

DOI 10.37882/2223-2966.2025.04.07

СИСТЕМА ПОВЫШЕНИЯ КАЧЕСТВА РАБОТЫ ОПТИКО-ЭЛЕКТРОННЫХ КОМПЛЕКСОВ УМНОГО ДОМА

SYSTEM FOR IMPROVING THE QUALITY OF OPERATION OF OPTICAL-ELECTRONIC COMPLEXES OF A SMART HOME

**P. Gessen
S. Saradgishvili**

Summary. This work is devoted to the problem of the quality of object identification and tracking in smart home ecosystems with the possibility of video processing. Due to the limitations in this area, the object of interest is approaches that remain computationally efficient with weak provision of computing resources. The authors conducted a study of both proven approaches and the latest methods in the field of tracking and identification of objects. Previously, the authors proposed a system that complements existing high-performance implementations to improve the accuracy characteristics of tracking and detection by adding scale and partial rotational invariance, a classical approach to extracting key features due to a histogram of directed segments, the support vector method as a classifier and further training it in the process. In the current work, an analysis of possible approaches to improving the previously presented method is carried out, existing solutions to the problem of collecting and determining key features of an object are studied, the implementation of a system with a neural network approach to their identification is provided, as well as the results of the work of original algorithms and using both proposed solutions, which show an improvement in the accuracy indicators of both tasks of interest. These improvements can be used in any smart home systems, although the possible scope of application may be much wider, to improve the quality of video surveillance, which may have a better impact on the quality of life of consumers of such products: from becoming a smart home even «smarter» in terms of predictive ability to improving security by improving surveillance systems.

Keywords: training, support vector machine, neural networks, clustering, tracking.

Гессен Павел Алексеевич

аспирант, Федеральное государственное автономное
образовательное учреждение высшего образования
«Санкт-Петербургский
политехнический университет Петра Великого»
gessen.pa@edu.spbstu.ru

Сараджишвили Сергей Эрикович

Кандидат технических наук, доцент, Федеральное
государственное автономное образовательное
учреждение высшего образования
«Санкт-Петербургский политехнический
университет Петра Великого»
saradzh_se@spbstu.ru

Аннотация. Данная работа посвящена проблеме качества выявления объектов и слежения за ними в экосистемах умного дома с возможностью видеообработки. Ввиду ограничений в данной области, объектом интереса являются подходы, которые остаются вычислительно эффективными при слабом обеспечении вычислительными ресурсами. Авторами проведено исследование как зарекомендовавших себя подходов, так и новейших методов в сфере слежения и выявления объектов. Ранее авторами была предложена система, которая дополняет существующие высокопроизводительные реализации для повышения точностных характеристик сопровождения и обнаружения за счет добавления масштабной и частичной поворотной инвариантности, классического подхода к извлечению ключевых особенностей за счет гистограммы направленных градиентов, метода опорных векторов в качестве классификатора и дообучения его в процессе работы. В текущей работе проводится анализ возможных подходов к улучшению ранее представленного метода, осуществляется изучение существующих решений задачи сбора и определения ключевых особенностей объекта, приводится реализация системы с нейросетевым подходом к их выявлению, а также приводятся результаты работы оригинальных алгоритмов и с использованием обоих предложенных решений, на которых видно улучшение точностных показателей обеих задач интереса. Данные улучшения представляется возможным использовать в любых системах умного дома, хотя возможная область применения может быть сильно шире, для повышения качества видеонаблюдения, что может в лучшую сторону повлиять на качество жизни потребителей подобных продуктов: от становления умного дома еще «умнее» в плане предсказательной способности до повышения безопасности за счет улучшения систем наружного наблюдения. Дополнительным преимуществом является простота внедрения в виду структуры предлагаемой системы: работа с любыми алгоритмами и модульность.

Ключевые слова: дообучение, метод опорных векторов, нейронные сети, кластеризация, слежение.

Введение

В настоящее время происходит активное развитие информационных технологий, что влияет на очень многие сферы жизни человека. Одним из самых значимых изменений является прогресс в оптических системах, что подтверждается их повсеместным использованием в совершенно различных областях: промышленность [1, с. 107], медицина [2, с. 3], транспорт [3] и прочее.

Одной из таких областей, которая может быть близка каждому, а не только узкопрофильным специалистам, является повседневная жизнь, а именно появление в ней экосистемы умного дома. Это направление активно развивается в последние годы. Сама система умного дома состоит из множества различных устройств с разным функционалом и способов их взаимодействия, включая и устройства с видеообработкой. Уже на данный момент используются различные приборы, функционирующие на основе данных с видеовхода, и еще больше находятся на этапе разработки. Помимо улучшения аппаратного обеспечения, разработка ведется и в направлении программного обеспечения. Однако общемировая практика таких исследований направлена в сторону улучшения точности решения задач в угоду производительности за счет использования большего объема вычислительных ресурсов [4; 5, с. 357], а это ведет к удорожанию производства, как следствие экономической нецелесообразности, и возможной несовместимости с определенными платформами, которые ограничены в своей производительности. При этом существующие программные решения имеют задел по производительности на текущих вычислительных мощностях.

Исходя из вышеописанного, актуальной задачей становится общее улучшение точностных характеристик существующих алгоритмов без сильного замедления производительности на текущих вычислителях. В статье [6, с. 160] предлагается решение поставленной проблемы. В данной статье авторами предлагается улучшить данную разработку путем еще большего улучшения показателей точности.

Таким образом цель настоящей работы — повысить показатель увеличения точностных характеристик исходной системы. Для этого должны быть выполнены следующие задачи: исследование существующих подходов и их сравнение, проектирование и реализация изменений, получение и анализ результатов.

Исследование предметной области

Рассмотрение подходов к решению задач выявления и слежения за объектами в видеопотоке приведено в статье [6, с. 158] и не является областью интереса

конкретной статьи. В данной работе предлагается улучшить изначальную систему. Для этого необходимо проанализировать исходные данные и определить области с наибольшим интересом. Обращая внимание на структуру исходного модуля, такими местами, которые потенциально могут принести наибольший качественный прирост, являются подход к классификации и подход к выделению ключевых особенностей объекта. В данной работе авторами предлагается изменить второй аспект ввиду его более раннего влияния на результат и меньшей комплексной интегрированности во всю систему.

В изначальной статье для описания объекта используется гистограмма направленных градиентов [7] — классический метод представления объекта, выбранный из-за высокой скорости работы и доказанной эффективности. Однако текущее состояние области таково, что для решения подобных задач в подавляющем большинстве случаев используются нейросетевые подходы. Рассмотрим их с точки зрения скорости работы, которая является одним из ключевых ограничений. Она напрямую зависит от количества параметров этих сетей.

Таблица 1.

Методы извлечения ключевых особенностей

Название	Количество параметров
AlexNet	60 M
VGG-16	138 M
VGG-19	144 M
Inception-V1 (GoogLeNet)	5 M
ResNet-18	11.7 M
ResNet-34	25.6 M
ResNet-50	26 M
ResNet-101	44.6 M
ResNet-152	230 M
Inception-v2	21.8 M
Inception-v3	21.8 M
Inception-ResNet-V2	55 M
Darknet-19	20.8 M
Xception	22.9 M
SqueezeNet	1.24 M
ShuffleNet	143 M
ShuffleNet-v2	2.3 M
DenseNet-100 (k = 12)	7.0 M
DenseNet-100 (k = 24)	27.2 M

Название	Количество параметров
DenseNet-250 (k = 24)	15.3 M
DenseNet-190 (k = 40)	25.6 M
DetNet	-
EfficientNet B0 до B7	от 5.3 M до 66 M
MobileNet	4.2 M
MobileNet-v2	3.4 M
WideResNet-40-4	8.9 M
WideResNet-16-8	11 M
WideResNet-28-10	36.5 M
SWideRNet	от 7.77 M до 946.69 M
HRNet W32, W48	28.5 M, 63.6 M
HRNet V2	-

Из представленного списка видно, что существует множество подходов к извлечению признаков из изображения, что доказывает их эффективность, хотя они все и основаны на сверточных нейронных сетях. Основным наблюдением в данной таблице является то, что такие сети в большинстве своем имеют большое количество параметров, что делает их «тяжелыми», то есть требовательными к аппаратной части для работы в режиме реального времени. Таким образом, наиболее подходящей реализацией является нейросеть SqueezeNet [7]. Эта сеть осуществляет классификацию изображения на заранее обученные классы, однако в рамках этой статьи она интересует авторов только способом извлечения ключевых признаков.

Предлагаемые улучшения

За основу берется предложенная в [6, с. 161] система улучшения характеристик существующих алгоритмов. Дальнейшие изменения можно разделить на 3 части в соответствии с последовательностью работы исходного подхода: обучение, корректировка, выявление.

Обучение

Данный этап является ключевым по влиянию на дальнейшее функционирование комплекса. Основным изменением является замена HOG-дескриптора [8, с. 887], который осуществлял сбор ключевых особенностей поданного на вход изображения объекта, на экстрактор ключевых особенностей нейронной сети SqueezeNet. Как было оговорено выше, он был выбран ввиду высокой вычислительной эффективности при высокой же эффективности выполнения задачи сбора признаков. Отдельно стоит отметить, что в данной работе была выбрана

предобученная на общем наборе данных реализация для более корректных результатов с точки зрения обобщенного подхода. По итогу, в начале работы осуществляется обучение классификатора на основе SVM [9, с. 25] с последующим дообучением в процессе. Также были изменены внутренние параметры классификатора для более точной работы с новым форматом входных данных.

Корректировка

Аналогично предыдущему блоку, изменению подвергся и этот: старый метод извлечения признаков заменен на новый. Теперь проверка на корректность положения объекта в кадре осуществляется на основе нейросетевого подхода. Остальные составные части, такие как кластеризация и классификация (за исключением вышеописанных правок) остались неизменными.

Выявление

Изменения аналогичные предыдущим — замена способа извлечения описания признаков объекта, теперь на вход классификатору поступает набор данных от нейронной сети, в дальнейшие этапы изменения не вносились.

Тестирование

Изменению подверглись два блока, которые необходимо рассматривать отдельно: наблюдение и выявление. Методики для их оценки приведены ниже. В обоих случаях будет использована специализированная среда оценки качества работы алгоритмов [10, с. 22], представляющая собой набор заранее размеченных видеопоследовательностей с различными объектами интереса в различных условиях.

Слежение

Для правильной и наиболее полной оценки качества, производимых предложенной системой улучшений на протяжении наблюдения и корректировки положения описывающего прямоугольника необходимо и достаточно на основании собранных показателей провести оценку метрики IoU (Intersection over Union) [11, с. 659]. В качестве улучшаемых алгоритмов были выбраны девять реализаций классических алгоритмов, которые хорошо показывают себя в условиях ограничений при слабых вычислительных мощностях. Подход к проведению тестирования для оценки результатов, применяемый на каждой из видеопоследовательностей:

- 1) Выбор основного алгоритма с добавленной оригинальной системой наблюдения.
- 2) Использование тестовой среды для оценки его характеристик.
- 3) Замена системы на улучшенную.

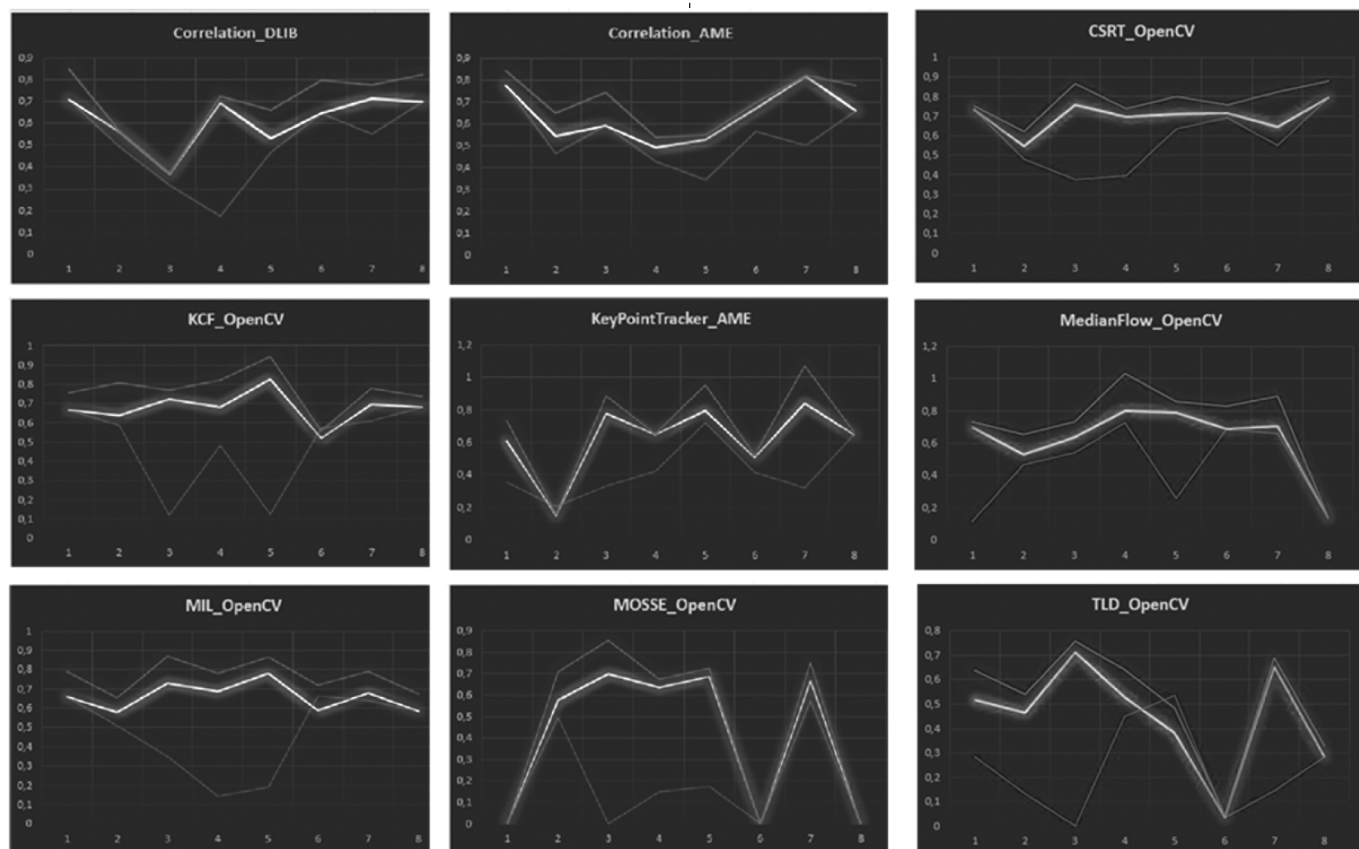


Рис. 1. Графики значений IoU слежения. Синим цветом обозначен оригинальный алгоритм, зеленым — комбинированный с оригинальной системой, голубым — улучшенная система

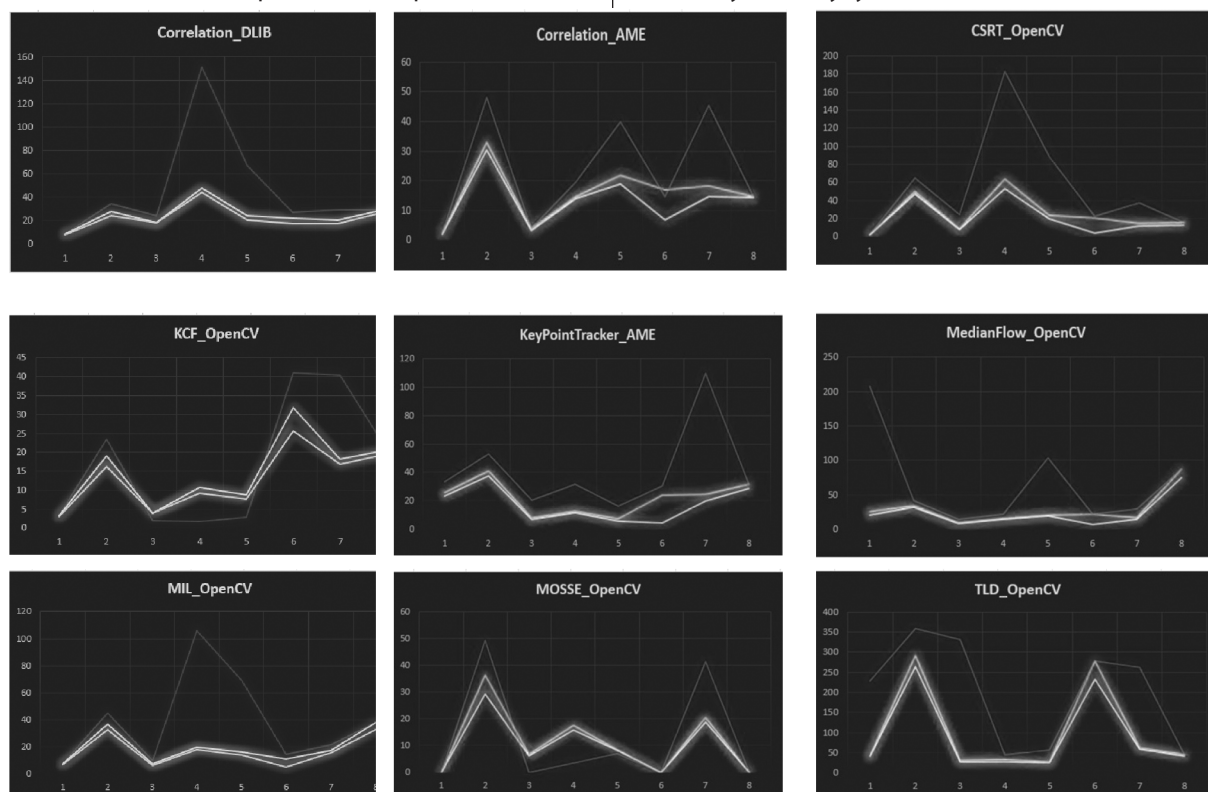


Рис. 2. Графики удаленности центров при слежении. Синим цветом обозначен оригинальный алгоритм, красным — комбинированный с оригинальной системой, зеленым — улучшенная система

- 4) Повторный запуск полученного комплекса и сбор метрик.
- 5) Анализ результатов первого и второго запусков.

Выявление

Ввиду того, что не все алгоритмы имеют возможность повторно выявлять объект интереса после его потери, метод исследования в данном случае будет несколько отличаться. Для оценки данного этапа будет использован многоагентный алгоритм в качестве эталона. Подход к проведению тестирования для оценки результатов, применяемый на каждой из видеопоследовательностей:

- 1) Обучение классификатора на протяжении 100–400 кадров при помощи размеченного истинного положения объекта.
- 2) Имитация срыва наблюдения путем перехода на другой кадр, разделенный с последним кадром обучения не менее, чем на 100 фреймов.
- 3) Получение информации о положении объекта от исследуемого способа.
- 4) Вычисление IoU и разницы дистанции центров от полученного результата и истинного положения объекта.

Данная последовательность действий применяется к трем алгоритмам: оригинальный многоагентный [12, с. 219], многоагентный с оригинальной системой и многоагентный с улучшенной системой.

Результаты

На рис. 1–4 приведены графики, полученные по результатам исследования алгоритмов по вышеописанным методикам. На всех графиках по оси абсцисс расположены порядковые номера видеопоследовательностей.

Собрав полученные результаты, можно увидеть, что для задачи слежения усредненное увеличение показателя IoU стало ~14 %, а разница между центрами в среднем стала меньше на ~16 %. Анализируя данные по задаче слежения, получаем следующие результаты: среднее значение IoU ~0.56 и средняя дистанция до центра ~21.9. Первое значение превосходит показатели оригинальной системы на 0.06, второе на 2.7. Все полученные результаты свидетельствуют об улучшении оригинальной системы в задачах корректировки и выявления, причем во втором случае можно говорить о высокой точности локализации объекта после его повторного обнаружения (для сравнения алгоритм tld на тех же данных имеет показатели 0.41 и 30.2). Однако скоростные показатели во время работы показали снижение ~32 %, что обусловлено использованием нейронной сети. Данное за-

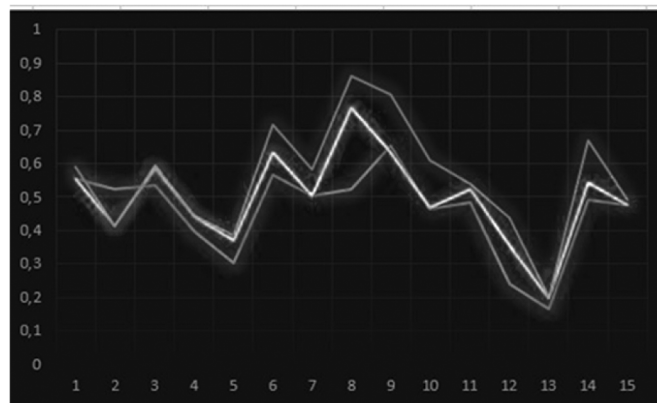


Рис. 3. Графики IoU для детекции многоагентного алгоритма (синий), оригинальной системы (оранжевый), улучшенной системы (серый)

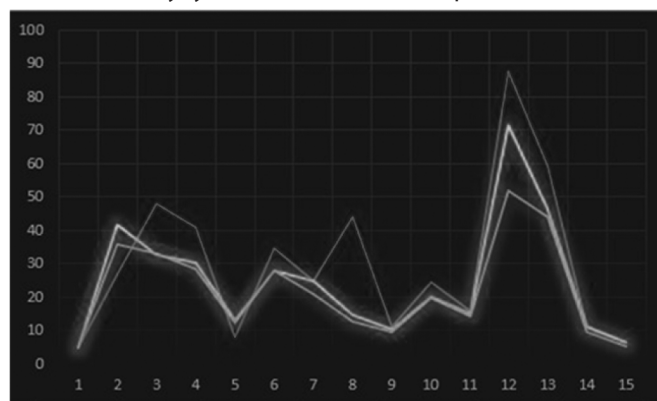


Рис. 4. Графики разницы центров для детекции многоагентного алгоритма (синий), оригинальной системы (оранжевый), улучшенной системы (серый)

медление существенно, особенно для алгоритмов, которые почти не имели запаса по производительности для работы в реальном времени, однако в некоторых случаях им можно пренебречь: в алгоритмах с большим запасом времени вычисления или если необходима повышенная точность в определенных задачах. Возможным вариантом решения проблемы замедления может быть отказ от обработки каждого кадра или оптимизация под целевую платформу.

Заключение

В работе были изучены существующие подходы извлечения ключевых особенностей из объектов на изображениях. Было предложено и реализовано усовершенствование системы улучшения точностных характеристик устройств умного дома. Проведены оценки качества работы предлагаемой системы.

ЛИТЕРАТУРА

1. Zhou L., Zhang L., Konz N. Computer Vision Techniques in Manufacturing // IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, vol. 53, no. 1, pp. 105–117, doi: 10.1109/TSMC.2022.3166397.
2. Gao J., Yang Y., Lin P., Park D. Computer Vision in Healthcare Applications // Journal of Healthcare Engineering, 2018, pp. 1–4, doi: 10.1155/2018/5157020.
3. Dilek E., Dener M. Computer Vision Applications in Intelligent Transportation Systems: A Survey // Sensor, 2023, 23(6):2938, doi: 10.3390/s23062938.
4. Rajib D., Mrinal K.B. A comprehensive survey on computer vision-based concepts, methodologies, analysis, and applications for automatic gun/knife detection. // Journal of Visual Communication and Image Representation, 2021, vol. 78, doi: 10.1016/j.jvcir.2021.103165.
5. Zhang Y., Wang L., Qi J., Wang D., Feng M., Lu H. Structured Siamese Network for Real-Time Visual Tracking // Computer Vision — ECCV 2018. ECCV 2018. Lecture Notes in Computer Science, 2018, vol. 11213, pp. 355–370, doi: 10.1007/978-3-030-01240-3_22.
6. Гессен П.А., Сараджишвили С.Э. Система улучшения точностных показателей видеобработки комплексов умного дома // Системный анализ в проектировании и управлении. 2024. №2. С. 157–166. doi: 10.18720/SPBPU/2/id24-159.
7. Iandola F., Han S., Moskewicz M., Ashraf K., Dally W., Keutzer K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size // ArXiv, 2016, URL: <https://arxiv.org/pdf/1602.07360> (дата обращения: 29.08.2024), doi: 10.48550/arXiv.1602.07360.
8. Dalal N., Triggs B. Histograms of oriented gradients for human detection // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, vol. 1, pp. 886–893, doi: 10.1109/CVPR.2005.177.
9. Shalev-Shwartz S., Singer Y., Srebro N. Pegasos: primal estimated subgradient solver for SVM // Mathematical Programming, 2011, vol. 127, pp. 3–30, doi: 10.1007/s10107-010-0420-4.
10. Бондаренко В.А., Гагарина А.Ю., Павлова В.А., Тупилов В.А. Программный комплекс автоматизации тестирования и валидации алгоритмов обнаружения и сопровождения объектов. // Известия Тульского государственного университета. Технические науки. 2021. №7. С. 16–28. doi: 10.24412/2071-6168-2021-7-16-28.
11. Rezaatofghi H., Tsoi N., Gwak J., Sadeghian A., Reid I., Savarese S. Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 658–666, doi: 10.1109/CVPR.2019.00075.
12. Бондаренко В.А., Ельцова Д.К., Лизин А.И., Павлова В.А., Созинова М.В., Тупилов В.А. Многоагентный алгоритм автоматического обнаружения и сопровождения недетерминированных объектов. // Известия ЮФУ. Технические науки. 2020. №1 (211). С. 218–232. doi: 10.18522/2311-3103-2020-1-218-232.

© Гессен Павел Алексеевич (gessen.pa@edu.spbstu.ru); Сараджишвили Сергей Эрикович (saradzh_se@spbstu.ru)

Журнал «Современная наука: актуальные проблемы теории и практики»