

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 2023

Работа выполнена в учреждении образования «Брестский государственный технический университет».

Научный руководитель

Головко Владимир Адамович, доктор технических наук, профессор, заведующий кафедрой интеллектуальных информационных технологий учреждения образования «Брестский государственный технический университет»

Официальные оппоненты:

Оппонирующая организация

Защита состоится

С диссертацией можно ознакомиться в библиотеке

Автореферат разослан

Ученый секретарь совета
по защите диссертаций
канд. физ.-мат. наук, доцент

ФИО

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 применении 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 распознавание маркировки продукта на конвейерной линии и обнаружение солнечных панелей на аэрофотоснимках.

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 прикладные нейросетевые системы компьютерного зрения (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 доказательстве 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.

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.

Разработаны нейросетевые системы компьютерного зрения (система обнаружения солнечных панелей на аэрофотоснимках, система распознавания маркировки продуктов на производственной линии), которые основываются на нейросетевых моделях, предобученных предложенным методом, что позволяет повысить качество решения задач.

Положения, выносимые на защиту

1. 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.

2. 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.

3. Нейросетевые системы компьютерного зрения, базирующиеся на использовании предобученных моделей – система обнаружения солнечных панелей на аэрофотоснимках (включая снимки низкого разрешения) с точностью до 87,46% и система распознавания (в реальном режиме времени) маркировки продукта на производственной линии, что позволяет повысить качество решения задач.

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 результаты диссертационной работы докладывались и обсуждались на следующих международных конференциях: «International Conference on Neural Networks and Artificial Intelligence» (Брест, 2014); «Информационное, программное и техническое обеспечение систем управления организационно-технологическими комплексами» (Луцк, 2015); «8th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications» (Варшава, 2015); «International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)» (2018); «International Conference on Pattern Recognition and Information Processing (PRIP)» (2019); «Вычислительные методы, модели и образовательные технологии» (Брест, 2013, 2014, 2015, 2016, 2019); «Современные проблемы

математики и вычислительной техники» (Брест, 2015); «Открытые семантические технологии проектирования интеллектуальных систем» (Минск, 2015, 2018, 2019, 2020, 2021, 2023); «11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)» (Краков, 2021).

Опубликованность результатов диссертации

По теме диссертационного исследования опубликовано 30 научных работ, среди которых: 14 статей в научных изданиях в соответствии с пунктом 19 Положения о присуждении ученых степеней и присвоении ученых званий (9,39 авторских листа), 3 статей в других научных изданиях, 9 статей в сборниках материалов научных конференций и 4 тезиса.

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, четырех глав с краткими выводами по каждой главе, заключения, библиографического списка, списка публикаций автора и приложений.

В *первой главе* рассмотрено краткое введение в теорию искусственных нейронных сетей, дана классификация их типов. Определено понятие глубокой нейронной сети. Рассмотрены задачи обучения глубоких нейронных сетей и существующие методы обучения. *Вторая глава* посвящена рассмотрению разработанного метода обучения ограниченной машины Больцмана и метода предобучения глубоких нейронных сетей на его основе. Также в этой главе предлагается алгоритм редуцирования параметров нейросетевых моделей. В *третьей главе* приведены экспериментальные результаты, обосновывающие полученные теоретические результаты. В *четвертой главе* рассмотрено практическое применение разработанных методов в интеллектуальных системах компьютерного зрения.

Общий объем диссертации составляет 116 страниц, из которых 91 страниц основного текста, 44 рисунка на 32 страницах, 17 таблиц на 9 страницах, библиография из 120 источников, включая 30 публикаций автора, приложения на 14 страницах.

ОСНОВНАЯ ЧАСТЬ

Во *введении* обоснована актуальность темы диссертационной работы, сформулированы цель и задачи исследования.

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 ней-

ронных 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:

$$\frac{\partial \ln P(x)}{\partial w_{ij}} = E[x_i y_j]_{\text{data}} - E[x_i y_j]_{\text{model}}$$

$$\frac{\partial \ln P(x)}{\partial T_i} = E[x_i]_{\text{data}} - E[x_i]_{\text{model}}$$

$$\frac{\partial \ln P(x)}{\partial T_j} = E[y_j]_{\text{data}} - E[y_j]_{\text{model}}$$

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):

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

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

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

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 пространстве синаптических связей 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 пространстве синаптических весов 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:

$$w_{ij}(t+1) = w_{ij}(t) - \alpha((y_j(1) - y_j(0))F'(S_j(1))x_i(1) + (x_i(1) - x_i(0))F'(S_i(1))y_j(0)),$$

$$T_i(t+1) = T_i(t) - \alpha(x_i(1) - x_i(0))F'(S_i(1)),$$

$$T_j(t+1) = T_j(t) - \alpha(y_j(1) - y_j(0))F'(S_j(1)).$$

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 применением 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

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.

Разработанные теоретические методы позволили решить следующие прикладные задачи компьютерного зрения, описанные в *четвертой главе*:

- 1) Задача обнаружения солнечных панелей на аэрофотоснимках;
- 2) Распознавание маркировки продукта на производственной линии.

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

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

1. Предлагаемая система 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.

Использование двухэтапности в решении поставленной задачи позволяет существенно ускорить обработку изображений, большую часть которых занимают объекты, отличные от искоемых (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 задачи определения наличия солнечных панелей на аэрофотоснимках применялась сверточная нейронная сеть, архитектура которой изображена на рисунке 1.

Для обучения были использованы изображения, полученные фотографированием со спутника (Google Maps). Для обучения и тестирования

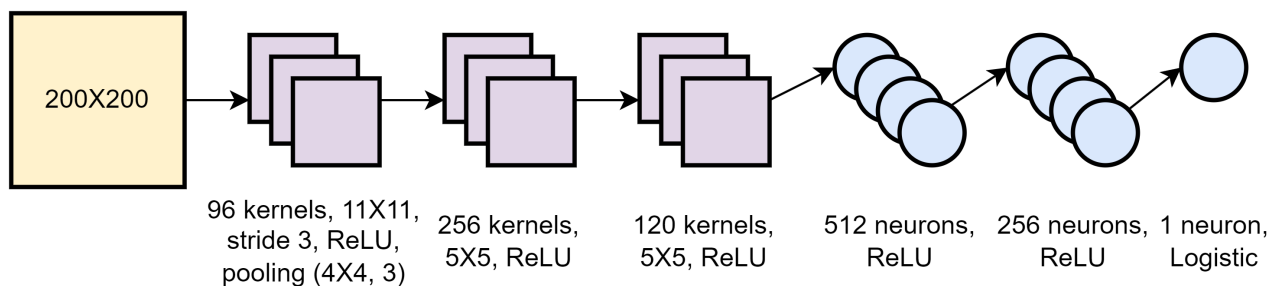


Рисунок 1. – Сверточная нейронная сеть для определения наличия солнечных панелей

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 2 отображены примеры используемых изображений.



Рисунок 2. – Примеры изображений, используемых для обучения

Сеть обучалась на протяжении 70 эпох методом обратного распространения ошибки, при этом были использованы следующие параметры: скорость обучения – 0,001, моментный параметр – 0,9, weight-decay – 0,0005, размер мини-батча – 20, dropout с вероятностью 0,5 для полносвязных слоев сверточной нейронной сети.

В результате проведенных экспериментов была получена точность обнаружения объектов в 87.46% (таблица 5 представляет confusion matrix, вычисленную для полученного бинарного классификатора на тестовой выборке).

Таблица 5. – Confusion matrix для обученного классификатора

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

2. Для задачи распознавания маркировки использовался конвейер 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 3.

Применение предложенного конвейера нейросетевых моделей позволило осуществить независимую настройку и повысить эффективность работы отдельных моделей. Помимо этого, применяемые модели для оценки положения продукта в кадре позволили уменьшить количество обрабатываемых кадров видеопотока и, как следствие, увеличить скорость работы конвейера. Предложенный алгоритм может работать с разными типами маркировки (DataMatrix, QR, алфавитно-цифровая и т.д.).

В *заключении* сформулированы основные научные результаты диссертации и рекомендации по их практическому применению. Приложение содержит акты внедрения результатов диссертации и исходный код, реализующий метод предобучения глубокой нейронной сети.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

1. Выявлена и доказана эквивалентность задач максимизации функции правдоподобия распределения входных данных, минимизации суммарной квадратичной 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 ([2], [3], [4], [5], [10], [18], [21]);

2. Разработан метод неконтролируемого предобучения глубоких нейронных сетей, базирующийся на минимизации квадратичной ошибки сети в скрытом и видимом слоях RBM-машины, что позволяет учитывать нелинейную природу нейронных элементов. Разработанный метод применен для

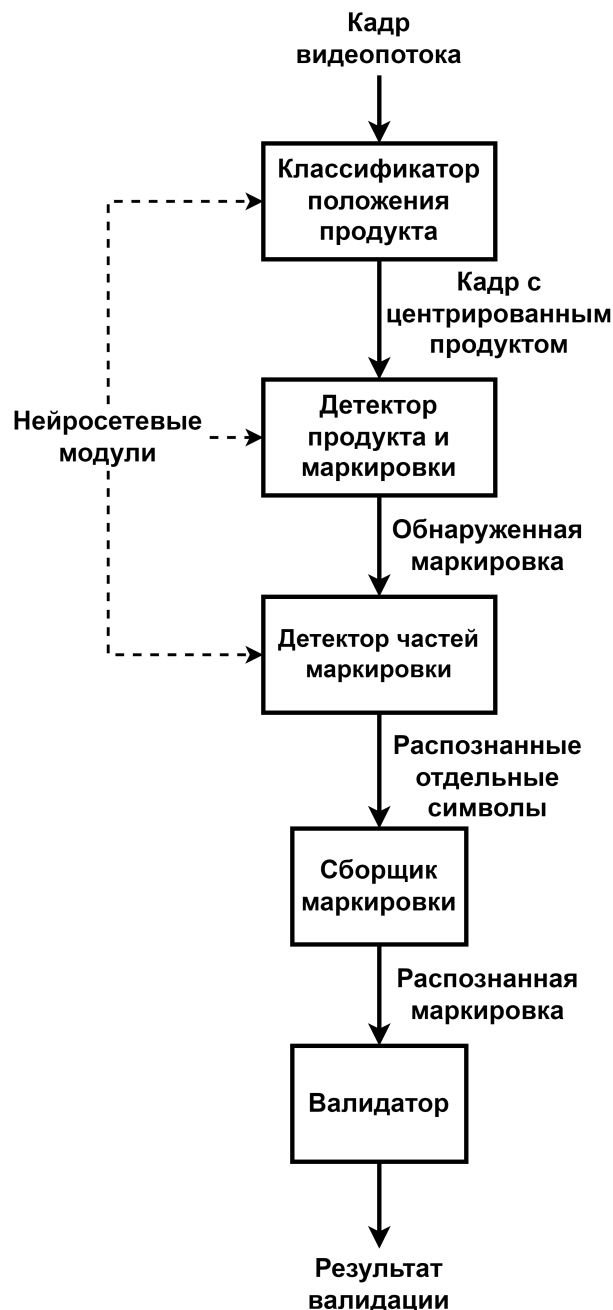


Рисунок 3. – Конвейер нейросетевых моделей для распознавания маркировки

обучения глубоких полносвязных и сверточных архитектур нейронных сетей и протестирован на выборках MNIST, CIFAR-10, CIFAR-100. Показано, что предложенный метод обладает большей эффективностью, чем классический ([1], [2], [3], [4], [5], [10], [18], [21], [27], [28], [19], [20], [29], [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 ([13], [26], [14]);

4. С использованием предложенного метода предобучения реализованы прикладные нейросетевые системы компьютерного зрения.

Разработана нейросетевая система обнаружения солнечных панелей на аэрофотоснимках, позволяющая обнаруживать солнечные панели с точностью 87,46% с возможностью использования фотографий низкого разрешения ([11], [22], [23]).

Разработана нейросетевая система распознавания маркировки продукта на производственной линии, базирующаяся на интеграции различных моделей глубоких сверточных нейронных сетей. Предложенный конвейер представляет собой цепочку взаимодействующих моделей нейросетевых классификаторов и детекторов, которые решают отдельные подзадачи обнаружения и распознавания маркировки или ее части. Подобная архитектура позволяет добавлять новые типы маркировок благодаря простой модульной структуре. Архитектуры нейросетей, используемые для построения конвейера, позволяют осуществлять обработку изображения в реальном времени ([6], [8], [7], [15], [24], [16], [25], [17], [9], [12]).

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».

Разработка проводилась с использованием свободного программного обеспечения и может быть использована в различных отечественных проектах без необходимости приобретения дорогостоящих программных средств.

СПИСОК ПУБЛИКАЦИЙ СОИСКАТЕЛЯ УЧЕНОЙ СТЕПЕНИ

Статьи в рецензируемых научных журналах

1. Golovko, V. A Learning Technique for Deep Belief Neural Networks / V. Golovko, A. Kroshchanka, U. Rubanau, S. Jankowski // Neural Networks and Artificial Intelligence. — Cham: Springer International Publishing, 2014. — Vol. 440. Communication in Computer and Information Science. — P. 136–146.
2. Головки, В. А. Персептроны и нейронные сети глубокого доверия: обучение и применение / В. А. Головки, А. А. Крощенко // Вестник Брестского государственного технического университета. — Брест, 2014. — Т. 5. — С. 2–12.
3. 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. — Вып. 19. — С. 6–12.
4. 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. — New York : Allerton Press, Inc., 2016. — P. 127–141.
5. Головки, В. А. Теория глубокого обучения: конвенциональный и новый подход / В. А. Головки, А. А. Крощенко, М. В. Хацкевич // Вестник Брестского государственного технического университета. — Брест, 2016. — № 5. — С. 7–16.
6. Головки, В. А. Интеграция искусственных нейронных сетей с базами знаний / В. А. Головки, В. В. Голенков, В. П. Ивашенко, В. В. Таберко, Д. С. Шаток, А. А. Крощенко, М. В. Ковалёв // Онтология проектирования. — EBSCO Publishing, 2018. — Т. 8. — № 3(29). — С. 366–386.
7. Golovko, V. Brands and caps labeling recognition in images using deep learning / V. Golovko, A. Kroshchanka, E. Mikhno // International Conference on Pattern Recognition and Information Processing. — Cham: Springer International Publishing, 2019. — P. 35–51.
8. Крощенко, А. А. Реализация нейросетевой системы распознавания маркировки продукции / А. А. Крощенко, В. А. Головки // Вестник Брестского государственного технического университета. — Брест, 2019. — № 5. — С. 9–12.
9. Golovko, V. Neuro-Symbolic Artificial Intelligence: Application for Control the Quality of Product Labeling / V. Golovko, A. Kroshchanka, M. Kovalev, V. Taberko, D. Ivaniuk // International Conference on Open Semantic Technologies for Intelligent Systems. — Cham: Springer International Publishing, 2020. — P. 81–101.

10. 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. — Berlin, Heidelberg : Springer-Verlag, 2021. — Vol. 31. — № 1. — P. 132–143.

11. Golovko, V. Deep convolutional neural network for detection of solar panels / V. Golovko, A. Kroshchanka, E. Mikhno, M. Komar, A. Sachenko // Data-Centric Business and Applications. — Cham: Springer International Publishing, 2021. — P. 371–389.

12. 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 (OSTIS-2021). — Minsk : BSUIR, 2021. — P. 219–224.

13. 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.

14. Kroshchanka, A. Reduction of neural network models in intelligent computer systems of a new generation / A. Kroshchanka // Open Semantic Technologies for Intelligent Systems (OSTIS-2023). — Minsk : BSUIR, 2023. — P. 127–132.

Статьи в других научных изданиях

15. Golovko, V. A. 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 // Open Semantic Technologies for Intelligent Systems (OSTIS-2018). — Minsk : BSUIR, 2018. — P. 133–146.

16. Golovko, V. 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). — Minsk : BSUIR, 2019. — P. 91–102.

17. Golovko, V. 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). — Minsk : BSUIR, 2020. — P. 175–182.

Статьи в сборниках материалов научных конференций

18. Golovko, V. A New Technique for Restricted Boltzmann Machine Learning / A. Kroshchanka, V. Turchenko, S. Jankowski, D. Treadwell //

Proceedings of the 8th IEEE International Conference IDAACS-2015, Warsaw 24-26 September 2015. — Warsaw, 2015. — P. 182–186.

19. 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 (OSTIS-2015) : материалы V междунар. науч.-техн. кофн. — Минск : БГУИР, 2015. — P. 481–486.

20. Крощенко, А. А. Применение нейронных сетей глубокого доверия в интеллектуальном анализе данных / А. А. Крощенко // сборник материалов IX Республиканской научной конференции молодых ученых и студентов «Современные проблемы математики и вычислительной техники». — Брест : БрГТУ, 2015. — С. 12–14.

21. Golovko, V. Theoretical Notes on Unsupervised Learning in Deep Neural Networks / V. Golovko, A. Kroshchanka // Proceedings of the 8th International Joint Conference on Computational Intelligence (IJCCI 2016). — SCITEPRESS, 2016. — P. 91–96.

22. Golovko, V. Convolutional neural network based solar photovoltaic panel detection in satellite photos / V. Golovko, S. Bezobrazov, A. Kroshchanka, A. Sachenko, M. Komar, A. Karachka // 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). — IEEE, 2017. — Vol. 1. — P. 14–19.

23. Golovko, V. Development of solar panels detector / V. Golovko, A. Kroshchanka, S. Bezobrazov, A. Sachenko, M. Komar, O. Novosad // International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T). — IEEE, 2018. — P. 761–764.

24. 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 : сборник материалов VII международной научно-практической конференции. — Брест : БрГУ, 2018. — С. 3–5.

25. Головки, В. А. Обнаружение и распознавание маркировки продукции с помощью нейросетевых алгоритмов / В. А. Головки, А. А. Крощенко // Вычислительные методы, модели и образовательные технологии : сборник материалов VIII международной научно-практической конференции. — Брест : БрГУ, 2019. — С. 3–6.

26. Kroshchanka, A. The Reduction of Fully Connected Neural Network Parameters Using the Pre-training Technique / A. Kroshchanka, V. Golovko // 11th IEEE International Conference on Intelligent Data Acquisition and

Advanced Computing Systems: Technology and Applications (IDAACS). — IEEE, 2021. — Vol. 2. — P. 937–941.

Тезисы

27. Крощенко, А. А. Методы глубокого обучения нейронных сетей / А.А. Крощенко // Материалы конференции «Вычислительные методы, модели и образовательные технологии». — Брест : БрГУ, 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 : сборник материалов международной научно-практической конференции. — Брест : БрГУ, 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 : сборник материалов международной научно-практической конференции. — Брест : БрГУ, 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 : сборник материалов международной научно-практической конференции. — Брест : БрГУ, 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. Разработан метод неконтролируемого предобучения глубоких нейронных сетей, позволяющий учитывать нелинейную природу нейронных элементов. Разработан алгоритм, позволяющий сократить количество настраиваемых параметров сети. Для демонстрации эффективности предложенных метода и алгоритма проведены вычислительные эксперименты с использованием выборок компьютерного зрения MNIST, CIFAR-10 и CIFAR-100.

Доказана теорема, что максимизация функции правдоподобия распределения входных данных в пространстве синаптических связей ограниченной машины Больцмана эквивалентна минимизации кросс-энтропии функции ошибки сети и минимизации суммарной квадратичной ошибки сети при использовании линейных нейронов.

Полученные теоретические результаты использовались при разработке практических решений в области компьютерного зрения.

Предложена нейросетевая система обнаружения солнечных панелей на аэрофотоснимках, базирующаяся на использовании сверточных нейронных сетей.

Для распознавания маркировки продукта на производственной линии предложена нейросетевая система, основывающаяся на применении конвейера нейросетевых моделей для решения отдельных подзадач обработки, который позволил осуществлять анализ маркировки в реальном времени.

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

РЭЗЮМЭ

Крашчанка 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. Распрацаваны метады некантралюемага праднавучання глыбокіх нейронавых сетак, які дазваляе ўлічваць нелінейную прыроду нейронавых элементаў. Распрацаваны алгарытм, які дазваляе скараціць колькасць наладжвальных параметраў сеткі. Для дэманстрацыі эфектыўнасці прапанаваных метадаў і алгарытму праведзены вылічальныя эксперыменты з выкарыстаннем агульнавядомых выбарак камп'ютэрнага зроку MNIST, CIFAR-10 і CIFAR-100.

Даказаная тэарэма, што максімізацыя функцыі праўдападабенства размеркавання ўваходных дадзеных у прасторы сінэптычных сувязяў абмежаванай машыны Больцмана эквівалентная мінімізацыі крос-энтрапіі функцыі памылкі сеткі і мінімізацыі сумарнай квадратычнай памылкі сеткі пры выкарыстанні лінейных нейронаў.

Атрыманыя тэарэтычныя вынікі выкарыстоўваліся пры распрацоўцы практычных рашэнняў у вобласці кампутэрнага зроку.

Прапанавана нейросеткавая сістэма выяўлення сонечных панэляў на аэрафотаздымках, які базуецца на выкарыстанні згорткавай нейронавай сеткі.

Для распазнання маркіроўкі прадукта на вытворчай лініі прапанавана нейросеткавая сістэма, якая засноўваецца на ўжыванні канвеера нейрасеткавых мадэляў для рашэння асобных падзадач апрацоўкі, які дазваляе ажыццяўляць аналіз маркіроўкі ў рэальным часе.

Рэкамендацыі па выкарыстанні і вобласць ужывання: інтэлектуальныя сістэмы відэааналіза, 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. A method for unsupervised pretraining of deep neural networks has been developed, which allows taking into account the nonlinear nature of neural elements. An algorithm has been developed to reduce the number of tunable neural network parameters. To demonstrate the effectiveness of the proposed method and algorithm, computational experiments were carried out using computer vision datasets MNIST, CIFAR-10 and CIFAR-100.

The theorem is proven that maximizing the likelihood function of the distribution of input data in the space of synaptic connections of a limited Boltzmann machine is equivalent to minimizing the cross-entropy of the network error function and minimizing the total squared error of the network when using linear neurons.

The theoretical results obtained were used to develop practical solutions in the area of computer vision.

A neural network system for detecting solar panels on aerial photographs is proposed, based on the use of convolutional neural networks.

To recognize product markings on a production line, a neural network system is proposed, based on the use of a pipeline of neural network models to solve separate processing subtasks, which made it possible to analyze markings in real time.

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.

Научное издание

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