

**Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Сибирский государственный индустриальный университет»
Администрация Правительства Кузбасса
Администрация г. Новокузнецка
Институт проблем управления им. Трапезникова РАН
Федеральный исследовательский центр угля и углехимии СО РАН
АНО «Научно-образовательный центр «Кузбасс»**

**СИСТЕМЫ АВТОМАТИЗАЦИИ
(в образовании, науке и производстве)
AS' 2022**

**ТРУДЫ ВСЕРОССИЙСКОЙ НАУЧНО–ПРАКТИЧЕСКОЙ
КОНФЕРЕНЦИИ
(с международным участием)**

15-16 декабря 2022 г.

**Новокузнецк
2022**

УДК 658.011.56
С 409

Редакционная коллегия:

д.т.н., проф. В.В. Зимин (ответственный редактор),
д.т.н., проф. С.М. Кулаков, д.т.н., проф. В.Ю. Островляничик,
д.т.н., проф. Л.Д. Павлова, д.т.н., доц. И.А. Рыбенко,
к.т.н., доц. В.И. Кожемяченко (технический редактор).

С 409 Системы автоматизации (в образовании, науке и производстве) AS'2022: труды Всероссийской научно–практической конференции (с международным участием), 15-16 декабря 2022 г. / Министерство науки и высшего образования Российской Федерации, Сибирский государственный индустриальный университет ; под общ. ред. В.В. Зимина. – Новокузнецк : Издательский центр СибГИУ, 2022. – 632 с.

ISBN 978-5-7806-0583-6

Труды конференции посвящены научным и практическим вопросам в области современных систем автоматизации и информатизации учебных, исследовательских и производственных процессов. Представлены результаты исследования, разработки и внедрения методического, математического, программного, технического и организационного обеспечения систем автоматизации и информационно-управляющих систем в различных сферах деятельности.

Сборник трудов ориентирован на широкий круг исследователей, научных работников, инженерно-технический персонал предприятий и научно-исследовательских лабораторий, преподавателей вузов, аспирантов и обучающихся по программам бакалавриата и магистратуры.

УДК 658.011.56

© Сибирский государственный
индустриальный университет, 2022

7. Домнышев А.В. Современные подходы в подготовке кадров для промышленного производства на основе автоматизации и цифровизации процесса обучения / А.В. Домнышев, научный руководитель О.А. Затепакин// // Лучшая научная статья 2021 : сборник статей XIV Всероссийская научно-практическая конференция «высокие технологии, наука и образование: актуальные вопросы, достижения и инновации», 23 января 2022 г. – Пенза. – URL: <http://library.sibsiu.ru>.
8. Сибел, Т Цифровая трансформация. Как выжить и преуспеть в новую эпоху / Томас Сибел. – Москва : Манн, Иванов и Фербер, 2021. – 256 с
9. Капустянова К.А., Гелета И.В. Совершенствование системы оценки конкурентоспособности персонала на предприятии//Символ науки.–2016. – №11-1. – С. 67–69

УДК 004.93

ИСПОЛЬЗОВАНИЕ ДЕТЕКЦИИ В СВЁРТОЧНЫХ НЕЙРОННЫХ СЕТЯХ ДЛЯ ПОВЫШЕНИЯ ТОЧНОСТИ КЛАССИФИКАЦИИ

Бычков А.Г., Киселёва Т.В., Маслова Е.В.

ФГБОУ ВО «Сибирский государственный индустриальный университет»

г. Новокузнецк, Россия, aleksds1@yandex.ru; kis@siu.sibsiu.ru;

Elenamaslova1805@yandex.ru

Аннотация. В работе рассматриваются структура свёрточной нейронной сети и математические методы, используемые для подсчёта её значений. Для повышения точности классификации предложены методы детекции. Также рассматриваются различные способы оптимизации исходного набора данных, что тоже влияет на скорость работы алгоритмов.

Ключевые слова: нейронные сети, свёрточные нейронные сети, детекция, подготовка данных.

Abstract. The paper discusses the structure of a convolutional neural network and the mathematical methods used to calculate its values. To improve the classification accuracy, detection methods are proposed. Various ways of optimizing the initial data set are also considered, which also affects the speed of the algorithms.

Keywords: neural networks, convolutional neural networks, detection, data preparation.

В настоящее время машинное обучение применяется во многих областях, как науки, так и промышленности. Машинное обучение (англ. machine learning, ML) – класс методов искусственного интеллекта, характерной чертой которых является не прямое решение задачи, а обучение в процессе применения решений множества сходных задач. Для построения таких методов используются средства математической статистики, численных методов, методов оптимизации, теории вероятностей, теории графов, различные технологии работы с данными в цифровой форме [1]. Данное определение подразумевает наличие определённых математических действий, которые могут потребовать недопустимо большое количество машинного времени. Это обстоятельство является одной из причин недостаточно быстрого внедрения технологий машинного обучения в различные сферы деятельности. Довольно часто нейронные сети и машинное обучение применяется для распознавания изображений [2]. В докладе рассмотрены возможные способы использования методов сегментации в машинном обучении в свёрточных нейронных сетях.

Суть свёрточных нейронных сетей состоит в том, что веса и значения активации какого-либо нейрона зависят не от всех входов, а только от некоторой окрестности «вокруг» него. После чего такое «движущееся окно» проходит по всему изображению.

В итоге для соседнего «пикселя» в свёрточном слое будут те же самые веса, но входной кусок картинки будет сдвинут.

Идея состоит в том, что в одной ячейке свёрточного слоя записывается определённый паттерн изображения (линии, углы, цветовые пятна и т.п.), которые затем прогоняются с целью нахождения похожих паттернов в других частях изображения, либо других изображениях в принципе.

Важной особенностью свёрточных нейронных сетей является Pooling layer или, как в некоторых источниках его называют, субдискретизирующий слой [3].

Он даёт возможность менять длину и высоту активационных карт. Поставив этот слой после ReLU, можно сжать получившиеся значения. Делается это по простому правилу. К примеру, используется Max Pool – выбирается максимальное значение из квадрата 2x2 и передаётся дальше. Это позволяет во время обратного прохода отдать весь градиент именно этому значению, а остальным – нули (максимум зависит от одного входа).

Такое сжатие позволяет создать информацию, которой оперирует сеть, более глобальной. В вышеуказанном примере в левом верхнем квадранте будет самым важным число 6, а остальные входы имеют меньшее значение, что позволит применять обученную в дальнейшем сеть к более широкому классу изображений, избегая тем самым переобучения модели. Это даёт возможность передавать дальше лишь самые выделяющиеся признаки на картинке, не вдаваясь в мелкие подробности (избегая переобучения), но при этом, обучая модель, видеть всю картинку целиком [3].

На вход подаётся изображение, дальше свёрточный слой его обрабатывает, затем изображение пропускается через нелинейность (ReLU). Может быть несколько комбинаций свёрточных и ReLU-слоёв. Далее происходит сжатие с помощью Pooling Level. Вся эта конструкция может повторяться далее, пока не получится свёрточный слой, достаточно малый для того, чтобы его можно было сделать «плоским» – перевести в одномерный массив, – связать с полносвязным слоем, который можно пропустить через Softmax и получить уже готовый результат.

Данная архитектура имеет огромное преимущество: она по умолчанию адаптирована к сдвигам изображений. Если нужный элемент переместился, к примеру, в другой угол изображения, то свёрточные веса тоже сдвинутся. Но в отличие от перцептрона есть ещё и Pooling-слои, которые обеспечат максимальное значение активации. Через несколько итераций таких слоёв в итоге на выходе окажется именно тот, который отвечает за распознавание именно этого элемента [3].

Подавляющее число ранних исследований связано с классификацией. Сначала бинарной, а потом и многовариантной, но всё же классификацией одного объекта. Тогда как на изображении может быть не только собака, но и велосипед, автомобиль и т.п. Кроме того, у той же собаки есть лапы, хвост и т.д. Возникла потребность извлечения автоматически больше информации. И потому возникли задачи сегментации и детекции.

Основное отличие между сегментацией и детекцией состоит в том, что детекция применяется в первую очередь для задач нахождения координат объекта как на рисунке 1.

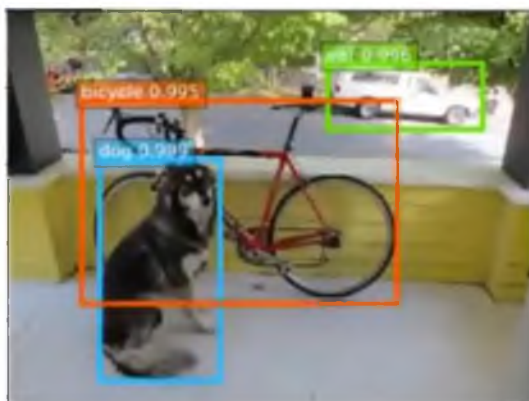


Рисунок 1 – Пример использования детекции

Детекцию можно использовать и для нахождения типа объекта, его атрибутов. Во многом из-за этого детекция порой используется чаще, чем сегментация. У детекции есть сложности с определением конкретной метрики и соответственно loss-функции. Для де-

текции важно в первую очередь точность определение координат (что прямоугольник, обводящий объект, расположен правильно) и confidence объекта (что корректно определён класс объекта). Обычно используют для метрики mAP – mean Average Precision.

Одним из основных подходов к детекции является Non Maximum Supression. Его суть состоит в том, чтобы сначала предсказать много боксов (тех самых прямоугольников), а потом отфильтровать лишние (рисунок 2). На текущий момент именно этот подход используется в подавляющем большинстве детекционных сетей [4].



Рисунок 2 – Фильтрация с помощью Non-Maximum Supression

Детекционные сети бывают двух видов: One-shot (быстрые) – YOLO, SSD, RetinaNet, SqueezeNet, DetectNet, и Two-shot (точные) – R-FCN, Fast RCNN, Faster-RCNN.

Различие между быстрыми и точными сетями важно по причине различных вариантов использования этих сетей. Скажем, если есть камера видеонаблюдения на улице, считывающая номера мимо проезжающих машин, то скорость куда важнее точности в данном случае. Но при решении задач, которые можно делать оффлайн, не на лету, таких как обработка спутниковых снимков, можно позволить себе использовать точные, но более медленные сети.

Идея One-shot сетей проста в реализации. Берётся исходное изображение, делится на регионы (ячейки), после чего для каждой такой ячейки в последнем Conv-слое сети предсказываются координаты бокса и класс объекта с центром в этой самой ячейке (рисунок 3). В обычной сегментационной сети на этом месте было бы предсказание класса каждого пикселя. Условно говоря, для бинарной классификации для каждой ячейки предсказываются 5 чисел: первые 4 – это координаты бокса, а последнее – вероятность класса. В случае мультиклассовой классификации будет не одно число в конце, а столько, сколько классов.

Данная сеть работает очень быстро, так как нужно предсказывать значительно меньшее количество раз. Скажем, вместо попиксельной классификации, когда может возникнуть ситуация с 1000x1000 вычислениями, в этом же случае будет 7x7 ячеек – 49 вычислений.

Одним из основных one-shot детекторов является YOLO (You Only Look Once), который представляет принципы Non Maximum Supression и деления на ячейки [5].

Для каждой ячейки в последнем conv-слое предсказываются координаты бокса и класс объекта с центром в ячейке. Для корректной работы данного алгоритма нужно иметь хорошо обученный энкодер, но зато при работе модели на практике получается великолепная производительность и приемлемая точность.

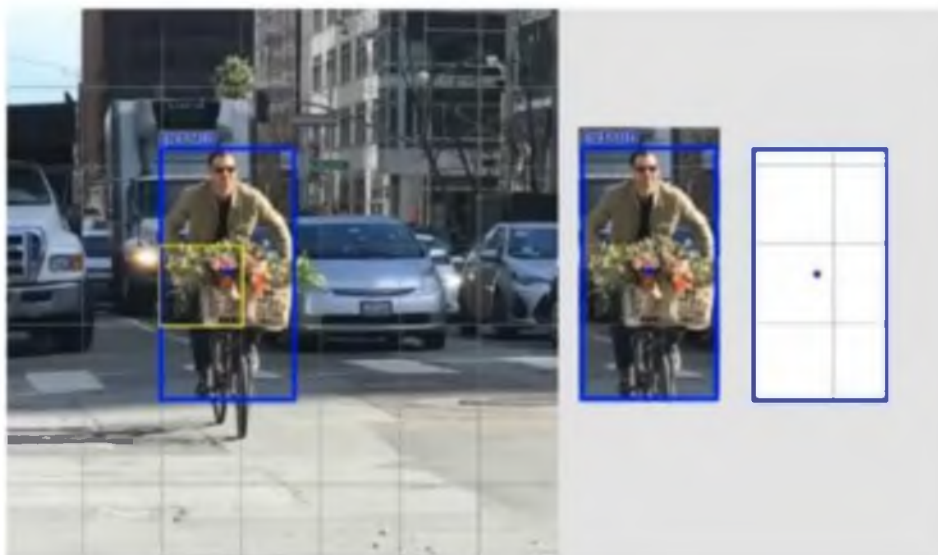


Рисунок 3 – Пример работы One-Shot сети

Другой популярный детектор, но уже с FPN, называется SSD (Single Shot Detector).

SSD предназначается для решения проблемы multiscale, когда на вход могут подаваться изображения разных размеров.

Далее идут Two-shot детекторы. Одним из основных является R-CNN (Region-based Convolutional Neural Network). Принцип работы показан на рисунке 4.

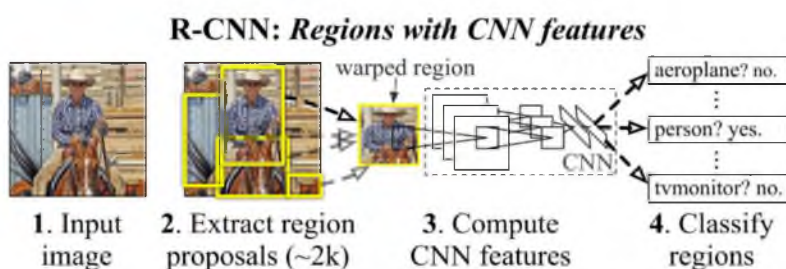


Рисунок 4 – Схема R-CNN

Условный алгоритм большинства Two-shot детекторов можно описать следующим образом. Берётся некоторое изображение, на нём некоторым способом (обычно *selective choice*) выбираются регионы, где возможно могут находиться объекты, которые нужно классифицировать.

Для каждого такого региона с помощью свёрточных нейросетей производится классификация.

В итоге получаем отдельные регионы с объектами и их классами. В центре стоят свёрточные нейронные сети, которые показывают хорошую точность на примере изображений. Но у такой архитектуры есть недостатки: он энергозатратный – требует большого количество времени на обучение. В методе *selective search* изображение сначала сегментируется на 2000 регионов, которые затем в ходе итерирования с помощью жадного алгоритма объединяются в более крупные регионы. Кроме того, сами свёрточные сети тоже требуют вычислительных мощностей. R-CNN не может быть использован для видео. Это вытекает из недостатка выше, так как все промежуточные методы энергозатратны, поэтому кадры просто не будут успевать обрабатываться. *Selective search* не является алгоритмом машинного обучения, поэтому могут возникнуть проблемы с выявлением потенциальных объектов на разных изображениях. На данный момент модель R-CNN устарела и не применяется.

Недостатки R-CNN привели авторов в 2015 году к улучшению модели [6]. Они назвали ее Fast R-CNN. В ее основе лежит следующая архитектура: изображение подается на вход сверточной нейронной сети и обрабатывается selective search. В итоге, имеем карту признаков и регионы потенциальных объектов. Координаты регионов потенциальных объектов преобразуются в координаты на карте признаков. Полученная карта признаков с регионами передается слою RoI (Region of Interest) pulling layer. Здесь на каждый регион накладывается сетка размером $H \times W$. Затем применяется MaxPulling для уменьшения размерности. Так, все регионы потенциальных объектов имеют одинаковую фиксированную размерность. Полученные признаки подаются на вход полносвязного слоя (Fully-connected layer), который передается двум другим полносвязным слоям. Первый с функцией активации softmax определяет вероятность принадлежности классу, второй – границы (смещение) региона потенциального объекта. Схема работы алгоритма приведена на рисунке 5.

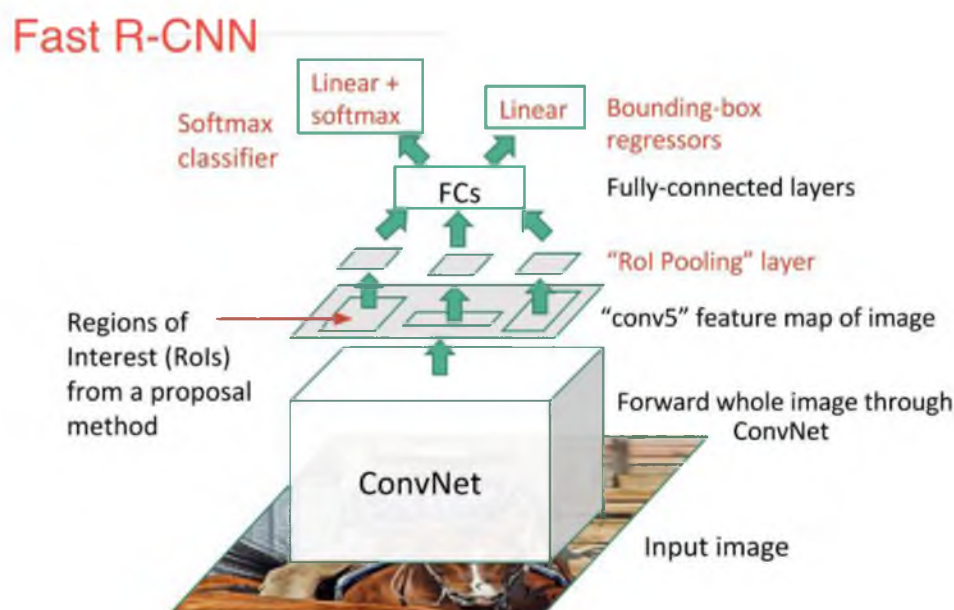


Рисунок 5 – Схема Fast R-CNN

Fast R-CNN показывает чуть более высокую точность и большие затраты времени обработки в отличие от R-CNN, так как не требуется подавать все регионы на сверточный слой. Но тем не менее, данный метод использует затратный Selective Search. Поэтому авторы пришли к Faster R-CNN.

Авторы продолжили улучшение Fast R-CNN и в 2016 предложили Faster R-CNN. Они разработали собственный метод локализации объекта взамен Selective Search – RPN (Region Proposal Networks) [7]. В основе RPN лежит система якорей. Архитектура Faster R-CNN образована следующим образом: изображение подается на вход сверточной нейронной сети. Так формируется карта признаков. Карта признаков обрабатывается слоем RPN. Здесь скользящее окно проходит по карте признаков. Центр скользящего окна связан с центром якорей. Якоря – это области, имеющие разные соотношения сторон и разные размеры. Авторы используют 3 соотношения сторон и 3 размера. На основе метрики IoU(intersection-over-union), степени пересечения якорей и истинных размеченных прямоугольников, выносится решение о текущем регионе – есть объект или нет. Далее используется алгоритм FastCNN: карта признаков с полученными объектами передаются слою RoI с последующей обработкой полносвязных слоев и классификацией, а также с определением смещения регионов потенциальных объектов. Схема Faster R-CNN показана на рисунке 6.

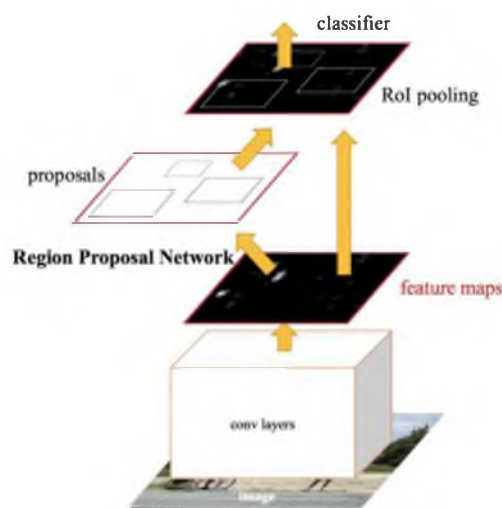


Рисунок 6 – Схема Faster R-CNN

Модель Faster R-CNN справляется немного хуже с локализацией, но работает быстрее Fast R-CNN. По различным исследованиям получалось, что R-CNN обрабатывает изображение около 50 секунд, Fast R-CNN – 2 секунды, а Faster R-CNN – около 0.2 секунды. Для улучшения работоспособности Faster R-CNN можно использовать следующие методы: добавить «тяжёлые» backbones (предобученные энкодеры из больших и громоздких сетей), добавить FPN, заменить RoI Pooling на RoI-align, производить аугментации изображений.

Таким образом, в данной работы рассмотрено текущее состояние области задач искусственного интеллекта. Особенность задач искусственного интеллекта состоит в том, что как только они будут решены, они переходят в разряд задач обычного вычисления.

Приведено определение свёрточных нейронных сетей. Описана их структура и основные отличия от обычного перцептрона: наличие свёрточных слоёв и pooling-слоёв.

Было дано определение принципов детекции в машинном обучении, а конкретно в распознавании образов.

Показана история развития свёрточных нейросетей на основе детекции, их эволюция. Описаны методы, которые используются для повышения эффективности работы сетей.

Библиографический список

1. Комашинский, В.И. Нейронные сети и их применение в системах управления и связи / В.И. Комашинский. – М.: Горячая линия-Телеком, 2002 – 94 с.
2. Бычков, А.Г. Методика оптимизации элементов нейронной сети на примере перцептрона / А.Г. Бычков, Т.В. Киселева, Е.В. Маслова // Системы управления и информационные технологии. – 2022. – № 1. – С. 4-8.
3. Червяко, Н.И. Применение нейронных сетей для задач прогнозирования и проблемы идентификации моделей прогнозирования / Н.И. Червяков // Нейрокомпьютеры: разработка и применение, 2003. – № 10. – С. 11-14.
4. Круглов, В.В. Искусственные нейронные сети. Теория и практика / В.В. Круглов. – М.: Горячая линия-Телеком, 2001. – 382 с.
5. Хайкин, С. Нейронные сети. Полный курс : пер. с англ. – 2-е изд. - М. : ИД «Вильямс», 2006. - 1104 с.
6. Николенко, С. Глубокое обучение / С. Николенко, А. Кадурын, Е. Архангельская. – СПб. : Питер, 2018. - 480 с.
7. Яковлев, Е. Л. Комплекс программ для навигации мобильных устройств внутри помещений с помощью нейронных сетей / Е. Л. Яковлев, А. Д. Хомоненко, П. А. Новиков // Информационно-управляющие системы. - 2016. - № 1. - С. 32–39.

<i>Каган Е.С., Гоосен Е.В., Колтинская С.А., Ложкин А.А., Михальченко М.А.</i>	
Проблемы и перспективы направления разработки инструментов количественной оценки стрессоустойчивости цепочек добавленной стоимости в угольной отрасли.....	391
<i>Гостевская А.Н., Маркидонов А.В.</i>	
Изменение поверхности ОЦК-кристалла при лазерной абляции	396
<i>Павлова Л.Д., Фрянов В.Н.</i>	
Моделирование предразрушения горных пород под влиянием микросейсмических воздействий на геомассив в окрестности подземных горных выработок и угольных целиков.....	400
<i>Бабушкина О.С., Калашников С.Н.</i>	
Итерационный метод решения уравнений над телом кватернионов.....	406
<i>Ермакова Л.А.</i>	
Опыт разработки плагинов в СУО Moodle для анализа работы пользователей	411
<i>Бегинина А.Г.</i>	
Информационная система ведения расписания для образовательного центра	413
<i>Бегинина А.Г.</i>	
Ценность информационных технологий в системах управления	417
<i>Bhattacharjee S., Chakraborty P., Roy M., Banerjee A., Pal M., Nikitenko M.S., Neogi B.</i>	
Analytical solution of ‘nonlinearly coupled electromechanical model equations’ of human cardiovascular muscle.....	420
СЕКЦИЯ 3. Информационные технологии в управлении организационными системами	
<i>Добронец Б.С., Попова О.А., Шломин А.А.</i>	
Прогнозная аналитика и большие данные в оценке рисков инвестиционных проектов.....	437
<i>Добронец Б.С., Попова О.А.</i>	
Информационно-аналитические подходы в анализе неструктурированных данных	442
<i>Домнъшев А.В.2, Затеякин О.А.</i>	
Опережающее развитие персонала как фактор повышения конкурентоспособности персонала и развития бизнеса	447
<i>Бычков А.Г., Киселёва Т.В., Маслова Е.В.</i>	
Использование детекции в свёрточных нейронных сетях для повышения точности классификации.....	453
<i>Блюмин С.Л.</i>	
Метаграфы и редукция Крона в моделировании оргсистем.....	459
<i>Щепкин А.В., Амелина К.Е.</i>	
Стимулирование публикационной активности.....	464
<i>Графкин А.В., Александрова М.И.</i>	
Разработка системы реализации алгоритма anti-tailgate для предотвращения несанкционированного прохода.....	469

Научное издание

**СИСТЕМЫ АВТОМАТИЗАЦИИ
(в образовании, науке и производстве)
AS' 2022**

**ТРУДЫ ВСЕРОССИЙСКОЙ НАУЧНО–ПРАКТИЧЕСКОЙ
КОНФЕРЕНЦИИ**
(с международным участием)

15-16 декабря 2022 г.

Под общей редакцией д.т.н., доц. В.В. Зимина

Техническое редактирование и компьютерная верстка В.И. Кожемяченко

Подписано в печать 05.12.2022 г.

Формат бумаги 60×84 1/16. Бумага писчая. Печать цифровая.
Усл. печ. л. 37.13. Уч.-изд. л. 40.40. Тираж ____ экз. Заказ ____.

Сибирский государственный индустриальный университет
654007, Кемеровская область – Кузбасс, г. Новокузнецк, ул. Кирова, зд. 42.
Издательский центр СибГИУ