёв С.Д. Руководство по эффективному программированию на платформе "Эльбрус". – М.: АО «МЦСТ, 2020.
19. Viola P. [et al.]. Robust real-time object detection // *International journal of computer vision*. – 2001. – Vol. 4, No. 34-47. – P. 4.
20. Бирюков А.А., Таранин М.В., Таранин С.В. Процессор 1879ВМ6Я. Реализация глубоких свёрточных нейронных сетей // *DSPA: Вопросы применения цифровой обработки сигналов*. – 2018. – Т. 8, №. 4. – С. 191-195.
21. Петричкович Я. и др. RoboDeus-50-ядерная гетерогенная СнК для встраиваемых систем и робототехники // *Электроника: Наука, технология, бизнес*. – 2020. – № 7. – С. 52-63.

## REFERENCES

1. Esenbaeva G.B. Tekhnicheskoe zrenie dlya mobil'nykh robotov [Technical vision for mobile robots], *Peredovye innovatsionnye razrabotki. Perspektivy i opyt ispol'zovaniya, problemy vnedreniya v proizvodstvo* [Advanced innovative developments. Prospects and experience of use, problems of introduction into production], 2019, pp. 73-75.
2. Gondimalla A. [et al.]. Sparten: A sparse tensor accelerator for convolutional neural networks, *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture*, 2019, pp. 151-165.
3. Deng L. [et al.]. Model compression and hardware acceleration for neural networks: A comprehensive survey, *Proceedings of the IEEE*, 2020, Vol. 108, No. 4, pp. 485-532.
4. Auten A., Tomei M., Kumar R. Hardware acceleration of graph neural networks, *2020 57th ACM/IEEE Design Automation Conference (DAC)*. IEEE, 2020, pp. 1-6.
5. Kim A.K. [i dr.]. Mikroprotsessory i vychislitel'nye komplekсы semeystva " El'brus" [Microprocessors and computing complexes of the Elbrus family]. Saint Petersburg: Izdatel'skiy dom " Piter", 2013.
6. Limonova E.E. [i dr.]. Otsenka bystrodeystviya sistemy raspoznavaniya na VLIW arkhitekture na primere platformy El'brus [Evaluation of the performance of the recognition system on the VLIW architecture on the example of the Elbrus platform], *Programmirovaniye* [Programming], 2019, No. 1, pp. 15-21.
7. Bocharov N.A., Zuev A.G., Slavin O.A. Proizvoditel'nost' mikroprotsessora El'brus-8SV dlya resheniya zadach tekhnicheskogo zreniya v usloviyakh ogranicheniy energopotrebleniya [The performance of the Elbrus-8SV microprocessor for solving problems of technical vision in conditions of limited energy consumption], *Izvestiya YuFU. Tekhnicheskie nauki* [Izvestiya SFedU. Engineering Sciences], 2021, No. 1 (218), pp. 259-271.

8. *Bocharov N.A. i dr.* Proizvoditel'nost' vychislitel'noy tekhniki s protsessorom «El'brus-8S» na zadachakh robototekhnicheskogo kompleksa [The performance of computing equipment with the Elbrus-8C processor on the tasks of the robotic complex], *Nanoindustriya* [Nanoindustry], 2018, No. 82, pp. 79.
9. *Kozhin A.S.* Osnovnye proektnye resheniya dlya protsessora «El'brus-16S» [Basic design solutions for the processor "Elbrus-16S"], *Nanoindustriya* [Nanoindustry], 2020, Vol. 13, No. S4, pp. 74-75.
10. *Bychkov I.N., Lobanov I.N., Molchanov I.A.* Vychislitel'naya tekhnika na osnove apparatno-programmnoy platformy «El'brus» dlya perspektivnykh informatsionnykh sistem [Computer technology based on the hardware and software platform "Elbrus" for advanced information systems], *Pribory* [Devices], 2018, No. 8, pp. 14.
11. *Kozhin A.S., Neyman-zade M.I., Tikhorskiy V.V.* Vliyanie podsistemy pamyati vos'miyadernogo mikroprotsessora «El'brus-8S» na ego proizvoditel'nost' [The influence of the memory subsystem of the eight-core microprocessor "Elbrus-8S" on its performance], *Voprosy radioelektroniki* [Questions of radio electronics], 2019, No. 3, pp. 13-21.
12. *Vizil'ter Yu.V. i dr.* Obrabotka i analiz izobrazheniy v zadachakh mashinnogo zreniya [Image processing and analysis in machine vision tasks], 2010.
13. *Vizil'ter Yu.V., Gorbatshevich V.S., Zheltov S.Yu.* Strukturno-funktsional'nyy analiz i sintez glubokikh konvolutsionnykh neyronnykh setey [Structural and functional analysis and synthesis of deep convolution neural networks], *Komp'yuternaya optika* [Computer optics], 2019, Vol. 43, No. 5, pp. 886-900.
14. *Albawi S., Mohammed T.A., Al-Zawi S.* Understanding of a convolutional neural network, *2017 international conference on engineering and technology (ICET)*. Ieee, 2017, pp. 1-6.
15. *O'Shea K., Nash R.* An introduction to convolutional neural networks, *arXiv preprint arXiv:1511.08458*, 2015.
16. *Krizhevsky A., Sutskever I., Hinton G.E.* Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, 2012, Vol. 25.
17. *Simonyan K., Zisserman A.* Very deep convolutional networks for large-scale image recognition, *arXiv preprint arXiv: 1409.1556*, 2014.
18. *Neyman-zade M.I., Korolev S.D.* Rukovodstvo po effektivnomu programmirovaniyu na platforme "El'brus" [Guide to effective programming on the Elbrus platform]. Moscow: AO «MTSST», 2020.
19. *Viola P. [et al.]*. Robust real-time object detection, *International journal of computer vision*, 2001, Vol. 4, No. 34-47, pp. 4.
20. *Biryukov A.A., Taranin M.V., Taranin S.V.* Protsessor 1879VM6YA. Realizatsiya glubokikh svertochnykh neyronnykh setey [Processor 1879VM6YA. Implementation of deep convolutional neural networks], *DSPA: Voprosy primeneniya tsifrovoy obrabotki signalov* [DSPA: Issues of application of digital signal processing], 2018, Vol. 8, No. 4, pp. 191-195.
21. *Petrichkovich Ya. i dr.* RoboDeus-50-yadernaya geterogennaya SnK dlya vstraivaemykh sistem i robototekhniki [RoboDeus-50-nuclear heterogeneous SnC for embedded systems and robotics], *Elektronika: Nauka, tekhnologiya, biznes* [Electronics: Science, Technology, Business], 2020, No. 7, pp. 52-63.

Статью рекомендовал к опубликованию д.т.н., профессор Н.Б. Парамонов.

**Суминов Константин Александрович** – ПАО «ИНЭУМ им. И.С. Брука»; г. Москва, Россия; e-mail: suminov.ka@phystech.edu; тел.: +79263563020; инженер-программист.

**Бочаров Никита Алексеевич** – e-mail: bocharov.na@phystech.edu; тел.: +79167346437; к.т.н.; начальник отдела.

**Suminov Konstantin Alexandrovich** – PJSC «Brook INEUM»; e-mail: suminov.ka@phystech.edu; Moscow, Russia; phone: +79263563020; software engineer.

**Bocharov Nikita Alexeevich** – e-mail: bocharov.na@phystech.edu; phone: +79167346437; cand. of eng. sc.; chief of department.