

**Бочаров Никита Алексеевич** – ПАО «ИНЭУМ им. И.С. Брука»; e-mail: bocharov.na@phystech.edu; г. Москва, Россия, тел.: +79167346437; к.т.н.; зам. руководителя управления; г.н.с.

**Бычков Игнат Николаевич** – e-mail: ignat\_b@ineum.ru, тел.: +74991353321; д.т.н.; зам. генерального директора.

**Корнев Павел Валерьевич** – e-mail: ineum@ineum.ru; тел.: +74991353321; соискатель.

**Парамонов Николай Борисович** – e-mail: paramonov\_n\_b@mail.ru; тел.: +74991355336; д.т.н.; профессор; руководитель управления; г.н.с.

**Bocharov Nikita Alexeevich** – JSC «INEUM»; e-mail: bocharov.na@phystech.edu; Moscow, Russia; phone: +79167346437; cand. of eng. sc.; deputy head of department; chief scientific officer.

**Bychkov Ignat Nikolaevich** – e-mail: ignat\_b@ineum.ru; phone: +74991353321; dr. of eng. sc.; deputy general director.

**Korenev Pavel Valerievich** – e-mail: ineum@ineum.ru; phone: +74991353321; candidate.

**Paramonov Nikolay Borisovich** – e-mail: paramonov\_n\_b@mail.ru; phone: +74991355336; dr. of eng. sc.; head of department; chief scientific officer.

УДК 004.89

DOI 10.18522/2311-3103-2024-1-267-276

**В.В. Ковалев, Н.Е. Сергеев**

## **РАСШИРЕНИЕ ПРИЗНАКОВОГО ПРОСТРАНСТВА В ЗАДАЧЕ ПОИСКА И РАСПОЗНАВАНИЯ МАЛОРАЗМЕРНЫХ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ**

*Одним из актуальных направлений при создании систем раннего обнаружения объектов является разработка алгоритмов поиска и распознавания малоразмерных объектов на изображениях. В задаче раннего обнаружения приходится распознавать объекты на дальних расстояниях от места их фиксации камерой. Образ на изображении таких объектов представлен малой компактной группой пикселей, которая претерпевает пространственные и яркостные изменения от кадра к кадру. Для успешного решения этой задачи целевые объекты реального мира должны иметь большие физические размеры. Кроме физических размеров объекта на образ объекта на изображении влияют большое количество факторов: разрешение матрицы камеры, фокусное расстояние объектива, светочувствительность матрицы и др. Вектор решения такой задачи направлен в сторону сверточных нейронных сетей. Однако, даже у передовых архитектур сверточных нейронных сетей поиск и распознавание малоразмерных объектов на изображениях вызывает трудности. Эта проблема напрямую связана с эффектом переобучения модели нейронной сети. Переобучение модели нейронной сети можно оценить на основе анализа кривых обучения. Для снижения вероятности переобучения применяют специальные методы, которые объединяет термин регуляризация. Однако, в распознавании малоразмерных объектов существующих методов регуляризации бывает недостаточно. В работе произведено исследование разработанного алгоритма предварительной обработки последовательности видеок кадров, увеличивающего исходное пространство признаков новым независимым признаком движения в кадре. Алгоритм предварительной обработки основан на пространственно-временной фильтрации последовательности видеок кадров, применение которого распространяется на широкий спектр архитектур сверточных нейронных сетей. Для исследования характеристик точности и распознавания сверточных нейронных сетей сформированы датасеты изображений в градациях серого и изображений с признаком движения на основе среды разработки 3D графики Unreal Engine 5. В работе приведен критерий малоразмерности объектов на изображениях. Произведено обучение и оценка характеристик точности тестовой модели сверточной нейронной сети и анализ динамики кривых обучения тестовой модели. Показано положительное влияние предложенного алгоритма предварительной обработки последовательности видеок кадров на интегральную точность обнаружения малоразмерных объектов.*

*Обнаружение малоразмерных объектов; сверточные нейронные сети; подвижные объекты; переобучение нейронных сетей.*

V.V. Kovalev, N.E. Sergeev

## EXPANSION OF THE FEATURE SPACE IN THE TASK OF SMALL OBJECT DETECTION IN IMAGES

*One of the current trends in creating early object detection systems is the development of algorithms for searching and recognizing small objects in images. In the early detection task, it is necessary to recognize objects at long distances from the place where they were recorded by the camera. The image in the image of such objects is represented by a small compact group of pixels, which undergoes spatial and brightness changes from frame to frame. To successfully solve this problem, real-world target objects must have large physical dimensions. In addition to the physical dimensions of the object, the image of the object in the image is influenced by a large number of factors: the resolution of the camera matrix, the focal length of the lens, the photosensitivity of the matrix, etc. The vector for solving this problem is directed towards convolutional neural networks. However, even advanced convolutional neural network architectures face challenges in finding and recognizing small objects in images. This problem is directly related to the effect of overtraining the neural network model. Retraining of a neural network model can be assessed based on learning curve analysis. To reduce the likelihood of overfitting, special methods are used, which are united by the term regularization. However, in recognizing small-sized objects, existing regularization methods are not enough. The work examines the developed algorithm for preprocessing a sequence of video frames, which increases the original feature space with a new independent feature of movement in the frame. The preprocessing algorithm is based on spatiotemporal filtering of a sequence of video frames, the application of which extends to a wide range of convolutional neural network architectures. To study the characteristics of accuracy and recognition of convolutional neural networks, datasets of grayscale images and images with a sign of motion were generated based on the 3D graphics development environment Unreal Engine 5. The work presents a criterion for the small size of objects in images. The accuracy characteristics of the test model of the convolutional neural network were trained and assessed, and the dynamics of the learning curves of the test model were analyzed. The positive influence of the proposed algorithm for preprocessing a sequence of video frames on the integral accuracy of detection of small-sized objects is shown.*

*Small objects detection; convolutional neural networks; moving objects; retraining of neural networks.*

**Введение.** В настоящее время одним из активно развивающихся направлений является разработка бортовых систем обнаружения объектов, предназначенных для установки на автономные носители. Такие системы требуют раннего обнаружения объектов. Раннее обнаружение подразумевает обнаружение объектов на дальних расстояниях от места их фиксации камерой. Объекты на дальних расстояниях представлены малой группой пикселей на цифровых изображениях. Размер объекта на изображении зависит от многих факторов: разрешение матрицы камеры, фокусное расстояние объектива, дальность до объекта, физические размеры объекта и др.

Типовыми алгоритмами, позволяющими решать задачу поиска и распознавания малоразмерных объектов на изображениях, являются сверточные нейронные сети (СНС). СНС условно можно разделить на одноэтапные и двухэтапные подходы. Двухэтапные СНС обеспечивают большую точность распознавания по сравнению с одноэтапными, но проигрывают в вычислительной сложности. Как правило, поиск и распознавание малоразмерных объектов осуществляется на изображениях в градациях серого или цветных изображениях в пространстве Red Green Blue (RGB). Стоит отметить, что даже у передовых архитектур СНС обнаружение малоразмерных объектов на изображениях вызывает трудности. Это связано с проблемой переобучения модели СНС. *Переобучение* – это негативный эффект, возникающий, когда алгоритм в процессе обучения формирует решающее правило слишком близко соответствующее тренировочному набору данных. Переобученная модель нейронной сети (НС) хорошо работает на данных, которые входили в тренировочную выборку, а на тестовых данных, которые модель НС никогда «не видела» показывает отрицательный результат. Это связано с тем, что объем при-

знакового пространства, характеризующий малоразмерные объекты на изображении, мал относительно вычислительной сложности модели НС. К тому же, современные архитектуры СНС содержат миллионы обучаемых параметров и с каждым годом вычислительные сложности алгоритмов продолжают расти в угоду более точным алгоритмам, что увеличивает вероятность переобучения моделей СНС.

Большое количество работ, посвященных проблеме поиска и распознавания малоразмерных объектов на изображениях, говорит об актуальности исследования данного направления. В 2016 году авторы Chen C., Liu M.-Y., Tuzel O., Xiao J. опубликовали одну из первых работ [1] посвященную поиску и распознаванию малоразмерных объектов на изображениях. В качестве алгоритма распознавания используется двухэтапная СНС R-CNN для предсказания ограничивающих рамок [2]. Увеличение точности нейросетевого алгоритма достигается за счет аугментации данных, применения контекстной информации и модификации архитектуры СНС. Последующие работы [3–8], посвященные распознаванию малоразмерных объектов на изображениях, используют подходы для увеличения точности распознавания объектов, в основе которых лежит парадигма подходов статьи [4].

Спектр применения двухэтапных СНС значительно уже одноэтапных СНС из-за высокой вычислительной сложности, особенно в автономных бортовых системах раннего обнаружения. Поэтому применяют одноэтапные СНС, среди которых можно выделить зарекомендованные архитектуры: You only look once (YOLO) [9], Single shot multibox detector (SSD) [10], CenterNet [11] и др. Так 2022 году авторы Akyon F. C., Onur Altinuc S., Temizel A. предложили алгоритм Slicing Aided Hyper Inference (SAHI) [12], который позволяет увеличить интегральную точность обнаружения малоразмерных объектов на изображениях. Фреймворк основан на базе одноэтапной СНС Fully convolutional one-stage object detection (FCOS) [13]. Алгоритм разбивает входное изображение на количество  $n$  подизображений с перекрытием. Каждое подизображение анализируется на наличие объектов интереса. Интегральное увеличение точности обеспечивается за счет аугментации данных и тонкой настройки (fine-tuning) модели СНС. Алгоритм SAHI подразумевает последовательную обработку изображений, что требует меньше ресурсов памяти вычислительной аппаратуры по сравнению с обработкой всего изображения. Но с другой стороны увеличивается время выполнения алгоритма. Применение SAHI целесообразно в обработке крупномасштабных изображений, где нет существенных ограничений, связанных со временем выполнения. В бортовых системах раннего обнаружения применение такого алгоритма не целесообразно.

Снижение вероятности переобучения за счет слияния мультиспектральных изображений [14, 15, 16], полученных с различных сенсорных датчиков является новым актуальным направлением регуляризации. Слияние изображений осуществляется различными способами: поканальное слияние на уровне пикселей, методом главных компонент, дискретным вейвлет преобразованием и др. Однако, не все бортовые системы оснащены необходимыми сенсорами.

Другим способом снижения вероятности переобучения СНС является увеличение признакового пространства за счет оценки последовательности видеок кадров рекурсивными алгоритмами Long Short-Term Memory (LSTM). С одной стороны, применение LSTM учитывает признаки, характеризующие временную зависимость, но с другой применение LSTM увеличивает вычислительную сложность СНС, обладают более сложной процедурой обучения и порой, могут выдавать непредсказуемый результат из-за своей рекурсивной природы.

Стоит отметить, что в рассмотренных работах не поднимается вопрос переобучения СНС и отсутствует причинно-следственная связь между переобучением модели СНС и неудовлетворительной точностью распознавания малоразмерных объектов на изображениях.

С учетом данных замечаний в настоящей статье предложен подход снижения вероятности переобучения СНС в задаче поиска и распознавания малоразмерных объектов, в основе которого лежит алгоритм предварительной обработки последовательности видеок кадров. Алгоритм увеличивает исходное признаковое пространство новыми признаками движения для двух моментов времени. Применение разработанного алгоритма распространяется на широкий спектр архитектур СНС при этом, несущественно увеличивая вычислительную сложность архитектуры СНС.

**Проблемы обучения сверточных нейронных сетей.** Обучение СНС базируется на методе обратного распространения ошибки, в основе которого лежит градиентный спуск. Обучение алгоритма происходит итеративным способом, который формально можно описать следующим образом:

- 1) из обучающей выборки случайным образом выбирается партия (батч) изображений, которая пропускается через СНС;
- 2) на основе полученного решения СНС и эталонного значения вычисляется значение ошибки, по формуле выбранной разработчиком;
- 3) методом обратного распространения ошибки вычисляются и применяются поправки к весовым коэффициентам СНС;
- 4) процесс обучения продолжается до тех пор, пока значение ошибки достигнет некоторого порогового значения.

На рис. 1 приведена возможная реализация зависимости значений ошибки НС от эпох обучения на тренировочном и тестовом выборках.

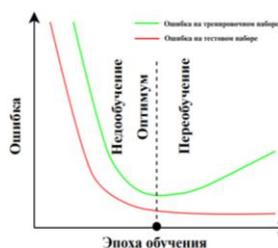


Рис. 1. Иллюстрация процесса обучения модели нейронной сети

Красная кривая иллюстрирует значение ошибки обучения НС на тренировочной выборке, а зеленая кривая значение ошибки обучения НС на тестовой выборке. До определенного момента обучения значения ошибки на тренировочной и тестовой выборке падают. На этом этапе модель НС находится в состоянии недообучения. После определенного момента ошибка обучения на тестовой выборке начинает увеличиваться. В этот момент модель НС начинает переобучаться. Модель НС в момент смены градиента направленности кривой ошибки на тестовой выборке имеет лучшее решающее правило в заданных рамках.

На рис. 2 приведена иллюстрация решающего правила (синяя кривая) модели НС на разных этапах обучения в двумерном пространстве признаков. На рис. 2 синяя кривая отображают решающие правила СНС, красные и зеленые окружности — признаки двух классов в двумерном пространстве.

Вероятность переобучиться модели СНС в задаче поиска и распознавания малоразмерных образов на изображениях больше, нежели для больших объектов. Это связано с тем, что малоразмерные объекты представлены малым количественным описанием на изображении. Для снижения вероятности переобучения применяют специальные подходы, которые объединяют термин *регуляризация*.

Для пояснения концепции предлагаемого алгоритма предварительной обработки приведем пример. На рис. 3 представлен случай отображения объектов двух классов в двумерном и трехмерном признаковом пространстве (рис. 3,а,б соответственно). Попытка построения классификатора в двумерном пространстве

неизбежно приведет к неоправданному увеличению сложности решающего правила, что тождественно переобучению, как было показано на рис. 3,а (синяя кривая  $D$ ). При этом, добавление нового признака (рис. 3,б) позволяет получить более простое решающее правило (синяя плоскость  $D$ ), однозначно разделяющее объекты двух классов  $B_1, B_2$ .

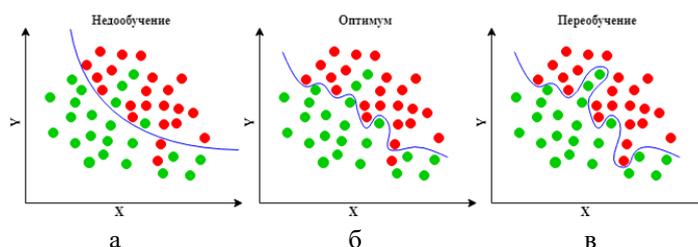


Рис. 2. Иллюстрация модели нейронной сети на разных этапах обучения

Таким образом, увеличение признакового пространства новым независимым признаком движения может снизить вероятность переобучения и повысить надежность классификации

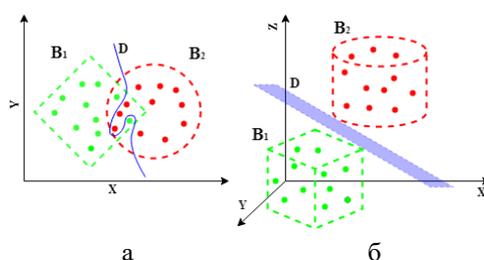


Рис. 3. Отображение множества объектов в виде точек

**Алгоритм поиска и распознавания малоразмерных объектов на изображениях.** Формально алгоритм поиска и распознавания малоразмерных объектов на изображениях можно представить в виде структурной схемы, представленной на рис. 4. Последовательное во времени изображение обрабатывается алгоритмом предварительной обработки, где формируется изображение с признаком движения. Извлечение семантических представлений высокого уровня происходит в экстракторе признаков MobileNet V2. Далее в агрегаторе признаков происходит объединение семантических представлений с разных масштабов пирамиды признаков методом Feature Pyramid Net (FPN) [17]. Декодирование зашифрованной информации по плотной сетке координат осуществляется с помощью алгоритма YOLO. Предсказанные ограничивающие рамки фильтруются алгоритмом неадекватного подавления Non Maximum Suppression (NMS) [18].

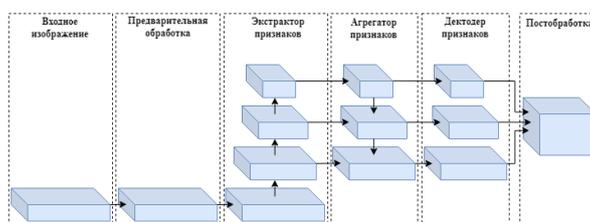


Рис. 4. Структурная схема алгоритма поиска и распознавания объектов на изображениях

В последнее время большое внимание уделяется методам регуляризации, основанных на изменении данных за счет алгоритмов предварительной обработки изображений. Вся совокупность алгоритмов предварительной обработки изображений условно можно разделить на категории, рассмотренные в [19].

**Разработка алгоритма предварительной обработки.** В данном разделе рассмотрен разработанный алгоритм предварительной обработки изображений, формирующий псевдоцветное изображение с признаком движения. Структурная схема алгоритма представлена на рис. 5 [20].

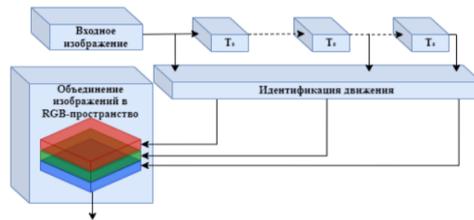


Рис. 5. Структурная схема алгоритма слияния последовательности изображений,  $T_0$  – период появления изображений камерой

На вход алгоритма поступают последовательные во времени изображения, где накапливаются в линии задержки с отводами. Далее три изображения, полученные в различные моменты времени: настоящий момент времени,  $T_0$  и  $3T_0$  поступают в блок идентификации движения. Временная задержка  $T_0$  определяется частотой выдачи кадров камерой. Блок идентификации движения формирует два разностных изображения:

$$d_{ij}(x, y) = |f(x, y, t_i) - f(x, y, t_j)|, \quad (1)$$

$$d_{ik}(x, y) = |f(x, y, t_i) - f(x, y, t_k)|, \quad (2)$$

где  $f(x, y, t_i)$ ,  $f(x, y, t_j)$ ,  $f(x, y, t_k)$  – дискретные изображения, полученные в моменты времени  $t_i$ ,  $t_j$ ,  $t_k$ ;  $x$ ,  $y$  – пространственные координаты. Изображение, полученное в момент времени  $t_i$ , является опорным изображением. Разностные изображения  $d_{ij}$ ,  $d_{ik}$  формируются путем сравнения опорного изображения с изображениями, полученными в моменты времени  $t_j = T_0$ ,  $t_l = 3T_0$ . Таким образом,  $d_{ij}$ ,  $d_{ik}$  характеризуют медленное и быстрое наличие движения. Алгоритм идентификации движения разработан с целью достижения компромисса между вычислительной сложностью и точностью идентификации движения. Подразумевается, что алгоритм предварительной обработки использует изображения со статичным фоном. Подвижность фона можно компенсировать применением стабилизационного алгоритма. Применение алгоритма предварительной обработки распространяется на широкий спектр СНС. Новизной предложенного алгоритма является использование двух разностных изображений

**Экспериментальное исследование динамики обучения сверточных нейронных сетей.** В экспериментальной части работы произведено исследование характеристик обучения, а также оценка характеристик точности поиска и распознавания малоразмерных объектов на изображениях тестовой модели СНС на основе критерия Mean Average Precision (mAP) [21], который характеризует интегральную точность обнаружения по всем классам:

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i, \quad (3)$$

где  $AP$  - Average Precision критерий, характеризующий среднюю точность поиска и распознавания объектов на изображениях одного класса. В свою очередь  $AP$  есть площадь под кривой точность-полнота (precision-recall).

Для проведения данного эксперимента сформирован датасет для задачи поиска и распознавания малоразмерных объектов. Датасет состоит из синтетических изображений, сформированных с помощью среды разработки 3D графики Unreal Engine 5. Сгенерировано 80000 изображений из 80 видеосюжетов с различными фонами и присутствием на них объектов малых размеров. Весь датасет разбит на тренировочную и тестовую выборки в процентном соотношении 70° к 30° соответственно. Для достижения релевантности экспериментальных результатов изображения в тренировочной и тестовой выборках имеют различные фона. Изображения характерны статичным задним фоном и наличием малоразмерных объектов пяти шести техники. Изображения представлены в градациях серого от 0 до 255 градаций яркости и имеют разрешение 640 на 640 элементов яркости с частотой появления 25 кадр/с. Критерий малоразмерности объектов на изображениях сформулирован следующим образом:

- 1) физический размер объектов реального мира не превышает 6 м;
- 2) отношение площади рамки, обрамляющей объект, и площади видеокадра находится в пределах от 0.08 до 0.58.

На основе исходных видеосюжетов получены тестовые и тренировочные выборки псевдоцветных изображений с признаком движения. Изображения в датасетах получены для того момента времени.

В экспериментальном исследовании используется тестовая модель СНС с вычислительной сложностью  $5.934 \cdot 10^9$  FLOPS. СНС была обучена на исходных и псевдоцветных тренировочных выборках изображений. Кривые ошибок обучения тестовой модели на тренировочном и тестовом датасетах от эпох обучения приведены на рис. 6.

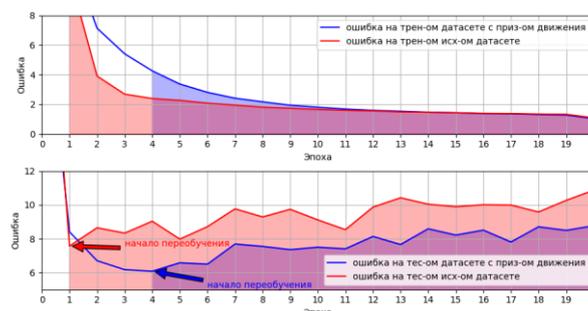


Рис. 6. Зависимости значения ошибок обучения тестовой модели СНС от эпох обучения

Из полученных зависимостей видно, что после определенной эпохи обучения кривые ошибок обучения на тренировочной и тестовой выборках начинают расходиться. В момент расхождения кривых СНС начинает переобучаться. Момент расхождения кривых для СНС, обученных на выборке псевдоцветных изображений начинается с пятой эпохи обучения, в то время как обучение СНС на исходной выборке изображений начинается со второй эпохи обучения. Это замечание говорит о том, что добавление признака движения позволяет снизить вероятность переобучения. Оценка точности тестовой модели СНС MobilenetV2-YOLO представлена на рис. 7.

Полученные оценки характеристик точности тестовой модели СНС MobilenetV2-YOLO не противоречат теоретическим предположениям: «увеличение p34-признакового пространства малоразмерных объектов снижает вероятность переобучения», что позволило увеличить интегральную точность обнаружения малоразмерных объектов на 23,65% mAP.

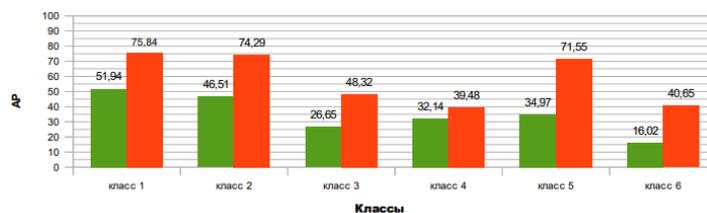


Рис. 7. Оценки AP архитектур СНС MobilenetV2-YOLO  
зеленый цвет – СНС обученная на изображениях в градациях серого; красный  
цвет – СНС обученная на псевдоцветных изображениях

**Заключение.** Разработан алгоритм предварительной обработки последовательности видеок кадров, который позволил снизить вероятность переобучения за счет увеличения исходного пространства признаков новыми независимыми признаками движения. Применение разработанного алгоритма позволило увеличить интегральную точность поиска и распознавания малоразмерных объектов на изображениях тестовой модели сверточной нейронной сети на 23,65% Mean Average Precision.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Chen C., Liu M.-Y., Tuzel O., Xiao J. R-CNN for small object detection // Asian Conference on Computer Vision, Springer, Taipei, Taiwan, November, 2016. – P. 214-230.
2. Bochkovskiy A., Wang C.-Y., Mark Liao H.-Y. YOLOv4: Optimal Speed and Accuracy of Object Detection // arXiv. – 2020. – Available at: <https://arxiv.org/abs/2004.10934>.
3. Hu G., Yang Z., Hu L., Huang L., Han J. Small Object Detection with Multiscale Features // International Journal of Digital Multimedia Broadcasting. – 2018.
4. Nguyen N.-D., Do T., Ngo T. D., Le D.-D. An evaluation of deep learning methods for small object detection // Journal of Electrical and Computer Engineering. – 2020.
5. Feng J. et al. Small Object Detection in Optical Remote Sensing Video with Motion Guided R-CNN // IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 2020. – P. 272-275.
6. Cao G., Xie X., Yang W., Liao Q., Shi G., Wu J. Feature-fused SSD: fast detection for small objects // Ninth International Conference on Graphic and Image Processing (ICGIP 2017). – 2018.
7. Zhang X., Izquierdo E., Chandramouli K. Dense and Small Object Detection in UAV Vision Based on Cascade Network // 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), Seoul, Korea (South), 2019. – P. 118-126.
8. Kisantal M., Wojna Z., Murawski J., Naruniec J., Cho K. Augmentation for small object detection // arXiv. – 2019. – Available at: 1902.07296.
9. Redmon J., Farhadi A. YOLOv3: An incremental improvement // arXiv. – 2018. – Available at: <https://arxiv.org/abs/1804.02767v1>.
10. Liu W., Anguelov D., Erhan D. et al. SSD: single shot multibox detector // European Conference on Computer Vision, vol. 9905 of Lecture Notes in Computer Science. – Springer, Cham, Switzerland, 2016. – P. 21-37.
11. Zhou X., Wang D. Objects as Points // arXiv. – 2019. – Available at: ArXiv, abs/1904.07850.
12. Akyon F.C., Onur Altinuc S., Temizel A. Slicing Aided Hyper Inference and Fine-Tuning for Small Object Detection // IEEE International Conference on Image Processing (ICIP), Bordeaux, France, 2022. – P. 966-970.
13. Zhi T., Chunhua S., Hao C., Tong H. FCOS: Fully convolutional one-stage object detection // Proceedings of the IEEE International Conference on Computer Vision (ICCV). – 2019. – P. 9627-9636.
14. Smeelen M., Schwering P., Toet A., Loog M. Semi-hidden target recognition in gated viewer images fused with thermal IR images // Inf. Fusion. – 2014. – P. 131-147.
15. Niu Y., Xu S., Wu L., et al. Airborne infrared and visible image fusion for target perception based on target region segmentation and discrete wavelet transform[J] // Mathematical Problems in Engineering. – 2012.
16. Liu S., Liu Z. Multi-Channel CNN-based Object Detection for Enhanced Situation Awareness // arXiv. – 2020. – Available at: <https://arxiv.org/abs/1712.00075>.

17. Lin T.-Y., Dollár P., Girshick R., He K., Hariharan B., Belongie S. Feature pyramid networks for object detection // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. – 2017. – P. 2117-2125.
18. Girshick R., Donahue J., Darrell T., Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. – 2014. – P. 580-587.
19. Верхаген К., Дейн Р., Грун Ф. и др. Распознавание образов. Состояние и перспективы. – М.: Радио и связь, 1985. – 104 с.
20. Ковалев В.В. Сергеев Н.Е. Методы решения проблемы переобучения нейронных сетей в задаче обнаружения малоразмерных объектов на изображениях // *Международный конгресс по интеллектуальным системам и информационным технологиям «IS&IT'23»*. – 2023. – № 1. – С. 38-45.
21. Padilla R., Netto S.L., Silva E.A.B. A Survey on Performance Metrics for Object-Detection Algorithms // *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, Niteroi, Brazil, 2020. – P. 237-242.

## REFERENCES

1. Chen C., Liu M.-Y., Tuzel O., Xiao J. R-CNN for small object detection, *Asian Conference on Computer Vision, Springer, Taipei, Taiwan, November, 2016*, pp. 214-230.
2. Bochkovskiy A., Wang C.-Y., Mark Liao H.-Y. YOLOv4: Optimal Speed and Accuracy of Object Detection, *arXiv*, 2020. Available at: <https://arxiv.org/abs/2004.10934>.
3. Hu G., Yang Z., Hu L., Huang L., Han J. Small Object Detection with Multiscale Features, *International Journal of Digital Multimedia Broadcasting*, 2018.
4. Nguyen N.-D., Do T., Ngo T. D., Le D.-D. An evaluation of deep learning methods for small object detection, *Journal of Electrical and Computer Engineering*, 2020.
5. Feng J. et al. Small Object Detection in Optical Remote Sensing Video with Motion Guided R-CNN, *IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 2020*, pp. 272-275.
6. Cao G., Xie X., Yang W., Liao Q., Shi G., Wu J. Feature-fused SSD: fast detection for small objects, *Ninth International Conference on Graphic and Image Processing (ICGIP 2017)*, 2018.
7. Zhang X., Izquierdo E., Chandramouli K. Dense and Small Object Detection in UAV Vision Based on Cascade Network, *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, Seoul, Korea (South), 2019, pp. 118-126.
8. Kisantal M., Wojna Z., Murawski J., Naruniec J., Cho K. Augmentation for small object detection, *arXiv*, 2019. Available at: 1902.07296.
9. Redmon J., Farhadi A. YOLOv3: An incremental improvement, *arXiv*, 2018. Available at: <https://arxiv.org/abs/1804.02767v1>.
10. Liu W., Anguelov D., Erhan D. et al. SSD: single shot multibox detector, *European Conference on Computer Vision, vol. 9905 of Lecture Notes in Computer Science*. Springer, Cham, Switzerland, 2016, pp. 21-37.
11. Zhou X., Wang D. Objects as Points // *arXiv*. – 2019. – Available at: [ArXiv, abs/1904.07850](https://arxiv.org/abs/1904.07850).
12. Akyon F.C., Onur Altinuc S., Temizel A. Slicing Aided Hyper Inference and Fine-Tuning for Small Object Detection, *IEEE International Conference on Image Processing (ICIP)*, Bordeaux, France, 2022, pp. 966-970.
13. Zhi T., Chunhua S., Hao C., Tong H. FCOS: Fully convolutional one-stage object detection, *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2019, pp. 9627-9636.
14. Smeelen M., Schwering P., Toet A., Loog M. Semi-hidden target recognition in gated viewer images fused with thermal IR images, *Inf. Fusion*, 2014, pp. 131-147.
15. Niu Y., Xu S., Wu L., et al. Airborne infrared and visible image fusion for target perception based on target region segmentation and discrete wavelet transform[J], *Mathematical Problems in Engineering*, 2012.
16. Liu S., Liu Z. Multi-Channel CNN-based Object Detection for Enhanced Situation Awareness, *arXiv*, 2020. Available at: <https://arxiv.org/abs/1712.00075>.
17. Lin T.-Y., Dollár P., Girshick R., He K., Hariharan B., Belongie S. Feature pyramid networks for object detection, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2117-2125.
18. Girshick R., Donahue J., Darrell T., Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 580-587.

19. Verkhagen K., Deyn R., Grun F. i dr. Raspoznavanie obrazov. Sostoyanie i perspektivy [Pattern recognition. State and prospects]. Moscow: Radio i svyaz', 1985, 104 p.
20. Kovalev V.V. Sergeev N.E. Metody resheniya problemy pereobucheniya neyronnykh setey v zadache obnaruzheniya malorazmernykh ob'ektov na izobrazheniyakh [Methods for solving the problem of re-training neural networks in the task of detecting small-sized objects in images], *Mezhdunarodnyy kongress po intellektual'nym sistemam i informatsionnym tekhnologiyam «IS&IT'23»* [International Congress on Intelligent Systems and Information Technologies "IS&IT'23"], 2023, No. 1, pp. 38-45.
21. Padilla R., Netto S.L., Silva E.A.B. A Survey on Performance Metrics for Object-Detection Algorithms, *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, Niteroi, Brazil, 2020, pp. 237-242.

Статью рекомендовал к опубликованию д.т.н., профессор А.В. Боженюк.

**Ковалев Владислав Владимирович** – Научно-конструкторское бюро вычислительных систем, e-mail: vlad.kovalev94@mail.ru; г. Таганрог, Россия; тел.: +79525864492; программист.

**Сергеев Николай Евгеньевич** – Южный федеральный университет; e-mail: nesergeev@sfedu.ru; г. Таганрог, Россия; тел.: +79281742585; кафедра вычислительной техники; д.т.н.; профессор.

**Kovalev Vladislav Vladimirovich** – Scientific design bureau of computer systems, e-mail: vlad.kovalev94@mail.ru; Taganrog, Russia; phone: +79525864492; programmer.

**Sergeev Nikolay Evgenievich** – Southern Federal University; e-mail: nesergeev@sfedu.ru; Taganrog, Russia; phone: +79281742585; the department of Computer Science; dr. of eng. sc.; professor.

УДК 004.932.2

DOI 10.18522/2311-3103-2024-1-276-285

**К.И. Морев, П.А. Ледерер**

### **ЭКСПЕРИМЕНТАЛЬНАЯ ОЦЕНКА ПОГРЕШНОСТЕЙ ВОССТАНОВЛЕНИЯ СТРУКТУРЫ НАБЛЮДАЕМОЙ СЦЕНЫ ИЗ СЕРИИ СНИМКОВ ДВИЖУЩЕЙСЯ КАМЕРЫ**

*Статья посвящена исследованию влияния использования различных математических моделей камер, следовательно, моделей формирования изображения сцены, при восстановлении 3-Д структуры сцены из набора 2-Д снимков в процессе движения камеры (восстановлении структуры из движения, далее – СИД). Сравнительная оценка проводится для двух моделей камер: классической модели камеры центральной проекции и, сравнительно новой, модели всенаправленной камеры. В статье дано краткое описание математической модели всенаправленной камеры, описанная модель используется в ходе экспериментов, также описаны способы представления изображений от всенаправленных камер. Дополнительно дается описание математической модели классической камеры центральной проекции. Описанная модель также используется в ходе проведения экспериментов. Используемые при решении задачи восстановления структуры из движения аналитические выкладки кратко упоминаются в статье. Также дается описание алгоритма получения 3-Д координат точек наблюдаемой сцены из последовательности снимков в движении. Проведенные в рамках исследования эксперименты подробно описаны в данной статье. Раскрыт процесс установки визуальных ориентиров, определения их истинных 3-Д координат. Описаны действия по формированию наборов данных для получения сравнительных оценок. В заключении работы дан анализ результатов экспериментов, выделены модели, позволяющие уменьшить погрешности восстановления 3-Д координат наблюдаемых визуальных ориентиров.*

*Математические модели камер; восстановление структуры из движения; визуальные ориентиры; всенаправленная камера; эпиполярная геометрия.*