

*Гессен Павел Алексеевич*<sup>1</sup>,  
аспирант;  
*Сараджишвили Сергей Эрикович*<sup>2</sup>,  
доцент, канд. техн. наук, доцент

## УСОВЕРШЕНСТВОВАННАЯ СИСТЕМА УЛУЧШЕНИЯ ТОЧНОСТНЫХ ПОКАЗАТЕЛЕЙ ВИДЕООБРАБОТКИ КОМПЛЕКСОВ УМНОГО ДОМА

<sup>1,2</sup> Россия, Санкт-Петербург,  
Санкт-Петербургский политехнический университет Петра Великого,  
<sup>1</sup> gessen.pa@edu.spbstu.ru, <sup>2</sup> saradzh\_se@spbstu.ru, <sup>2</sup> ssaradg@yandex.ru

**Аннотация.** В целях повышения качественных характеристик системы улучшения точностных показателей устройств экосистемы умного дома с видеовходом, обеспечивающих выявление объектов в кадре, а также наблюдение за ними в процессе работы с дообучением, произведено исследование мирового опыта в данной области. Основной проблемой являются подходы, которые позволяют с наибольшей точностью, при сохранении высокой вычислительной эффективности, обеспечивать обучение классификаторов на лету без априорного знания о типе объекта наблюдения. Существует система, решающая данный спектр задач. Авторами исследования предложены улучшения такой системы, проведена оценка эффективности предложенных улучшений в задачах корректировки положения и выявления объектов с достаточными скоростными показателями с предварительным online-обучением в процессе первичной встречи цели.

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

*Pavel A. Gessen*<sup>1</sup>,  
Postgraduate Student;  
*Sergei E. Saradgishvili*<sup>2</sup>,  
Associate Professor, Candidate of Technical Sciences

## UPGRADED SYSTEM FOR IMPROVING THE VIDEO PROCESSING ACCURACY INDICATORS OF SMART HOME COMPLEXES

<sup>1,2</sup> Peter the Great St.Petersburg Polytechnic University, St. Petersburg, Russia,  
<sup>1</sup> gessen.pa@edu.spbstu.ru, <sup>2</sup> saradzh\_se@spbstu.ru, <sup>2</sup> ssaradg@yandex.ru

**Abstract.** To improve the quality characteristics of the system for improving the accuracy of smart home ecosystem devices with video input, which ensure the identification of objects in the frame, as well as monitoring them in the process of working with additional training, a study of world experience in this field has been conducted. The main problem is

approaches that allow for the most accurate, while maintaining high computational efficiency, to provide on-the-fly training of classifiers without a priori knowledge of the type of object of observation. There is a system that solves this problem. The authors of the study proposed improvements to such a system, evaluated the effectiveness of the proposed improvements in the tasks of correcting the position and identifying objects with sufficient speed indicators with preliminary online training during the initial meeting of the goal.

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

## **Введение**

В настоящее время практически всеобъемлющее использование различных устройств для съемки, хранения и воспроизведения визуальной информации является неотъемлемой частью жизни. Они используются в различных областях, таких как: промышленность [13], медицина [6], транспорт [5], робототехника [12], безопасность [9].

В экосистемах умного дома уже сейчас используются приборы, основывающиеся на данных с видеовхода, и еще больше находятся в процессе реализации. Однако общемировая разработка новых алгоритмов направлена на увеличение точностных характеристик за счет использования больших вычислительных ресурсов [3, 9, 14], что может привести к несовместимости с некоторым аппаратным обеспечением из-за их ограниченных возможностей по производительности, а также экономической нецелесообразности.

Таким образом, актуальной задачей становится разработка независимой системы, которая способствовала бы повышению точности наблюдения за целью и возможности ее повторного выявления. В работе [1] предлагается решение этой задачи. В данной же статье предлагается улучшить предложенную систему.

## **1. Постановка задачи**

Цель работы: увеличить точность наблюдения объекта для предложенной системы.

Задачи, необходимые для достижения цели:

- а) исследовать существующие алгоритмы и методы по описанию объекта и провести их сравнительный анализ;
- б) предложить метод по улучшению точностных характеристик системы;
- в) реализовать метод в рамках предложенной системы;
- г) показать эффективность применения предложенного метода и его реализации на практике.

## **2. Исследование предметной области**

Подробное рассмотрение характеристик и параметров решения задач выявления и наблюдения приведены в статье [2] и находятся вне поля интересов данного текста. Обращая внимание на структуру рассматриваемой системы, двумя самыми очевидными местами, которые можно было бы

концептуально улучшить, являются следующие: способ выявления ключевых особенностей объекта и метод классификации. Авторами данной работы предлагается улучшить именно первый аспект, поэтому будет произведено рассмотрение способов описания объекта, применяемых в настоящее время. В оригинальной статье используется классический метод сбора ключевых особенностей, однако текущее положение сферы использует для этого нейросетевые подходы, их и предлагается рассмотреть. Результаты представлены в таблице 1.

Таблица 1

**Методы извлечения ключевых особенностей**

| Название                 | Количество параметров |
|--------------------------|-----------------------|
| 1                        | 2                     |
| AlexNet                  | 60 М                  |
| VGG-16                   | 138 М                 |
| VGG-19                   | 144 М                 |
| Inception-V1 (GoogleNet) | 5 М                   |
| ResNet-18                | 11,7 М                |
| ResNet-34                | 25,6 М                |
| ResNet-50                | 26 М                  |
| ResNet-101               | 44,6 М                |
| ResNet-152               | 230 М                 |
| Inception-v2             | 21,8 М                |
| Inception-v3             | 21,8 М                |
| Inception-ResNet-V2      | 55 М                  |
| Darknet-19               | 20,8 М                |
| Xception                 | 22,9 М                |
| SqueezeNet               | 1,24 М                |
| ShuffleNet               | 143 М                 |
| ShuffleNet-v2            | 2,3 М                 |
| DenseNet-100 (k = 12)    | 7,0 М                 |
| DenseNet-100 (k = 24)    | 27,2 М                |
| DenseNet-250 (k = 24)    | 15,3 М                |
| DenseNet-190 (k = 40)    | 25,6 М                |
| DetNet                   | -                     |
| EfficientNet B0 до B7    | от 5,3 М до 66 М      |
| MobileNet                | 4,2 М                 |
| MobileNet-v2             | 3,4 М                 |
| WideResNet-40-4          | 8,9 М                 |
| WideResNet-16-8          | 11 М                  |
| WideResNet-28-10         | 36,5 М                |
| SWideRNet                | от 7,77 М до 946,69 М |
| HRNet W32, W48           | 28,5 М, 63,6 М        |
| HRNet V2                 | -                     |

Существует много нейросетей, способных извлекать ключевые признаки, что доказывает их работоспособность, однако большая их часть достаточно тяжеловесна, следовательно, не может достаточно эффективно работать в системах умного дома, где мощности ограничены.

### **3. Предлагаемые изменения системы**

В работе [2] предложена система, которая решает проблему улучшения классических алгоритмов наблюдения и выявления, однако есть возможность для ее совершенствования. В данной статье за основу берется предложенная в [2] система и производится улучшение ее точностных показателей за счет изменения механизма описания объектов интереса, при этом данное решение также работает в совокупности с каким-то существующим алгоритмом наблюдения.

**Обучение.** Основные изменения в оригинальной системе производятся на данном этапе. Вместо сбора HOG-дескрипторов [4] для выявления ключевых особенностей цели предлагается использовать предобученную на общем наборе данных легковесную нейронную сеть. В данной работе по результатам анализа была выбрана сеть SqueezeNet [8] из-за с одной стороны малого количества параметров, что позволяет сохранять вычислительную эффективность, с другой стороны эффективности выполняемой задачи. Опционально можно использовать и более “тяжелые” сети, однако замедление скорости работы становится критичным, что не позволяет использовать их для систем с малой вычислительной мощностью. Таким образом на первых кадрах происходит обучение SVM-классификатора [11] за счет сбора признаков с текущего положения объекта (отклик алгоритма слежения принимается за истинное положение объекта).

**Корректировка.** Данный аспект работы [2] изменился за счет изменения способа сбора ключевых особенностей. Уточнение положения объекта в кадре происходит на основе последовательности откликов обученного классификатора. На протяжении всего времени слежения предлагаемая система осуществляет проверку корректности и точности описания истинного положения объекта в кадре. Для этого используются обученный классификатор и метод кластеризации ключевых точек. На каждом кадре после кластеризации находится такая описывающая рамка, которая лучшим образом определяет объект, для определения которого используется нейросеть SqueezeNet вместо HOG-дескриптора. Дальнейшие этапы корректировки изменениям не подвергались.

**Выявление.** Данный аспект работы [2] также изменился за счет изменения способа сбора ключевых особенностей. Изображения внутри описывающих рамок полученных кластеров описываются нейронной сетью и передаются на вход классификатору для проверки. Дальнейшие этапы выявления изменениям не подвергались.

## 4. Результаты

### 4.1. Наблюдение

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) [10]. Подход к проведению тестирования для оценки результатов:

- 1) выбор основного алгоритма с добавленной оригинальной системой наблюдения;
- 2) использование тестовой среды для оценки его характеристик;
- 3) замена системы на улучшенную;
- 4) повторный запуск полученного комплекса и сбор метрик;
- 5) анализ результатов первого и второго запусков.

В качестве улучшаемых алгоритмов сопровождения были выбраны девять различных реализаций, которые представляют собой интерес для решения задач в условиях ограничений, а данные по их работе без, с использованием оригинальной и улучшенной систем представлены на рисунке 1 и рисунке 2.

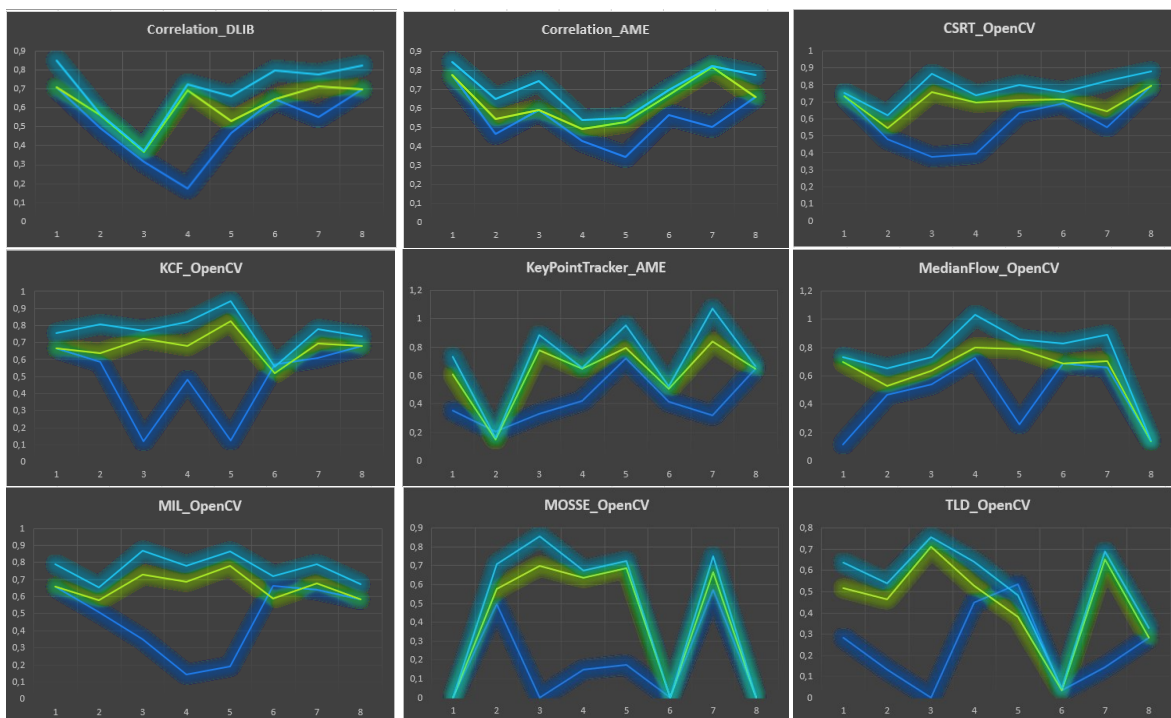


Рис. 1. Графики значений IoU: синим цветом обозначен оригинальный алгоритм, зеленым – комбинированный с оригинальной системой, голубым – улучшенная система

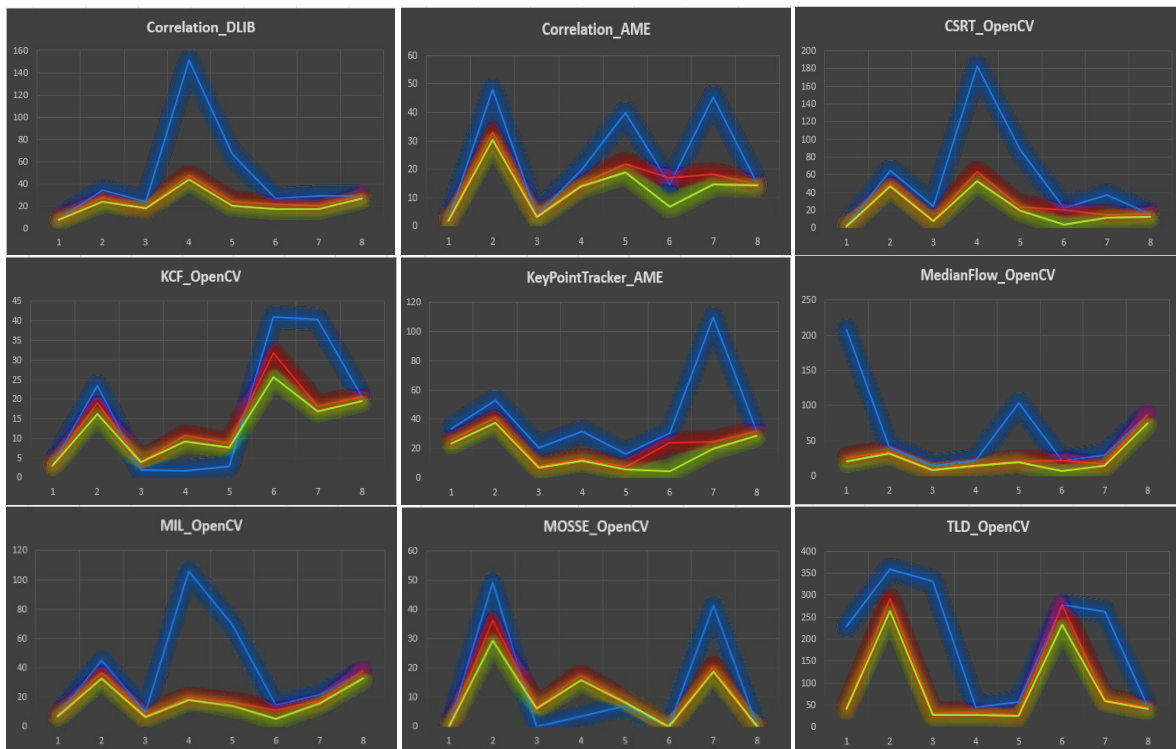


Рис. 2. Графики удаленности центров: синим цветом обозначен оригинальный алгоритм, красным – комбинированный с оригинальной системой, зеленым – улучшенная система

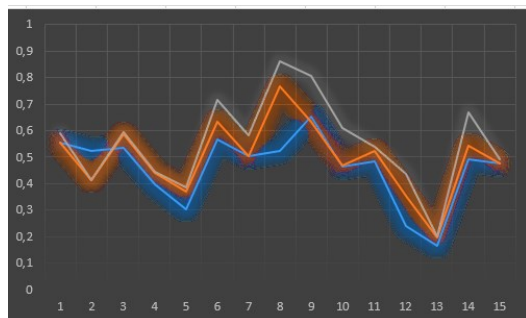


Рис. 3. Графики IoU для детекции многоагентного алгоритма (синий), оригинальной системы (оранжевый), улучшенной системы (серый)

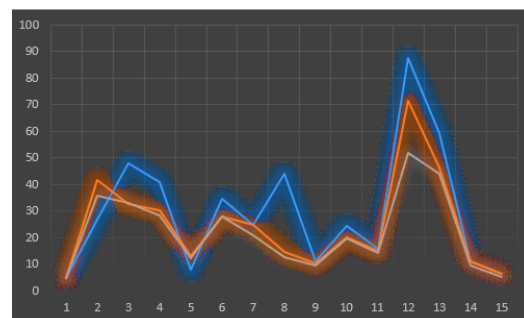


Рис. 4. Графики разницы центров для детекции многоагентного алгоритма (синий), оригинальной системы (оранжевый), улучшенной системы (серый)

## 4.2. Выявление

Для оценки данного аспекта будет использована параметрическая F-мера (F-score [7]) и многоагентный алгоритм [1].

Подход к проведению тестирования для оценки результатов:

- 1) выбор видеозаписи для проверки;
- 2) обучение классификатора на протяжении 100–400 кадров при помощи размеченного истинного положения объекта;
- 3) имитация срыва наблюдения путем перехода на другой кадр, разделенный с последним кадром обучения не менее, чем на 100 фреймов;
- 4) сбор фрагментов кадра двумя способами: метод скользящих окон разного масштаба и предложенный метод кластеризации;

- 5) проверка полученных описывающих прямоугольников классификатором;
- 6) вычисление F-меры, IoU и разницы дистанции центров.

### **4.3. Выводы**

Анализируя приведенные результаты по наблюдению, можно заметить, что в среднем по всем алгоритмам произошло улучшение показателя IoU на протяжении времени на  $\sim 14\%$  (чем выше зеленая линия над синей, тем лучше), а приближение центра к истинному на  $\sim 16\%$  (чем ниже красная линия под синей, тем лучше).

Анализируя приведенные результаты по выявлению, получаем среднее значение IoU  $\sim 0,56$ , что на  $0,06$  лучше оригинальной системы, и дистанции до центра  $\sim 21,9$ , что на  $2,7$  лучше, чем у оригинальной системы. Данные показатели свидетельствуют о достаточной точности локализации объекта после его обнаружения (keypoint detector –  $0,458$  и  $29,8$ , tld –  $0,41$  и  $30,2$ ).

Во время работы замечено замедление скорости обработки данных  $\sim 22\%$  из-за использования более долгой в обработке нейронной сети.

### **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.

### **Список литературы**

1. Бондаренко В. А., Ельцова Д. К., Лизин А. И., Павлова В. А., Созинова М. В., Тупиков В. А. Многоагентный алгоритм автоматического обнаружения и сопровождения недетерминированных объектов. // Известия ЮФУ. Технические науки. – 2020. – № 1(211). – С. 218–232.
2. Гессен П. А., Сараджишвили С. Э. Система улучшения точностных показателей видеобработки комплексов умного дома // Системный анализ в проектировании и управлении: сб. науч. трудов XXVII Междунар. науч.-практич. конф., г. Санкт-Петербург, 13–14 октября 2023 г.: в 2-х частях. – Ч. 2. – СПб.: ПОЛИТЕХ-ПРЕСС, 2024. – С. 157–165.
3. Bertinetto L., Valmadre J., Henriques J. F., Vedaldi A., Torr P. H. S. Fully-convolutional siamese networks for object tracking. // In: G. Hua, H. Jégou (eds.) Computer Vision – ECCV 2016 Workshops. ECCV 2016. Lecture Notes in Computer Science. – Vol. 9914. – Cham: Springer, 2016. – Pp. 850–865. – DOI: [https://doi.org/10.1007/978-3-319-48881-3\\_56](https://doi.org/10.1007/978-3-319-48881-3_56).
4. Dalal N., Triggs B. Histograms of oriented gradients for human detection // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 2005. – Vol. 1. – Pp. 886–893. – DOI:10.1109/CVPR.2005.177.
5. Dilek E., Dener M. Computer vision applications in intelligent transportation systems: a survey // Sensors. – 2023. – Vol. 23(6). – P. 2938. – DOI:10.3390/s23062938.
6. Gao Junfeng, Yang Yong, Lin Pan, Park Dong. Computer vision in healthcare applications // Journal of Healthcare Engineering. – 2018. – Pp. 1–4. – DOI:10.1155/2018/5157020.
7. Goutte C., Gaussier E. A probabilistic interpretation of precision, recall and f-score, with implication for evaluation // European Conference on Information Retrieval. – Springer, 2005. – Pp. 345–359.

8. Iandola F. N., Han Song, Moskewicz M. W, Ashraf Kh., Dally W. J., Keutzer K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size // ArXiv. – Submitted on February 24, 2016. – DOI:<https://doi.org/10.48550/arXiv.1602.07360>.
9. Debnath R., Bhowmik M. K. 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(12). – Paper 103165. – ISSN 1047-3203. – DOI:[10.1016/j.jvcir.2021.103165](https://doi.org/10.1016/j.jvcir.2021.103165).
10. Rezatofighi 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), Long Beach, CA, USA, 2019. – Pp. 658–666.
11. Shalev-Shwartz S., Singer Y., Srebro N. Pegasos: primal estimated sub-gradient solver for SVM // Math. Program. – 2020. – № 127. – Pp. 3–30.
12. Strat T., Chellappa R., Patel V. Vision and robotics // AI Magazine. – 2020. – Vol. 41. – No. 2. – Pp. 49–65. – DOI:[10.1609/aimag.v41i2.5299](https://doi.org/10.1609/aimag.v41i2.5299).
13. Zhou L., Zhang L., Konz N. Computer vision techniques in manufacturing // IEEE Transactions on Systems, Man, and Cybernetics: Systems. – 2023. – Vol. 53. – No. 1 (Jan. 2023). – Pp. 105–117. – DOI:[10.1109/TSMC.2022.3166397](https://doi.org/10.1109/TSMC.2022.3166397).