

УДК 004.932.2

## Применение свёрточных нейронных сетей в задаче устранения теней с фотографий

*Алексейчук А.С.\**

Московский авиационный институт (национальный исследовательский университет) (МАИ), г. Москва, Российская Федерация  
ORCID: <https://orcid.org/0000-0003-4167-8347>  
e-mail: alexejchuk@gmail.com

*Мукин Ю.Д.\*\**

Московский авиационный институт (национальный исследовательский университет) (МАИ), г. Москва, Российская Федерация  
ORCID: <https://orcid.org/0009-0003-6804-2039>  
e-mail: yurimukind@gmail.com

В статье предлагается метод удаления теней на фотографиях при помощи методов глубокого обучения. Предложенный метод состоит из нескольких этапов: разделение изображения на прямоугольные фрагменты размером 32x32 пикселя, локализация теней на каждом фрагменте, восстановление цвета затененных объектов и обратное объединение фрагментов в целое изображение. Локализация теней рассматривается как задача семантической сегментации, для ее решения разработана и обучена нейронная сеть архитектуры «энкодер-декодер». Для восстановления цвета объектов на выявленных затененных участках используется еще одна нейронная сеть на основе архитектуры CDNet. Приводятся примеры обработки изображений с помощью разработанного метода, в том числе изображений с беспилотного летательного аппарата, продемонстрировано высокое качество восстановления затененных областей.

**Ключевые слова:** машинное зрение, глубокое обучение, обработка изображений, свёрточные нейронные сети, локализация теней, семантическая сегментация.

**Для цитаты:**

*Алексейчук А.С., Мукин Ю.Д.* Применение свёрточных нейронных сетей в задаче устранения теней с фотографий // Моделирование и анализ данных. 2024. Том 14. № 1. С. 41–51. DOI: <https://doi.org/10.17759/mda.2024140103>

\**Алексейчук Андрей Сергеевич*, кандидат физико-математических наук, доцент кафедры «Математическая кибернетика», Московский авиационный институт (национальный исследовательский университет) (МАИ), г. Москва, Российская Федерация, доцент кафедры «Цифровое образование», Московский государственный психолого-педагогический



университет (ФГБОУ ВО МГППУ), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0003-4167-8347>, e-mail: [alexejchuk@gmail.com](mailto:alexejchuk@gmail.com)

**\*\*Мукин Юрий Дмитриевич**, студент, Московский авиационный институт (национальный исследовательский университет) (МАИ), г. Москва, Российская Федерация, ORCID: <https://orcid.org/0009-0003-6804-2039>, e-mail: [yurimukind@gmail.com](mailto:yurimukind@gmail.com)

## 1. ВВЕДЕНИЕ

В современном мире все более широкое распространение получают различные беспилотные летательные аппараты. Получаемые с них данные могут быть использованы для решения самых разных задач, однако эти данные подвержены быстрому устареванию из-за изменения угла падения солнечных лучей и варьирования цветовой температуры. Если изменение цветовой температуры не является значимой проблемой, так как может быть нивелировано при помощи стандартных методов обработки изображения, то изменение угла падения солнечных лучей может существенно повлиять на качество решения задачи [4]. Это связано с тем, что при изменении освещения также меняется насыщенность и форма теней, что снижает точность работы многих алгоритмов обнаружения объектов, а также препятствует работе некоторых алгоритмов выделения границ.

## 2. ПОСТАНОВКА И МЕТОД РЕШЕНИЯ ЗАДАЧИ

В данной статье рассмотрен алгоритм устранения теней, использующий в качестве входных данных изображения в формате RGB, способный локализовать затененные области и устранить или минимизировать эффект затенения.

Алгоритм можно разбить на два основных шага: локализация теней и удаление теней. Удаление теней состоит в восстановлении исходной цветовой информации на затененных участках. Оба эти шага могут быть выполнены при помощи свёрточных нейронных сетей.

Из-за большого объёма данных, которые содержат изображения, полученные с современных камер, и вычислительной сложности свёрточных нейронных сетей, обработка изображений высокого разрешения в исходном виде становится чрезвычайно ресурсозатратной задачей, а также существенно ограничивает список систем, на которых возможно применение итоговой модели. Существует два подхода решения данной проблемы: разделить изображение на части (плитки, «тайлы») меньшего размера (применить тайлинг) либо произвести масштабирование изображения. Оба подхода имеют свои плюсы и минусы, которые накладывают ограничения на спектр задач, в которых они применяются.

При применении масштабирования сохраняется геометрия объектов на изображении, но при этом теряется часть информации (могут быть утеряны мелкие детали и искажены переходные цвета), что снижает точность локализации теней.

В свою очередь, применение тайлинга не приводит к потере пространственной информации, но неудачный выбор размера плитки может привести к тому, что каждый

отдельный тайл не будет содержать достаточное количество признаков для выявления границ теней, что может значительно понизить итоговое качество работы модели, нацеленной на классификацию объектов. Проблема слишком малого размера плитки стоит особенно остро при работе с исходными изображениями высокого разрешения, что актуально для кадров, полученных с современных беспилотных летательных аппаратов. Например, на рис. 1 показано целое изображение и его фрагмент, не содержащий теней, но и не содержащий признаков, позволяющих однозначно это определить.

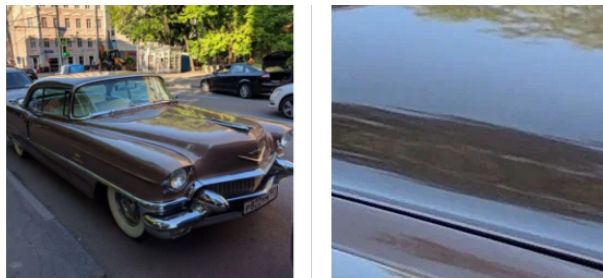


Рис. 1. Пример изображения и одной из плиток 256 x 256 пикселей

Поскольку в данной работе первоочередную роль играет обнаружение и устранение теней, что является искажениями цветового характера, было принято решение использовать тайлинг. Для обучения сетей был взят набор данных ISTD\_DATASET, представленный в статье [1].

### 3. ЛОКАЛИЗАЦИЯ ТЕНЕЙ

Задача локализации теней состоит в бинарной классификации каждого пикселя изображения (1 – «пиксель затенён»; 0 – «пиксель не затенён»). На основе этой классификации вычисляется маска тени, которая является бинарным изображением, то есть содержит только значения 0 и 1 (рис. 2). Таким образом, построение маски тени можно рассматривать как задачу семантической сегментации.

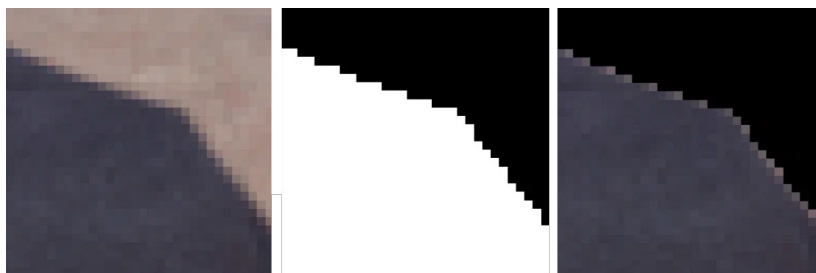


Рис. 2. Исходное изображение с тенью (слева), маска тени (посередине), наложение маски на изображение (справа)



Архитектура предлагаемой нейронной сети для локализации теней изображена на рисунке 3. Она представляет собой свёрточную нейронную сеть, состоящую из двух основных блоков – энкодера и декодера.

Энкодер содержит 6 свёрточных слоев связанных функцией активации  $ReLU(x) = \max(0, x)$ , а также вспомогательные слои BatchNorm2d и MaxPool2d, которые минимизируют внутренний ковариантный сдвиг и выделяют максимальные значения в окне заданного размера для каждого канала.

Декодер состоит из 7 слоев свёртки с дробным шагом, которые позволяют увеличить размер каждого отдельного канала за счет уменьшения количества каналов, и слоев нормализации. На последнем слое применяется функция активации  $\sigma(x) = \frac{1}{1 + e^{-x}}$ , что гарантирует попадание итогового значения в диапазон  $[0, 1]$ .

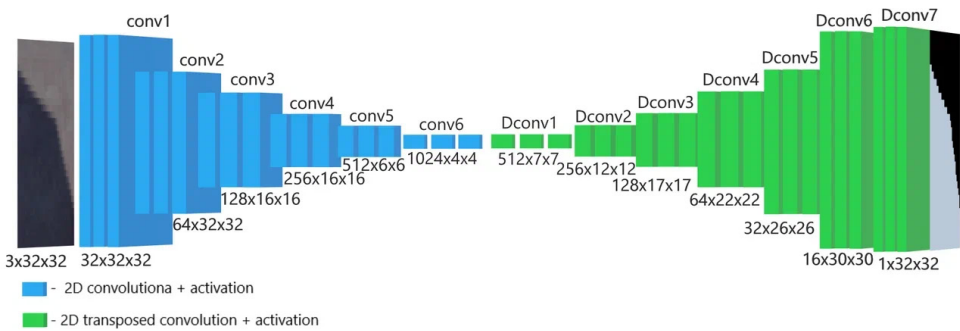


Рис. 3. Свёрточная нейронная сеть, состоящая из энкодера и декодера

В качестве функции потерь в этой модели используется категориальная кросс-энтропия:

$$I(x, y) = -\frac{1}{N} \sum_{n=1}^N \sum_{c=1}^C w_c \log \frac{\exp(x_{n,c})}{\sum_{i=1}^C \exp(x_{n,i})} y_{n,c},$$

где  $x_{n,i}$  – выходные данные последнего слоя модели с сигмоидной функцией активации (тензор вероятностей того, что  $i$ -й пиксель является затененным),  $y_{n,c}$  – истинная разметка (реальные классы пикселей),  $w_c$  – веса классов, по умолчанию инициализированные единицами,  $C$  – количество классов,  $N$  – размерность мини-пакета.

При помощи оптимизатора Adam с параметром скорости обучения  $lr = 0,0001$  было выполнено 25 эпох обучения модели. На графике на рис. 4 отображена эволюция функции потерь и метрики DICE. DICE, известный также как коэффициент Сёрнсена, представляет собой критерий качества, используемый в области компьютерного зрения для измерения сходства между двумя разметками изображений. Он основан на ошибках классификации первого и второго рода и может быть рассчитан по формуле:

$$DICE = \frac{2TP}{2TP + FN + FP},$$

где  $TP$  – количество истинно положительных результатов (классифицирован как 1, является 1),  $FN$  – количество ложноотрицательных результатов (классифицирован как 0, является 1),  $FP$  – количество ложноположительных результатов (классифицирован как 1, является 0).

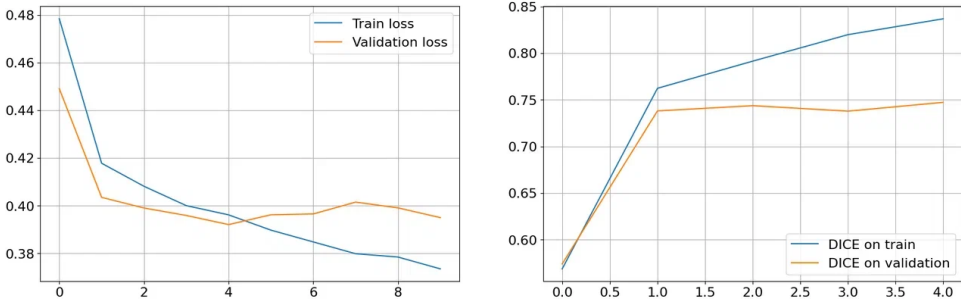


Рис. 4. Эволюция кросс-энтропии и критерия DICE на тестовой и обучающей выборках

По графикам видно, что модель явно испытывает переобучение, однако метрика DICE на тестовой выборке достаточно высока – около 75%. На рис. 5 приведен пример работы модели на данных из тестовой выборки.

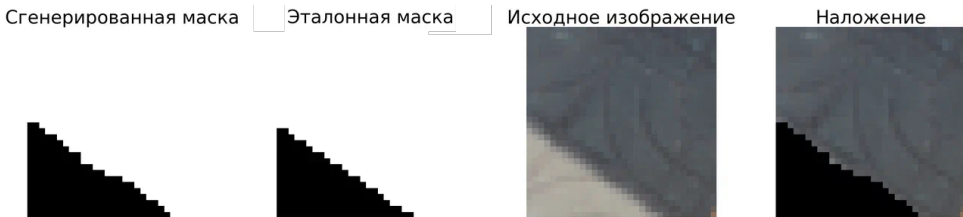


Рис. 5. Пример работы модели на тестовых данных

## 4. УДАЛЕНИЕ ТЕНЕЙ

Процесс удаления теней представляет собой восстановление оригинального цвета деталей затененных участков изображения. Для восстановления цвета используется архитектура CDNet, предложенная в работе [3]. Данная сеть состоит из классических свёрточных слоев и свёрточных слоев с дробным шагом. CDNet образована двумя и четырьмя скрытыми слоями соответственно (рис. 6). Входной слой получает обрабатываемое изображение в формате RGB, а выходной – генерирует изображение в формате RGB той же размерности, но со скорректированными значениями цветовых каналов.

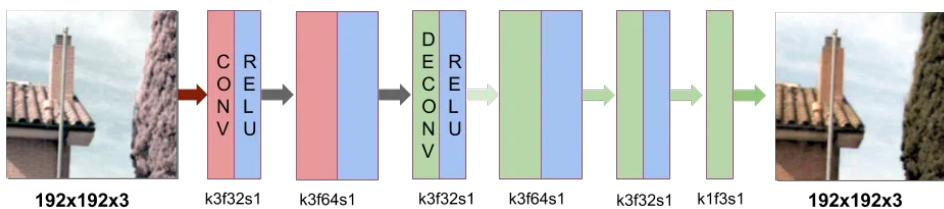


Рис. 6. Исходная архитектура CDNet

В качестве функции ошибки используется расстояние между цветами эталонного и обработанного изображений, которое можно рассчитать по формуле:

$$d = \sqrt{(R_{\text{обп}} - R_{\text{исх}})^2 + (G_{\text{обп}} - G_{\text{исх}})^2 + (B_{\text{обп}} - B_{\text{исх}})^2}.$$

После 25 эпох обучения, выполненных при помощи оптимизатора Adam с параметром скорости обучения  $lr = 0.0001$ , были получены следующие значения функции ошибки (рис. 7).

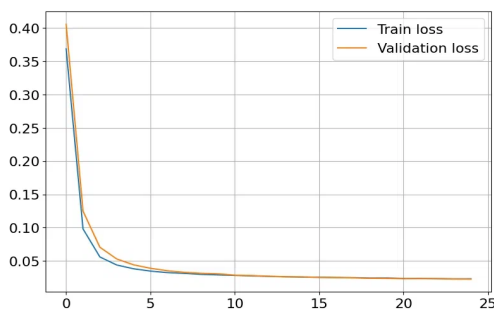


Рис. 7. Эволюция функции ошибки на тестовой и обучающей выборках

На тестовой и обучающей выборках было достигнуто почти идентичное значение функции ошибки, что свидетельствует о низкой степени переобучения, и без значительных архитектурных изменений добиться заметного улучшения невозможно. На рис. 8 приведен пример восстановления цвета в затененной области одного тайла изображения. Видно почти полное совпадение цветов в обработанной и эталонной областях.

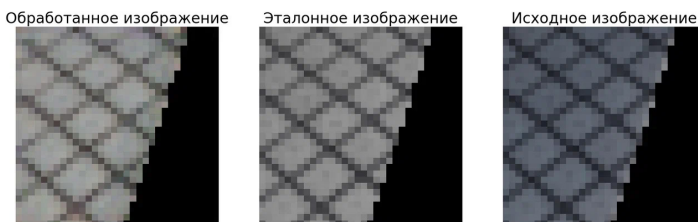


Рис. 8. Пример восстановления цветов

## 5. ОБЪЕДИНЕНИЕ МОДЕЛЕЙ

Для выполнения поставленной задачи необходимо объединить результаты работы обеих моделей, а также собрать разобранное на плитки изображение обратно в целое. Таким образом, для удаления теней необходимо выполнить следующие шаги: 1 – разделить изображение на плитки, 2 – сгенерировать маску для каждой плитки с помощью нейросетевой модели семантической сегментации, 3 – объединить каждую плитку со своей маской, 4 – восстановить цвет с помощью нейросетевой модели восстановления цвета, 5 – скомбинировать исходные плитки с плитками после удаления тени, 6 – объединить плитки в целое изображение.

Для применения маски и объединения исходного и восстановленного изображений использованы следующие функции:

$$\text{apply\_mask}(R_i, G_i, B_i, M_i) = \begin{cases} R_i, G_i, B_i, & \text{если } M_i = 1, \\ 0, 0, 0, & \text{если } M_i = 0; \end{cases}$$

$$\text{combine\_images}(R_i, G_i, B_i, \hat{R}_i, \hat{G}_i, \hat{B}_i, M_i) = \begin{cases} R_i, G_i, B_i, & \text{если } M_i = 0, \\ \hat{R}_i, \hat{G}_i, \hat{B}_i, & \text{если } M_i = 1, \end{cases}$$

где  $R_i, G_i, B_i$  – значения каналов  $i$ -го пикселя исходного изображения,  $\hat{R}_i, \hat{G}_i, \hat{B}_i$  – значения каналов  $i$ -го пикселя восстановленного изображения,  $M_i$  – значение маски этого пикселя.

На рис. 9 приведены примеры работы описанного алгоритма, включая промежуточные маски теней.

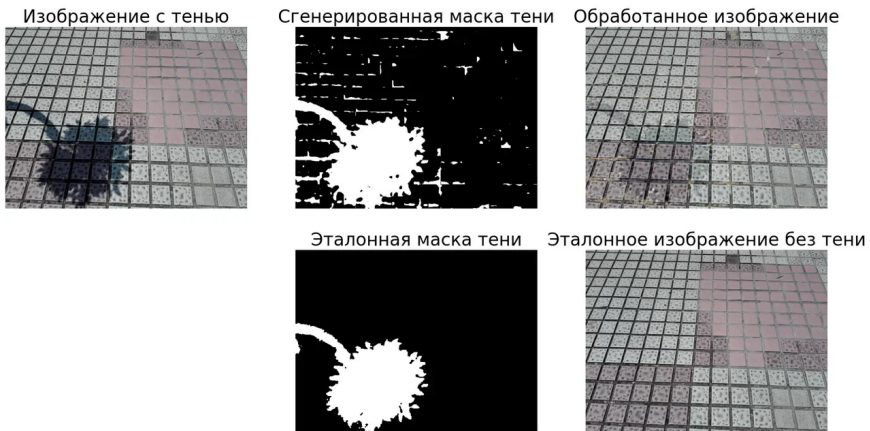


Рис. 9. Результат работы алгоритма на данных из тестовой выборки

На рис. 10 приведен пример обработки изображения, полученного с БПЛА. Видно высокое качество восстановления цвета на затененных высокими зданиями участках.



Рис. 10. Результат работы алгоритма на изображении с БПЛА

## 6. ВЫВОДЫ

Предложенный в данной работе алгоритм позволяет эффективно обнаруживать и удалять тени с изображений. Применение нейросетевого подхода позволяет не накладывать дополнительные ограничения по углу съемки и перспективе на исходные изображения, что даёт возможность использовать его не только для обработки снимков с БПЛА и спутников, но и с наземных аппаратов. Благодаря использованию относительно легких архитектур данный алгоритм может быть выполнен на широком спектре устройств и, теоретически, может значительно улучшить качество работы многих моделей глубокого обучения, предназначенных для распознавания образов, которые, в том числе, используются в автопилотах.

Однако с относительной простотой используемых моделей связаны и некоторые проблемы алгоритма. Детектор теней при классификации пикселей основывается на относительной яркости конкретной зоны, из-за чего могут возникать псевдо-тени, которые на самом деле являются просто более темными участками. К сожалению, полностью избежать этой проблемы не представляется возможным, поэтому для улучшения результатов можно совершенствовать модель восстановления цвета. Её доработка может включать расширение обучающей выборки, архитектурные улучшения и замену функции потерь на такую, которая учитывала бы неравномерность вклада в конечную яркость трёх базовых цветов.

В заключение можно отметить, что нейросетевые алгоритмы играют ключевую роль в разработке инновационных методов обработки изображений. Эффективность нейросетевого подхода к удалению теней подчеркивается его способностью к автоматическому извлечению и адаптации к контекстуальным особенностям изображения. Такой алгоритм не только значительно повышает качество конечного результата, но и обеспечивает устойчивую и точную обработку в широком спектре условий.

### *Литература*

1. Zhang X., Zhao Y., Gu C., Lu C., Zhu S. SpA-Former: An Effective and lightweight Transformer for image shadow removal // International Joint Conference on Neural Networks (IJCNN). 2023. P. 1–8.





2. *Simonyan K., Zisserman A.* Very Deep Convolutional Networks for Large-Scale Image Recognition // 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA. 2015. Conference Track Proceedings.
3. *Soria X., Sappa A., Hammoud R.* Wide-Band Color Imagery Restoration for RGB-NIR Single Sensor Images // Sensors (Basel, Switzerland). 2018. Vol.18. № 7. P. 2059.
4. *Кузьмин С.А.* Устранение влияния теней на точность выделения объектов в видеопоследовательностях // Журнал Радиоэлектроники [Электронный журнал]. № 5. 2012. URL: [jre.cplire.ru/jre/may12/2/text.html](http://jre.cplire.ru/jre/may12/2/text.html) (дата обращения 01.02.2024).



# Application of Convolutional Neural Networks in the Problem of Removing Shadows from Photographs

**Andrei S. Alekseychuk\***

Moscow Aviation Institute (MAI), Moscow, Russia

ORCID: <https://orcid.org/0000-0003-4167-8347>

e-mail: [alexejchuk@gmail.com](mailto:alexejchuk@gmail.com)

**Yuriy D. Mukin\*\***

Moscow Aviation Institute (MAI), Moscow, Russia

ORCID: <https://orcid.org/0009-0003-6804-2039>

e-mail: [yurimukind@gmail.com](mailto:yurimukind@gmail.com)

The article proposes a method for removing shadows from photographs using deep learning methods. The proposed method consists of several stages: dividing the image into rectangular fragments of 32x32 pixels, localizing shadows on each fragment, restoring the color of shadowed objects, and combining the fragments back into a whole image. Shadow localization is considered as a semantic segmentation problem; to solve it, a neural network of encoder-decoder architecture has been developed and trained. To restore the color of objects in identified shaded areas, another neural network based on the CDNet architecture is used. Examples of image processing using the developed method are given, including images from a drone, and the high quality of restoration of shaded areas is demonstrated.

**Keywords:** computer vision, deep learning, image processing, convolutional neural networks, shadow localization, semantic segmentation.

## For citation:

Alekseychuk A.S., Mukin Yu. D. Application of Convolutional Neural Networks in the Problem of Removing Shadows from Photographs. *Modelirovanie i analiz dannykh = Modelling and Data Analysis*, 2024. Vol. 14, no. 1, pp. 41–51. DOI: <https://doi.org/10.17759/mda.2024140103> (In Russ., abstr. in Engl.).

## References

1. Zhang X., Zhao Y., Gu C., Lu C., Zhu S. SpA-Former: An Effective and lightweight Transformer for image shadow removal // International Joint Conference on Neural Networks (IJCNN). 2023. P. 1–8.

\***Andrei S. Alekseychuk**, PhD (physics and mathematics), Associate Professor, Moscow Aviation Institute (National Research University) (MAI), Moscow, Russia, Associate Professor of the Department of Digital Education, Moscow State Psychological and Pedagogical University, Moscow, Russia, ORCID: <https://orcid.org/0000-0003-4167-8347>, e-mail: [alexejchuk@gmail.com](mailto:alexejchuk@gmail.com)

\*\***Yuriy D. Mukin**, student, Moscow Aviation Institute (National Research University) (MAI), Moscow, Russia, ORCID: <https://orcid.org/0009-0003-6804-2039>, e-mail: [yurimukind@gmail.com](mailto:yurimukind@gmail.com)



2. Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition // 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA. 2015. Conference Track Proceedings.
3. Soria X., Sappa A., Hammoud R. Wide-Band Color Imagery Restoration for RGB-NIR Single Sensor Images // Sensors (Basel, Switzerland). 2018. Vol.18. № 7. P. 2059.
4. Kuzmin C.A. Ustranenie vliyaniya tenej na tochnost vydeleniya obyektov v videoposledovatel'nostyah [Eliminating the influence of shadows on the accuracy of object selection in video sequences] // Zhurnal Radioelektroniki=Radioelectronics Journal [Online journal]. № 5. 2012. Available at: [jre.cplire.ru/jre/may12/2/text.html](http://jre.cplire.ru/jre/may12/2/text.html) (Accessed: 01.02.2024). (In Russ.).

Получена 27.02.2024  
Принята в печать 12.03.2024

Received 27.02.2024  
Accepted 12.03.2024