

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Объект авторского права

УДК 004.032.26

КРОЩЕНКО

Александр Александрович

**МЕТОДЫ ОБУЧЕНИЯ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ
ДЛЯ ЗАДАЧ КОМПЬЮТЕРНОГО ЗРЕНИЯ**

АВТОРЕФЕРАТ

диссертации на соискание ученой степени

кандидата технических наук

по специальности 05.13.17 – Теоретические основы информатики

Минск 2024

Научная работа выполнена
в УО «Брестский государственный технический университет».

НАУЧНЫЙ РУКОВОДИТЕЛЬ – Головкин Владимир Адамович,
доктор технических наук, профессор,
заведующий кафедрой интеллектуальных
информационных технологий
УО «Брестский государственный
технический университет».

ОФИЦИАЛЬНЫЕ ОППОНЕНТЫ: Старовойтов Валерий Васильевич,
доктор технических наук, профессор,
главный научный сотрудник
лаборатории идентификации систем
ГНУ «Объединенный институт проблем
информатики НАН Беларуси»;

Богущ Рихард Петрович,
доктор технических наук, доцент,
заведующий кафедрой
вычислительных систем и сетей
УО «Полоцкий государственный
университет имени Евфросинии Полоцкой».

ОППОНИРУЮЩАЯ ОРГАНИЗАЦИЯ –
УО Белорусский государственный
университет информатики
и радиоэлектроники.

Защита состоится **21 марта 2024 года** в 14:30 часов на заседании совета
по защите диссертаций Д 02.01.02 при Белорусском государственном универ-
ситете (г. Минск, ул. Ленинградская 8, корпус юридического факультета),
ауд. 407. Телефон ученого секретаря 209-57-09.

Почтовый адрес: пр-т Независимости 4, Минск, 220030.

Телефон ученого секретаря: 209-57-09; e-mail: Kochyn@bsu.by.

С диссертацией можно ознакомиться в Фундаментальной библиотеке
Белорусского государственного университета.

Автореферат разослан «____» февраля 2024 года.

Ученый секретарь совета по защите диссертаций
кандидат технических наук доцент

В.П. Кочин

ВВЕДЕНИЕ

Практические приложения компьютерного зрения с каждым годом становятся все более разнообразными. В современном мире компьютерное зрение используется повсеместно – в сложных производственных и медицинских системах, в интеллектуальных системах интернета вещей и развлекательных приложениях для мобильных устройств.

В качестве основы при разработке таких систем все чаще находят применение глубокие нейросетевые модели. Данные модели показывают впечатляющие результаты при решении самых разнообразных задач компьютерного зрения – распознавания, детекции и сегментации объектов на фото- и видеоизображениях, получения аннотаций для фотографий и генерации изображений по текстовому описанию. Глубокие нейронные сети, применяемые для решения подобных задач, содержат миллионы настраиваемых параметров и, для некоторых архитектур, десятки слоев нейронных элементов.

Обучение подобных «тяжелых» моделей с нуля является нетривиальной задачей. Оно часто сопряжено с риском переобучения, результатом которого является отличная приспособленность сети к данным из обучающей выборки, но плохая обобщающая способность, то есть неэффективность модели для данных, не использовавшихся при обучении. Чаще всего переобученность возникает при применении малой обучающей выборки. Другой проблемой является ресурсоемкость процесса обучения таких моделей, даже при использовании современных технических средств.

Проблемы обучения глубоких нейронных сетей активно изучаются в зарубежных научных школах. В нашей стране такие исследования также проводятся. Однако, нужно отметить, что такие исследования часто носят эмпирический характер, поэтому разработка строгих математически обоснованных методов обучения остается важной задачей теории нейронных сетей.

В диссертационной работе разработаны подходы для неконтролируемого предобучения глубоких нейронных сетей и редуцирования параметров моделей. Предложены алгоритмы для решения практических задач теории компьютерного зрения – обнаружения и локализации солнечных панелей на аэрофотоснимках, обнаружения и распознавания маркировки на поточных производственных линиях. Предложенные методы и алгоритмы позволят улучшить работу интеллектуальных систем, использующих полносвязные и сверточные нейросетевые модели.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Связь работы с научными программами (проектами), темами

Тема диссертации соответствует приоритетному направлению научно-технической деятельности согласно пункту 1 перечня приоритетных направлений научной, научно-технической и инновационной деятельности на 2021-2025 годы (Указ Президента Республики Беларусь от 07 мая 2020 г. № 156).

Исследования по теме диссертационной работы проводились в рамках научных программ:

1. НИР МОРБ «Алгоритмы интеллектуального анализа и обработки больших объемов данных на основе нейронных сетей глубокого доверия» (ГБ 15/203, № госрегистрации 20150743),

2. ГПНИ «Информатика и космос, научное обеспечение безопасности и защиты от чрезвычайных ситуаций» по заданию «Нейросетевые методы обработки комплексной информации и принятия решений на основе интеллектуальных многоагентных систем» (№ госрегистрации 20140547),

3. ГПНИ «Информатика и космос, научное обеспечение безопасности и защиты от чрезвычайных ситуаций» по заданию «Методы и алгоритмы интеллектуальной обработки и анализа большого объема данных на основе нейронных сетей глубокого доверия» (задание 1.6.05, № госрегистрации 20163595),

4. НИР «Методы и алгоритмы построения интеллектуальных систем анализа и обработки данных», этап «Разработка гибридных интеллектуальных систем на основе нейросимволического подхода» (решение НТС УО «Брестский государственный технический университет» от 12.11.2021, протокол № 6, № 22202052022070),

5. НИР БРФФИ «Модели и исследование 3-D оцифровки на основе фактических данных и анализа гетерогенных данных» (№ Ф22КИ-046 от 05.11.2021 г., № госрегистрации 20220090).

Цель, задачи, объект и предмет исследования

Целью исследования является разработка эффективных методов и алгоритмов для обучения глубоких нейронных сетей, используемых для решения задач компьютерного зрения, включающих обнаружение солнечных панелей на аэрофотоснимках и распознавание маркировки продукта на конвейерной линии.

Указанная цель определяет следующие *задачи исследования*:

1) Разработать метод неконтролируемого предобучения глубоких нейронных сетей, позволяющий повысить эффективность обучения моделей;

- 2) Разработать алгоритм редуцирования параметров глубоких нейронных сетей, позволяющий упростить структуру моделей;
- 3) Провести сравнительный анализ эффективности разработанных метода обучения и алгоритма редуцирования;
- 4) Разработать нейросетевую систему компьютерного зрения для решения прикладных задач (обнаружения солнечных панелей на аэрофотоснимках и распознавания маркировки продукта на конвейерной линии) с применением предлагаемого метода обучения.

Объектом исследования являются нейросетевые системы компьютерного зрения. *Предметом исследования* выступают методы и алгоритмы обучения глубоких нейронных сетей и их применение к задачам компьютерного зрения.

Научная новизна

Научная новизна состоит в установлении эквивалентности задач максимизации функции правдоподобия распределения входных данных, минимизации кросс-энтропийной функции ошибки и суммарной квадратичной ошибки при использовании линейных нейронов в пространстве синаптических связей ограниченной машины Больцмана, что позволяет учитывать нелинейную природу нейронных элементов.

Разработан метод обучения ограниченной машины Больцмана на основе доказанных эквивалентностей, применение которого для предобучения ГНС позволяет расширить класс обучаемых моделей и повысить обобщающую способность ГНС.

Разработан алгоритм редуцирования параметров нейросетевой модели, основывающийся на неконтролируемом предобучении, что позволяет уменьшить количество настраиваемых параметров модели без потери обобщающей способности.

Разработана нейросетевая система компьютерного зрения, которая основана на предлагаемом методе предобучения нейросетевых моделей, применение которой позволяет улучшить качество решения задач классификации. Продемонстрирована эффективность системы на примере решения задач обнаружения солнечных панелей на аэрофотоснимках и распознавания маркировки продуктов на производственной линии.

Положения, выносимые на защиту

1. Установление эквивалентности задач максимизации функции правдоподобия распределения входных данных, минимизации кросс-энтропийной функции ошибки и суммарной квадратичной ошибки при использовании ли-

нейных нейронов в пространстве синаптических связей ограниченной машины Больцмана, что позволяет учитывать нелинейную природу нейронных элементов.

2. Метод обучения ограниченной машины Больцмана, базирующийся на эквивалентности задач максимизации функции правдоподобия распределения входных данных и минимизации суммарной квадратичной ошибки при использовании линейных нейронов в пространстве синаптических связей сети, что позволяет расширить класс обучаемых моделей и повысить обобщающую способность глубоких нейронных сетей.

3. Алгоритм редуцирования параметров глубокой нейронной сети, базирующийся на неконтролируемом предобучении сети, что позволяет упростить ее архитектуру (путем сокращения числа настраиваемых параметров модели) без потери обобщающей способности.

4. Нейросетевая система компьютерного зрения, основывающаяся на предлагаемом методе предобучения нейросетевых моделей, применение которой позволяет улучшить качество решения задач классификации.

Личный вклад соискателя ученой степени в результаты диссертации

Основные положения диссертации получены соискателем лично. Соавтором основных публикаций автора является научный руководитель д.т.н., профессор В.А. Головкин, который осуществлял определение целей и постановку задач исследований, выбор методов исследований, принимал участие в планировании работ и обсуждении результатов. В диссертационную работу не включены результаты, которые были получены другими соавторами или с другими соавторами. Материалы совместных публикаций использованы соискателем в объеме авторского вклада.

Апробация диссертации и информация об использовании ее результатов

Основные положения и полученные результаты диссертационной работы докладывались и обсуждались на следующих конференциях: «8th International Conference on Neural Networks and Artificial Intelligence» (Брест, 3-6 июня 2014 г.); «8th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications» (Варшава, 24-26 сентября 2015 г.); «International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)» (Харьков, 9-12 октября 2018 г.); «14th International Conference on Pattern Recognition

and Information Processing (PRIP)» (Минск, 21-23 мая 2019 г.); «Вычислительные методы, модели и образовательные технологии» (Брест, 22-23 октября 2013 г., 15-16 октября 2014 г., 22 октября 2015 г., 21 октября 2016 г., 18 октября 2019 г.); «Современные проблемы математики и вычислительной техники» (Брест, 19-21 ноября 2015 г.); «Открытые семантические технологии проектирования интеллектуальных систем» (Минск, 19-21 февраля 2015 г., 15-17 февраля 2018 г., 21-23 февраля 2019 г., 19-22 февраля 2020 г., 16-18 сентября 2021 г., 20-22 апреля 2023 г.); «11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)» (Краков, 22-25 сентября 2021 г.); «Международный конгресс по информатике: информационные системы и технологии (CSIST'2022)» (Минск, 27-28 октября 2022 г.).

По результатам диссертации получено 3 акта о внедрении.

Опубликованность результатов диссертации

Основные результаты диссертационного исследования опубликованы в 30 научных работах, среди которых: 8 статей в научных изданиях в соответствии с пунктом 19 Положения о присуждении ученых степеней и присвоении ученых званий (общим объемом 3,87 авторского листа), 5 статей в других научных изданиях, 13 статей в сборниках материалов научных конференций и 4 тезисов.

Структура и объем диссертации

Диссертация состоит из перечня сокращений и обозначений, введения, общей характеристики работы, четырех глав, заключения, списка использованных источников и двух приложений.

Полный объем диссертации составляет 117 страниц, из которых 44 рисунка на 33 страницах, 17 таблиц на 9 страницах, 2 приложения на 13 страницах. Список использованных источников состоит из 110 наименований, включая 30 публикаций автора (на 5 страницах).

ОСНОВНАЯ ЧАСТЬ

В *первой главе* рассмотрено краткое введение в теорию нейронных сетей с классификацией их основных типов, определено понятие глубокой нейронной сети. Осуществлена постановка задачи обучения глубоких нейронных сетей, приведен обзор литературы по теме диссертации, описаны основные подходы к обучению нейросетевых моделей. В контексте проблемы обучения глубоких моделей рассмотрены два метода неконтролируемого предобучения – автоэнкодерный и метод, базирующийся на применении ограниченной машины Больцмана (RBM – Restricted Boltzmann Machine).

В случае применения неконтролируемого предобучения глубокая нейронная сеть обучается в два этапа. На первом этапе осуществляется ее предобучение путем формирования последовательности простых моделей из параметров ГНС, каждая из которых обучается на данных, получаемых от предыдущих моделей. На втором этапе с помощью настроенных параметров моделей инициализируется исходная ГНС, а затем дообучается выбранным методом обучения. Достоинство рассматриваемых методов в том, что с их помощью формируется хорошая начальная инициализация параметров нейронной сети, которая позволяет добиться меньшего значения ошибки, с которого начинает обучаться модель на втором этапе. Таким образом, сам процесс последующего обучения сети ускоряется, позволяя добиться более высокой обобщающей способности к завершению второго этапа. Помимо этого свойства, также становится не критичным теоретическое требование о сравнимости количества параметров нейронной сети с размером обучающей выборки. При использовании предобучения обучающая выборка может быть меньше суммарного количества параметров сети, без снижения ее способности эффективно обучаться.

Отметим, что теоретическая база, на основе которой получены правила обучения ограниченной машины Больцмана, переусложнена, и указанные правила могут быть получены, исходя из выбора другой целевой функции и скорректированного варианта применения правил обучения. Кроме этого, данный подход не учитывает возможность использования нелинейных функций активации нейронных элементов.

В качестве целевой функции, применяемой при выводе правил обучения RBM, используется функция логарифмического правдоподобия:

$$\ln P(x) = \ln \sum_y e^{-E(x,y)} - \ln \sum_{x,y} e^{-E(x,y)},$$

где $E(x, y)$ – энергия системы в момент, когда видимый и скрытый нейрон ограниченной машины Больцмана находятся в состоянии (x, y) .

Вычисляя частные производные данной функции по весовым коэффициентам и порогам, можно прийти к следующим уравнениям:

$$\begin{aligned} \frac{\partial \ln P(x)}{\partial w_{ij}} &= E[x_i y_j]_{\text{data}} - E[x_i y_j]_{\text{model}}, \\ \frac{\partial \ln P(x)}{\partial T_i} &= E[x_i]_{\text{data}} - E[x_i]_{\text{model}}, \\ \frac{\partial \ln P(x)}{\partial T_j} &= E[y_j]_{\text{data}} - E[y_j]_{\text{model}}, \end{aligned}$$

где w_{ij} – веса RBM,

T_i, T_j – пороговые элементы видимого и скрытого слоев.

В силу сложности вычисления математического ожидания в приведенных выше формулах в классическом методе обучения RBM используется процедура контрастного расхождения (Contrastive Divergence - CD). Для обозначения количества проходов по слоям RBM, используется параметр стадийности - k (CD- k).

Таким образом, формулы обучения для классического метода имеют вид (случай CD-1):

$$w_{ij}(t+1) = w_{ij}(t) + \alpha(x_i(0)y_j(0) - x_i(1)y_j(1)),$$

$$T_i(t+1) = T_i(t) + \alpha(x_i(0) - x_i(1)),$$

$$T_j(t+1) = T_j(t) + \alpha(y_j(0) - y_j(1)),$$

где α – скорость обучения RBM,

$x_i(0)$ – исходные данные,

$x_i(1)$ – исходные восстановленные данные,

$y_j(0)$ – данные, полученные со скрытого слоя RBM,

$y_j(1)$ – восстановленные данные, полученные со скрытого слоя RBM.

Во **второй главе** рассматривается альтернативный подход к предобучению глубокой нейронной сети, также базирующийся на применении ограниченных машин Больцмана, но в контексте другой целевой минимизируемой функции ошибок. Данный подход позволил получить схожие, а при некоторых условиях и эквивалентные правила обучения ограниченной машины Больцмана, что доказано в приведенных в данной главе теоремах.

В качестве целевых минимизируемых функций ошибок используются функции MSE (среднеквадратичная функция ошибок – Mean Squared Error) и CE (кросс-энтропийная функция потерь – Cross-Entropy).

Исходя из выбранных целевых функций, доказываем эквивалентность максимизации функции правдоподобия распределения входных данных и минимизации суммарной квадратичной ошибки сети при использовании линейных нейронов в параметрическом пространстве ограниченной машины Больцмана:

Теорема. Максимизация функции правдоподобия распределения входных данных $P(x)$ эквивалентна минимизации кросс-энтропийной функции и специальному случаю минимизации среднеквадратичной ошибки в одном и том же параметрическом пространстве ограниченной машины Больцмана:

$$\max \ln P(x) = \min CE_s = \min E_s.$$

Таким образом, очевидное преимущество в использовании MSE в качестве функции ошибок перед применением функции логарифмического правдоподобия и кросс-энтропийной функции в том, что при использовании MSE могут быть получены как линейные, так и нелинейные правила обучения RBM.

В результате доказательства теоремы были получены различные правила модификации весов и порогов нейронной сети, которые являются базисом предлагаемого метода обучения:

- 1) правила обучения для случаев CD-1 и CD- k (нелинейный случай);
- 2) правила обучения для случаев CD-1 и CD- k (линейный случай);
- 3) правила для группового обучения (случаи CD-1, CD- k) для нелинейного случая;
- 4) правила для группового обучения (случаи CD-1, CD- k) для линейного случая.

Так, например, для нелинейной ограниченной машины Больцмана правила модификации синаптических весов и порогов в случае CD-1 будут следующими:

$$\begin{aligned} w_{ij}(t+1) &= w_{ij}(t) - \\ &\quad - \alpha((y_j(1) - y_j(0))F'(S_j(1))x_i(1) + (x_i(1) - x_i(0))F'(S_i(1))y_j(0)), \\ T_i(t+1) &= T_i(t) - \alpha(x_i(1) - x_i(0))F'(S_i(1)), \\ T_j(t+1) &= T_j(t) - \alpha(y_j(1) - y_j(0))F'(S_j(1)), \end{aligned}$$

где $F'(x)$ определяет производную функции активации,

$S_i(1)$ – взвешенная сумма, посчитанная для i -го нейрона видимого слоя,

$S_j(1)$ – взвешенная сумма, посчитанная для j -го нейрона скрытого слоя.

Также в работе приводятся правила предобучения для сверточной ограниченной машины Больцмана (CRBM – Convolutional Restricted Boltzmann Machine), соответствующие представленным выше случаям.

В данной главе предлагается алгоритм для редуцирования (сокращения числа) параметров полносвязной и сверточной нейронных сетей, который основывается на применении метода предобучения с использованием ограниченной машины Больцмана. Данный алгоритм позволяет уменьшить количество параметров сети, ограничив его некоторым эффективным подмножеством.

В *третьей главе* приводится описание и результаты проведенных вычислительных экспериментов, обосновывающих предложенный метод предобучения. На общеизвестных выборках (MNIST, CIFAR-10, CIFAR-100) проводится сравнение обобщающей способности моделей, полученных примени-

ем классического и предложенного методов предобучения глубоких нейронных сетей. Также в этой главе описываются экспериментальные результаты, полученные для описанного алгоритма редуцирования параметров моделей. Проведенные эксперименты демонстрируют эффективность разработанных подходов.

Для выборки MNIST, сверточной НС с архитектурой слоев 40X5X5 – 40X5X5 – 640X320 – 320X160 – 160X10 и параметров обучения, представленных в таблице 1, были получены результаты для предложенного метода предобучения глубокой нейронной сети, представленные в таблице 2.

Таблица 1 – MNIST: основные параметры обучения

Этап	Параметр	Значение
Предобучение	Скорость обучения	0,000125
	Размер мини-батча	128
	Моментный параметр	[0,5; 0,9]
	Количество эпох обучения	30
Обучение	Скорость обучения	0,001
	Размер мини-батча	128
	Моментный параметр	0,9
	Количество эпох обучения	50

Таблица 2 – MNIST: результаты обучения

Метод обучения	Эффективность, %
BP	99.367
REBA	99.371
HREBA	99.458
C-RBM	99.447

Лучший средний показатель был достигнут гибридным методом HREBA, сочетающим в себе предобучение классическим и предложенным подходом, при этом максимальная эффективность была получена этим же методом и составила **99.53** %.

Для выборки MNIST, архитектуры 784-800-800-10 и параметров обучения, представленных в таблице 3, были получены результаты для редуцированной версии сети, представленные в таблице 4.

В таблице 4 **количество параметров** определяет количество параметров, сохранившихся после выполнения редуцирования; **редуцировано параметров** – процент редуцированных параметров модели.

На основе предложенных методов была осуществлена разработка нейросетевой системы компьютерного зрения и решены следующие прикладные задачи, описанные в *четвертой главе*:

Таблица 3 – Основные параметры обучения

Этап	Параметр	Значение
Обучение	Скорость обучения	0.05-0.1
	Размер мини-батча	100
	Моментный параметр	0.9
	Количество эпох обучения	50-100
Предобучение	Скорость обучения	0.05-0.2
	Размер мини-батча	32-100
	Моментный параметр	[0.5, 0.9]
	Количество эпох обучения	10

Таблица 4 – 784-800-800-10, MNIST

Тип	Эффективность, %, C-RBM / REBA	Количество параметров, C-RBM / REBA	Редуцировано параметров, %, C-RBM / REBA
без редуц.	98.63 / 98.33	1276810 / 1276810	0/0
t=0.2	98.61 / 98.27	233760 / 279635	81.69 / 78.1
t=0.5	98.03 / 98.05	32524 / 32817	97.45 / 97.43
t=0.8	97.1 / 96.48	17061 / 12217	98.66 / 99.04

- 1) задача обнаружения солнечных панелей на аэрофотоснимках;
- 2) задача распознавания маркировки продукта на производственной линии.

Указанные задачи решались как с использованием общеизвестных предобученных архитектур нейронных сетей, так и с применением собственных архитектур, предобученных с помощью предложенного метода.

1. Предлагаемое решение **задачи обнаружения солнечных панелей** включает два основных компонента, в составе которых используются предобученные глубокие нейронные сети:

- 1) классификатор для оценки наличия солнечной панели на аэрофотоснимке;
- 2) детектор для локализации солнечной панели.

Алгоритм решения представлен на рисунке 1.

Использование двухэтапности в решении поставленной задачи позволяет существенно ускорить обработку изображений, большую часть которых занимают объекты, отличные от искомым (солнечных панелей), т.к. нейросетевые модели, используемые на этапе локализации, как правило, более ресурсоемки. Таким образом, модель первого этапа выполняет роль фильтра для последующей обработки. Тем самым предлагаемый алгоритм можно применять для последовательной обработки больших изображений.

Для реализации *первого компонента* (классификатора) применялась

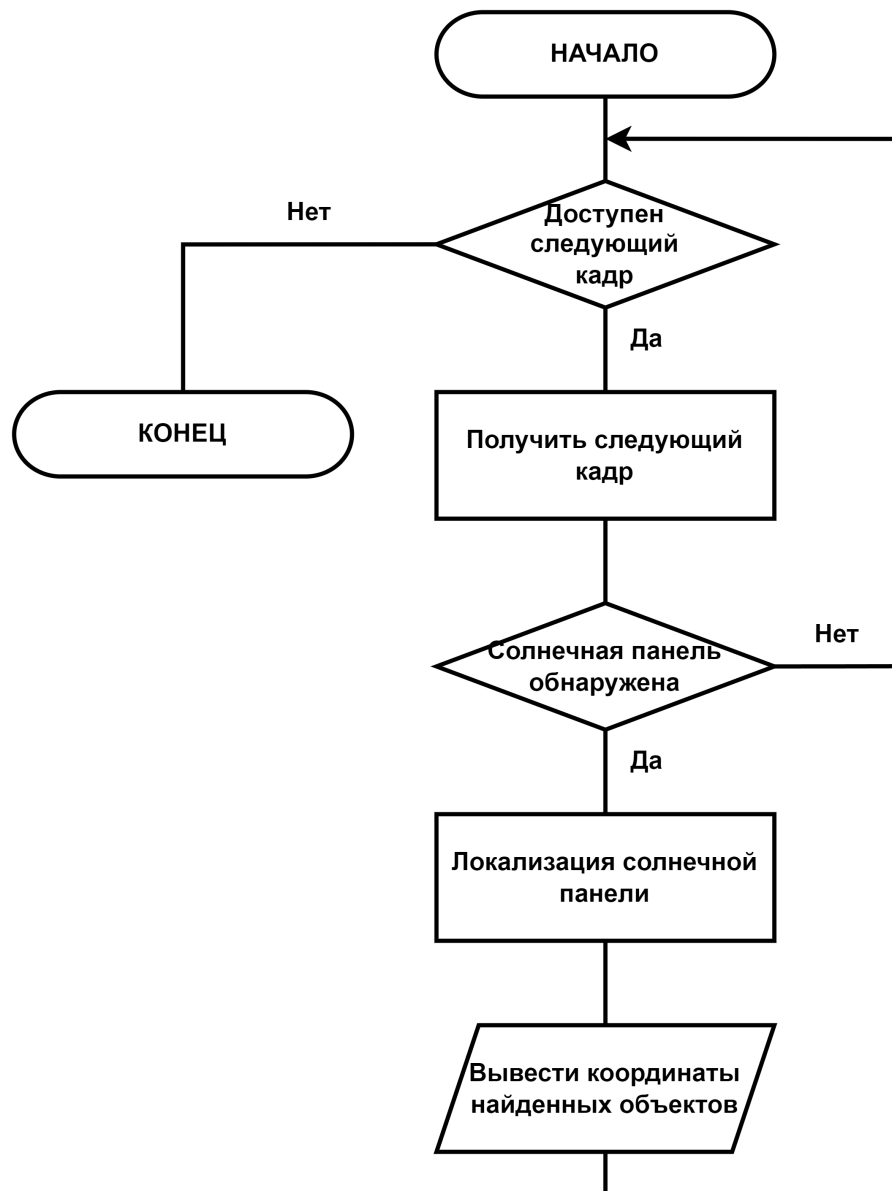


Рисунок 1 – Блок-схема решения задачи обнаружения солнечных панелей

сверточная нейронная сеть, архитектура которой изображена на рисунке 2.

Для обучения были использованы изображения, полученные из системы Google Maps. Для обучения и тестирования была использована выборка из 3347 3-канальных изображений размером 200X200 пикселей (из которых 1643 содержали солнечные панели, а 1704 не содержали). Полная выборка делилась на обучающую и тестирующую выборки в соотношении 4 к 1.

На рисунке 3 отображены примеры используемых изображений.

После обучения классификатора была получена точность обнаружения солнечных панелей в **87.46%** (таблица 5 представляет confusion matrix, вычисленную для полученного бинарного классификатора на тестовой выборке).

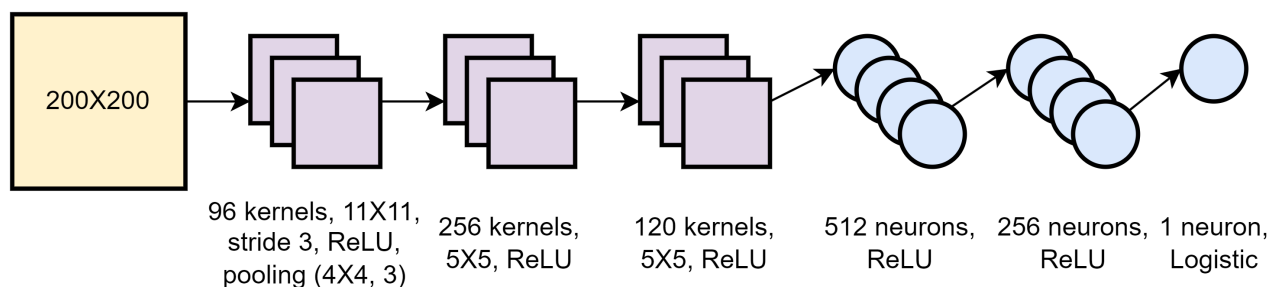


Рисунок 2 – Сверточная нейронная сеть для определения наличия солнечных панелей



Рисунок 3 – Примеры изображений, используемых для обучения

Таблица 5 – Confusion matrix для обученного классификатора

	Предсказано отсутствие	Предсказано наличие
Фактически отсутствует	325	32
Фактически присутствует	52	261

Эффективность классификатора, обученного без процедуры предобучения, составила **80,15%** на исследуемой выборке. Таким образом, использование предлагаемого метода предобучения позволило добиться прироста эффективности в **7,31%**.

Для реализации *детектора солнечной панели* применялась глубокая нейронная сеть Faster-RCNN (Faster Region-based Convolutional Neural Network) с классификатором ResNet-50, предобученным на выборке COCO.

2. Для задачи **распознавания маркировки** использовался конвейер (пайплайн) нейронных сетей, каждая из которых решает определенную

подзадачу:

- 1) оценка положения продукта в кадре;
- 2) обнаружение продукта и маркировки;
- 3) распознавание маркировки;
- 4) «сборка» маркировки и ее проверка.

Для достижения поставленных задач применялся конвейер, представленный на рисунке 4.

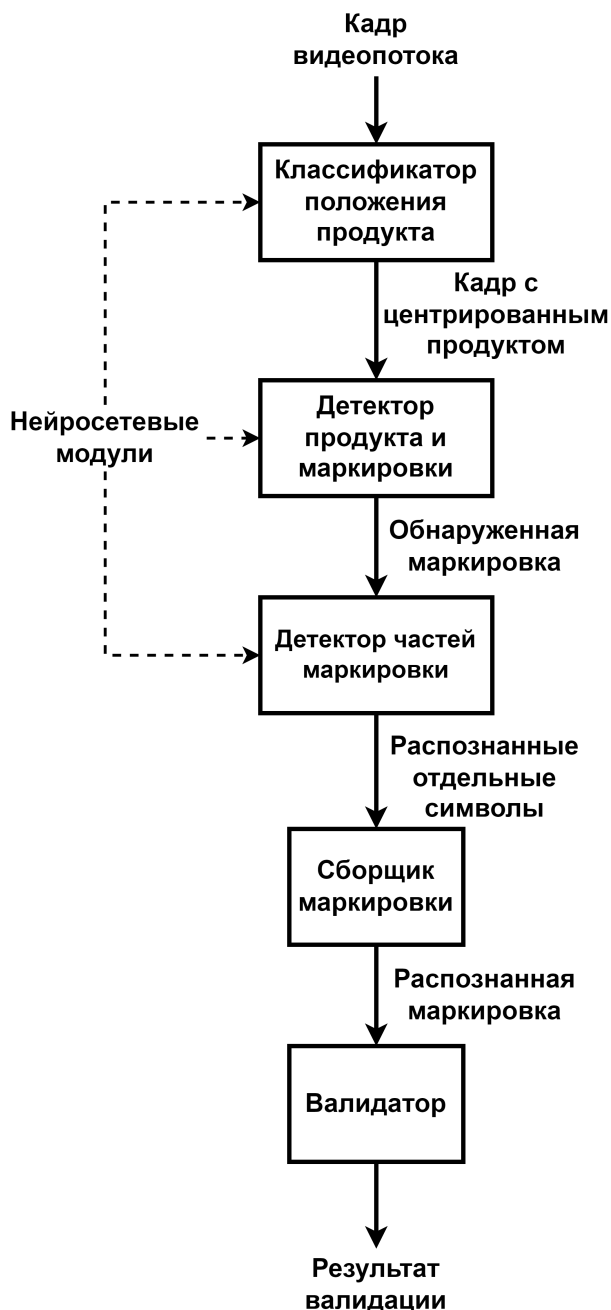


Рисунок 4 – Конвейер нейросетевых моделей для распознавания маркировки

Точность распознавания классификатора положения продукта составила **93.27%**. Для обучения данной модели применялось предобучение по предла-

гаемому методу REBA. Для классификатора, обученного без использования предобучения эффективность составила **88,03%**, таким образом, использование предобучения позволило увеличить эффективность распознавания на **5,24%** на исследуемой выборке.

Оба детектора (продуктов/маркировок и частей маркировки) обучались на основе предобученных моделей (использовалось предобучение I типа).

Итоговая эффективность детектора товара и маркировки составила **99%** (**mAP = 0.99**), детектора отдельных цифр – **92%** (**mAP = 0.92**). Результаты эффективности распознавания отдельных цифр представлены в таблице 6.

Таблица 6 – Эффективность обнаружения отдельных классов цифр

Class label	AP
0	0.9218
1.	0.9107
2.	0.9354
3.	0.9286
4.	0.9265
5.	0.9137
6.	0.9274
7.	0.9167
8.	0.9646
9.	0.8975
mAP	0.92429

Результаты работы детектора продукта и маркировки и детектора частей маркировки изображены на рисунках 5 и 6.

Применение предложенного конвейера нейросетевых моделей позволило осуществить независимую настройку и повысить эффективность работы отдельных моделей. Помимо этого, использование предлагаемого классификатора для оценки положения продукта в кадре позволило уменьшить количество обрабатываемых кадров видеопотока и увеличить скорость работы конвейера. Предложенный алгоритм может работать с разными типами маркировки (DataMatrix, QR, алфавитно-цифровая и т.д.) при условии обучения необходимых моделей распознавания.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

1. Установлена эквивалентность задач максимизации функции правдоподобия распределения входных данных, минимизации суммарной квадра-

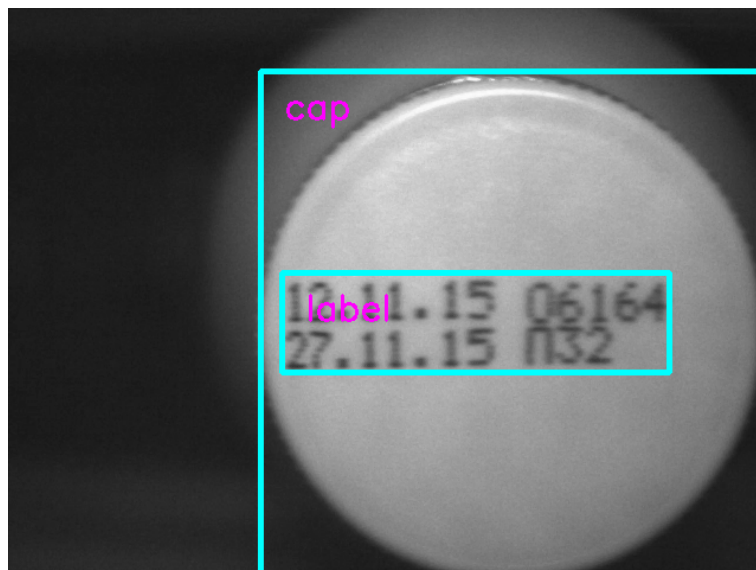


Рисунок 5 – Обнаруженный продукт и маркировка

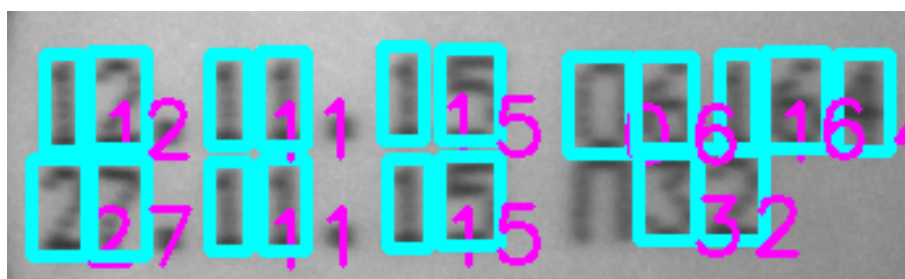


Рисунок 6 – Обнаруженные цифры в маркировке

точной ошибки сети при использовании линейных нейронов и минимизации кросс-энтропийной функции ошибки сети в пространстве синаптических связей ограниченной машины Больцмана. Из полученных теоретических результатов следует, что природа неконтролируемого обучения в RBM-сети является идентичной при использовании различных целевых функций [1–3, 5, 9, 16, 18].

2. Разработан метод обучения ограниченной машины Больцмана на основе доказанных эквивалентностей, применение которого для предобучения ГНС позволяет расширить класс обучаемых моделей и повысить обобщающую способность ГНС. Разработанный метод применен для обучения глубоких нейронных сетей и протестирован на выборках MNIST, CIFAR-10, CIFAR-100. Показано, что предложенный метод обладает большей эффективностью, чем классический [1–3, 5, 9, 14–18, 27–30].

3. Предложен алгоритм редуцирования параметров глубокой нейронной сети, базирующийся на неконтролируемом предобучении сети и позволяющий сократить количество настраиваемых параметров сети и упростить ее архитектуру без потери обобщающей способности. Проведены вычислительные

эксперименты, доказывающие эффективность предложенного метода [7,8,26].

4. Разработана нейросетевая система компьютерного зрения, основанная на предлагаемых методах предобучения и редуцирования, с помощью которой решены прикладные задачи компьютерного зрения, а именно:

- задача обнаружения солнечных панелей на аэрофотоснимках – предлагаемое решение позволяет повысить эффективность при определении наличия солнечной панели на 7,31% с последующей локализацией солнечной панели [13, 19, 22];
- задача распознавания маркировки продукта на производственной линии – предлагаемое решение позволяет повысить эффективность при оценке положения продукта с маркировкой в кадре на 5,24% с последующей локализацией маркировки с точностью 99% и обнаружением ее отдельных составляющих (цифр) с точностью 92%. Предлагаемое решение позволяет осуществлять обработку изображений в реальном времени [4, 6, 10–12, 20, 21, 23–25].

Рекомендации по практическому использованию результатов

Научные и практические результаты диссертационной работы использованы в научно-исследовательских работах, учебном процессе, а также в ряде прикладных систем.

Разработанные алгоритмы обнаружения и локализации солнечных панелей на аэрофотоснимках, методы предобучения нейросетевых моделей и инструментальные средства внедрены и используются в компании ООО «Intelligent Semantic Systems» при разработке интеллектуальных систем в составе подсистем компьютерного зрения.

Предложенный метод предобучения глубоких нейронных сетей применялся при обучении модели классификатора кадров для нейросетевой системы распознавания маркировки продукта на производственной линии ОАО «Савушкин продукт».

Кроме того, научные и практические результаты диссертационной работы используются в учебном процессе учреждения образования «Брестский государственный технический университет».

Реализация программных модулей осуществлялась с использованием свободного ПО и может быть непосредственно применена при разработке отечественных проектов в области компьютерного зрения без необходимости приобретения дорогостоящих программных средств.

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ УЧЕНОЙ СТЕПЕНИ

Статьи в рецензируемых научных журналах

1. Головки, В.А. Перцептроны и нейронные сети глубокого доверия : обучение и применение / В.А. Головки, А.А. Крощенко // Вестник Брестского государственного технического университета. Физика, математика, информатика. — 2014. — № 5 (89). — С. 2–12.
2. Golovko, V. The Nature of Unsupervised Learning in Deep Neural Networks : A New Understanding and Novel Approach / V. Golovko, A. Kroshchanka, D. Treadwell // Optical Memory and Neural Networks. — 2016. — Vol. 25, № 3. — P. 127–141.
3. Головки, В.А. Теория глубокого обучения : конвенциональный и новый подход / В.А. Головки, А.А. Крощенко, М.В. Хацкевич // Вестник Брестского государственного технического университета. Физика, математика, информатика. — 2016. — № 5 (101). — С. 7–16.
4. Крощенко, А.А. Реализация нейросетевой системы распознавания маркировки продукции / А.А. Крощенко, В.А. Головки // Вестник Брестского государственного технического университета. Физика, математика, информатика. — 2019. — № 5 (118). — С. 9–12.
5. Golovko, V.A. Deep Neural Networks : Selected Aspects of Learning and Application / V.A. Golovko, A.A. Kroshchanka, E.V. Mikhno // Pattern Recognition and Image Analysis. — 2021. — Vol. 31, № 1. — P. 132–143.
6. Kroshchanka, A. Neural network component of the product marking recognition system on the production line / A. Kroshchanka, D. Ivaniuk // Open Semantic Technologies for Intelligent Systems : research papers collection / Belar. State Univ. of Informatics and Radioelectr.; ed. : V.V. Golenkov (ed.-in-chief) [et al.]. — Minsk, 2021. — Iss. 5. — P. 219–224.
7. Kroshchanka, A.A. Method for Reducing Neural-Network Models of Computer Vision / A.A. Kroshchanka, V.A. Golovko, M. Chodyka // Pattern Recognition and Image Analysis. — Berlin, Heidelberg : Springer-Verlag, 2022. — Vol. 32, № 2. — P. 294–300.
8. Kroshchanka, A.A. Reduction of Neural Network Models in Intelligent Computer Systems of a New Generation / A. Kroshchanka // Open Semantic Technologies for Intelligent Systems (OSTIS-2023) : research papers collection / Belar. State Univ. of Informatics and Radioelectr.; ed. : V.V. Golenkov (ed.-in-chief) [et al.]. — Minsk, 2023. — Iss. 7. — P. 127–132.

Статьи в других научных изданиях

9. Головки, В.А. Метод обучения нейронной сети глубокого доверия и применение для визуализации данных / В.А. Головки, А.А. Крощенко // Комп'ютерно-інтегровані технології: освіта, наука, виробництво. — Луцьк, 2015. — № 19. — С. 6–12.

10. Интеграция искусственных нейронных сетей с базами знаний / В.А. Головки, В.В. Голенков, В.П. Ивашенко, В.В. Таберко, Д.С. Иванюк, А.А. Крощенко, М.В. Ковалёв // Онтология проектирования. — 2018. — Т. 8. — № 3 (29). — С. 366–386.

11. Principles of decision-making systems building based on the integration of neural networks and semantic models / V. Golovko, A. Kroshchanka, V. Ivashenko, M. Kovalev, V. Taberko, D. Ivaniuk // Open Semantic Technologies for Intelligent Systems (OSTIS-2019) : research papers collection / Belar. State Univ. of Informatics and Radioelectr.; ed. : V.V. Golenkov (ed.-in-chief) [et al.]. — Minsk, 2019. — Iss. 3. — P. 91–102.

12. Implementation of an intelligent decision support system to accompany the manufacturing process / V. Golovko, A. Kroshchanka, M. Kovalev, V. Taberko, D. Ivaniuk // Open Semantic Technologies for Intelligent Systems (OSTIS-2020) : research papers collection / Belar. State Univ. of Informatics and Radioelectr.; ed. : V.V. Golenkov (ed.-in-chief) [et al.]. — Minsk, 2020. — Iss. 4. — P. 175–182.

13. Deep Convolutional Neural Network for Detection of Solar Panels / V. Golovko, A. Kroshchanka, E. Mikhno, M. Komar, A. Sachenko // Data-Centric Business and Applications. ICT Systems-Theory, Radio-Electronics, Information Technologies and Cybersecurity. — Cham: Springer International Publishing, 2021. — P. 371–389.

Статьи в сборниках материалов научных конференций

14. A Learning Technique for Deep Belief Neural Networks / V. Golovko, A. Kroshchanka, U. Rubanau, S. Jankowski // Neural Networks and Artificial Intelligence : proc. of the 8th Internat. Conf. ICNNAI 2014, Brest, Belarus, June 3-6, 2014 / V. Golovko, A. Imada (eds.). — Springer, 2014. — P. 136–146.

15. Головки, В.А. Применение нейронных сетей глубокого доверия для выделения семантически значимых признаков / В.А. Головки, А.А. Крощенко // Открытые семантические технологии проектирования интеллектуальных систем : материалы V Междунар. науч.-технич. конф. OSTIS-2015, Минск, 19-21 февраля 2015 г. / УО «Бел. гос. ун-т информатики и радиоэлектроники», ГУ «Администрация Парка высоких технологий»; редкол.: В.В. Голенков (отв. ред.) [и др.]. — Минск, 2015. — P. 481–486.

16. A New Technique for Restricted Boltzmann Machine Learning [Electronic resource] / V. Golovko, A. Kroshchanka, V. Turchenko, S. Jankowski, D. Treadwell // Proceedings of the 8th IEEE International Conference of Intelligent Data Acquisition and Advanced Computing Systems : Technology and Applications (IDAACS), Warsaw, Poland, Sept. 24-26, 2015 / Research Institute for Intelligent Computer Systems, Ternopil National Economic University and V.M. Glushkov Inst. of Cybernetics, National Academy for Sciences of Ukraine, Warsaw University of Technology. — P. 182–186. — Mode of access : <https://ieeexplore.ieee.org/document/7340725>. — Date of access : 05.06.2023.

17. Крощенко, А.А. Применение нейронных сетей глубокого доверия в интеллектуальном анализе данных / А.А. Крощенко // Современные проблемы математики и вычислительной техники : сб. материалов IX Респ. науч. конф. молодых ученых и студентов, Брест, 19-21 ноября 2015 г. / Мин. обр. Респ. Бел., УО «Брест. гос. технич. ун-т»; редкол.: В.С. Рубанов (гл. ред.) [и др.]. — Брест, 2015. — С. 12–14.

18. Golovko, V. Theoretical Notes on Unsupervised Learning in Deep Neural Networks [Electronic resource] / V. Golovko, A. Kroshchanka // Proceedings of the 8th Internat. Joint Conf. on Computational Intelligence (IJCCI 2016), Porto, Portugal, Nov. 9-11 2016. — P. 91–96. — Mode of access : <https://scitepress.org/papers/2016/60843/60843.pdf>. — Date of access : 05.06.2023.

19. Convolutional Neural Network Based Solar Photovoltaic Panel Detection in Satellite Photos [Electronic resource] / V. Golovko, S. Bezobrazov, A. Kroshchanka, A. Sachenko, M. Komar, A. Karachka // The 9th IEEE Internat. Conf. on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, Bucharest, Romania, Sept. 21-23 2017. — P. 14–19. — Mode of access : <https://ieeexplore.ieee.org/document/8094501>. — Date of access : 05.06.2023.

20. Integration of artificial neural networks and knowledge bases / V.A. Golovko, A.A. Kroshchanka, V.V. Golenkov, V.P. Ivashenko, M.V. Kovalev, V.V. Taberko, D.S. Ivaniuk // Открытые семантические технологии проектирования интеллектуальных систем : материалы VIII Междунар. науч.-технич. конф., Минск, 15-17 февраля 2018 г. / Мин. обр. Респ. Бел., УО «Бел. гос. ун-т информатики и радиоэлектроники»; редкол.: В.В. Голенков (гл. ред.) [и др.]. — Минск, 2018. — Вып. 2. — С. 133–146.

21. Головки, В.А. Нейросетевые модели глубокого обучения для решения задач распознавания объектов на изображении / В.А. Головки, А.А. Крощенко // Вычислительные методы, модели и образовательные технологии : сб. материалов VII Междунар. науч.-практич. конф., Брест, 19 октября 2018 г.

/ Брест. гос. ун-т имени А.С. Пушкина; под общ. ред. А.А. Козинского. — Брест, 2018. — С. 3–5.

22. Golovko, V. Development of Solar Panels Detector [Electronic resource] / V. Golovko, A. Kroshchanka, S. Bezobrazov, A. Sachenko, M. Komar, O. Novosad // 2018 International Scientific-Practical Conference «Problems of Infocommunications. Science and Technology (PIC S&T 2018)», Kharkiv, Ukraine, 9-12 October 2018 / Institute of Electrical and Electronics Engineers, Inc. — Mode of access : <https://www.semanticscholar.org/paper/Development-of-Solar-Panels-Detector-Golovko-Kroshchanka/914c0b8c159c64100609fc8455636b1e3f8568cb>. — Date of access : 05.06.2023.

23. Golovko, V. Brands and caps labeling recognition in images using deep learning [Electronic resource] / V. Golovko, A. Kroshchanka, E. Mikhno // Pattern Recognition and Information Processing : revised selected papers of the 14th Internat. Conf. PRIP 2019, Minsk, May 21-23 2019; Eds. : S.V. Ablameyko, V.V. Krasnoproshin, M.M. Lukashevich. — P. 35–51. — Mode of access : https://www.researchgate.net/publication/337459978_Brands_and_Caps_Labeling_Recognition_in_Images_Using_Deep_Learning. — Date of access : 05.06.2023.

24. Головки, В.А. Обнаружение и распознавание маркировки продукции с помощью нейросетевых алгоритмов / В.А. Головки, А.А. Крощенко // Вычислительные методы, модели и образовательные технологии: сб. материалов VII Междунар. науч.-практич. конф., Брест, 18 октября 2019 г. / Брест. гос. ун-т имени А.С. Пушкина; под общ. ред. А.А. Козинского. — Брест, 2019. — С. 3–6.

25. Neuro-Symbolic Artificial Intelligence: Application for Control the Quality of Product Labeling [Electronic resource] / V. Golovko, A. Kroshchanka, M. Kovalev, V. Taberko, D. Ivaniuk // Open Semantic Technologies for Intelligent Systems : revised selected papers of the 10th Intern. Conf. (OSTIS 2020), Minsk, Febr. 19-22, 2020; / Eds. : V. Golenkov, V. Krasnoproshin, V. Golovko, E. Azarov. — P. 81–101. — Mode of access : https://libeldoc.bsuir.by/bitstream/123456789/42395/1/Golovko_Neuro_Symbolic.pdf. — Date of access : 05.06.2023.

26. Kroshchanka, A. The Reduction of Fully Connected Neural Network Parameters Using the Pre-training Technique / A. Kroshchanka, V. Golovko // Proceedings of the 11th IEEE Intern. Conf. on Intelligent Data Acquisition and Advanced Computing Systems : Technology and Applications (IDAACS), Cracow, Poland, September 22-25, 2021 / Cracow University of Technology [et al.]. — Vol. 2. — P. 937–941. — Mode of access : https://www.researchgate.net/publication/357613109_The_Reduction_of_Fully_Connected_Neural_

Тезисы

27. Крощенко, А.А. Методы глубокого обучения нейронных сетей / А.А. Крощенко // Вычислительные методы, модели и образовательные технологии : сб. матер. региональной науч.-практич. конф., Брест, 22-23 окт. 2013 г. / Брест. гос. ун-т им. А.С. Пушкина; под общ. ред. О.В. Матысика. — Брест, 2013. — С. 21–22.

28. Головки, В.А. Об одном методе обучения нейронных сетей глубокого доверия / В.А. Головки, А.А. Крощенко // Вычислительные методы, модели и образовательные технологии : сб. матер. Междунар. науч.-практич. конф., Брест, 15-16 окт. 2014 г. / Брест. гос. ун-т им. А.С. Пушкина; под общ. ред. О.В. Матысика. — Брест, 2014. — С. 98–99.

29. Головки, В.А. Применение нейронных сетей глубокого доверия в интеллектуальном анализе данных / В.А. Головки, А.А. Крощенко // Вычислительные методы, модели и образовательные технологии : сб. матер. Междунар. науч.-практич. конф., Брест, 22 окт. 2015 г. / Брест. гос. ун-т им. А.С. Пушкина; под общ. ред. О.В. Матысика. — Брест, 2015. — С. 97–98.

30. Крощенко, А.А. Применение глубокой нейронной сети для решения задачи распознавания образов / А.А. Крощенко // Вычислительные методы, модели и образовательные технологии : сб. матер. Междунар. науч.-практич. конф., Брест, 21 окт. 2016 г. / Брест. гос. ун-т им. А.С. Пушкина; под общ. ред. О.В. Матысика. — Брест, 2016. — С. 132–133.

РЕЗЮМЕ

Крощенко Александр Александрович

МЕТОДЫ ОБУЧЕНИЯ ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ЗАДАЧ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Ключевые слова: глубокие нейронные сети, компьютерное зрение, сверточные нейронные сети, ограниченная машина Больцмана, детекция объектов.

Целью работы является разработка методов и алгоритмов для обучения глубоких нейронных сетей, применяемых для решения прикладных задач компьютерного зрения.

Принципы и методы исследования: в работе использовались методы теории нейронных сетей и машинного обучения, методы обработки изображений, детекции и распознавания образов.

Полученные результаты и их новизна. Установлена эквивалентность задач максимизации функции правдоподобия распределения входных данных, минимизации кросс-энтропийной функции ошибки и суммарной квадратичной ошибки при использовании линейных нейронов в пространстве синаптических связей ограниченной машины Больцмана, что позволяет учитывать нелинейную природу нейронных элементов.

Разработан метод обучения ограниченной машины Больцмана на основе доказанных эквивалентностей, применение которого для предобучения ГНС позволяет расширить класс обучаемых моделей и повысить обобщающую способность ГНС.

Разработан алгоритм редуцирования параметров нейросетевой модели, основывающийся на неконтролируемом предобучении, что позволяет уменьшить количество настраиваемых параметров модели без потери обобщающей способности.

Разработана нейросетевая система компьютерного зрения, которая основана на предлагаемом методе предобучения нейросетевых моделей, применение которой позволяет улучшить качество решения задач классификации. Продемонстрирована эффективность системы на примерах решения задач обнаружения солнечных панелей на аэрофотоснимках и распознавания маркировки продуктов на производственной линии.

Рекомендации к использованию и область применения: интеллектуальные системы видеонаблюдения, используемые для анализа маркировки продукта на производственных линиях и анализа аэрофотоснимков.

РЭЗІЮМЭ

Крошчанка Аляксандр Аляксандравіч

МЕТАДЫ НАВУЧАННЯ ГЛЫБОКІХ НЕЙРОНАВЫХ СЕТАК ДЛЯ ЗАДАЧ КАМП'ЮТЭРНАГА ЗРОКУ

Ключавыя словы: глыбокія нейронавыя сеткі, камп'ютэрны зрок, згортачныя нейронавыя сеткі, гібрыдныя інтэлектуальныя сістэмы, абмежаваная машына Больцмана, дэтэкцыя аб'ектаў.

Мэта працы: распрацоўка метадаў і алгарытмаў для навучання глыбокіх нейронавых сетак, якія выкарыстоўваюцца для вырашэння прыкладных задач камп'ютэрнага зроку.

Прынцыпы і метады даследавання: у працы выкарыстоўваліся метады тэорыі нейронавых сетак і машыннага навучання, метады апрацоўкі відарысаў, дэтэкцыі і распазнання выяў.

Атрыманыя вынікі і іх навізна. Усталяванне эквівалентнасці задач максімізацыі функцыі праўдападабенства размеркавання ўваходных дадзеных, мінімізацыі крос-энтрапійнай функцыі памылкі і сумарнай квадратычнай памылкі пры выкарыстанні лінейных нейронаў у прасторы сінаптычных сувязяў абмежаванай машыны Больцмана, што дазваляе ўлічваць нелінейную прыроду нейронавых элементаў.

Распрацаваны метады навучання абмежаванай машыны Больцмана на аснове даказаных эквівалентнасцяў, ужыванне якога для праднавучання ГНС дазваляе пашырыць клас выкарыстоўваных мадэляў і павысіць абагульняльную здольнасць ГНС.

Распрацаваны алгарытм скарачэння параметраў нейрасеткавай мадэлі, які засноўваецца на некантралюемым праднавучанні, што дазваляе паменшыць колькасць наладжвальных параметраў мадэлі без страты абагульняючай здольнасці.

Распрацавана нейрасеткавая сістэма камп'ютэрнага зроку, якая заснавана на прапанаваным метады праднавучання нейрасеткавых мадэляў, прымяненне якой дазваляе палепшыць якасць рашэння задач класіфікацыі. Прадэманстравана эфектыўнасць сістэмы на прыкладах рашэння задач выяўлення сонечных панэляў на аэрафотаздымках і распазнавання маркіроўкі прадуктаў на вытворчай лініі.

Рэкамендацыі па выкарыстанні і вобласць ужывання: інтэлектуальныя сістэмы відэаназірання, якія выкарыстоўваюцца для аналізу маркіроўкі прадукта на вытворчых лініях і аналіза аэрафотаздымкаў.

SUMMARY

Kroshchenko Aleksandr Aleksandrovich

METHODS FOR TRAINING DEEP NEURAL NETWORKS FOR COMPUTER VISION PROBLEMS

Keywords: deep neural networks, computer vision, convolutional neural networks, hybrid intelligent systems, restricted Boltzmann machine, object detection.

The purpose of the work is development of methods and algorithms for training deep neural networks used to solve applied problems of computer vision.

Principles and methods of research: the methods of the theory of neural networks and machine learning, methods of image processing, detection and pattern recognition were used in the work.

The obtained results and their novelty. The equivalence of the problems of maximizing the likelihood function of the distribution of input data, minimizing the cross-entropy error function and the total squared error when using linear neurons in the space of synaptic connections of a restricted Boltzmann machine has been established, which makes it possible to take into account the nonlinear nature of neural elements.

A method for training a restricted Boltzmann machine based on proven equivalences has been developed, the use of which for pre-training a DNN allows one to expand the class of trained models and increase the generalizing ability of the DNN.

An algorithm has been developed for pruning the parameters of a neural network model, based on unsupervised pre-training, which makes it possible to reduce the number of adjustable parameters of the model without loss of generalization ability.

A neural network computer vision system has been developed, which is based on the proposed method of pre-training neural network models, the use of which can improve the quality of solving classification problems. The effectiveness of the system is demonstrated by examples of solving the problems of detecting solar panels on aerial photographs and recognizing product labels on a production line.

Recommendations on the use and field of application: intelligent video surveillance systems used to analyze product labeling on production lines, analysis of aerial photographs, intelligent face identification systems with the possibility of automatic additional training.