

УДК 004.032.26
DOI: 10.18799/29495407/2023/2/17

МОДЕЛЬ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ LENET5 ДЛЯ ОБНАРУЖЕНИЯ И КЛАССИФИКАЦИИ ОБЪЕКТОВ ВОЗДУШНОГО ПРОСТРАНСТВА НА ИЗОБРАЖЕНИЯХ

Клековкин Вадим Александрович¹,
vak37@tpu.ru

Марков Николай Григорьевич¹,
markovng@tpu.ru

¹ Национальный исследовательский Томский политехнический университет,
Россия, 634050, г. Томск, пр. Ленина, 30.

Для решения задачи обнаружения и классификации летающих объектов трех классов на изображениях разработана новая модель сверточной нейронной сети, архитектура которой является модификацией известной архитектуры сверточной нейронной сети LeNet5. Для ее обучения, валидации и исследования эффективности разработаны два датасета. Первый из них содержит размеченные изображения с одиночными летающими объектами трех классов: беспилотный летательный аппарат самолетного типа, включая «летающее крыло», беспилотный летательный аппарат вертолетного типа и Птица. Второй датасет наряду с изображениями первого датасета включает размеченные изображения, содержащие два и более летающих объектов этих же классов. Исследования предложенной модели сверточной нейронной сети по точности классификации летающих объектов на изображениях тестовых выборок этих датасетов показали, что модель дает высокие результаты только при решении задач распознавания летающих объектов на изображениях, когда на каждом анализируемом изображении имеется по одному такому объекту.

Ключевые слова: сверточные нейронные сети, архитектура сверточной нейронной сети LeNet5, классификация летающих объектов, беспилотный летательный аппарат самолетного типа.

Введение

В настоящее время практически важными являются задачи мониторинга воздушного пространства с целью выявления в нем и классификации летающих объектов. В числе таких объектов в воздушном пространстве рассматриваются птицы, беспилотные летательные аппараты (БПЛА) различных типов (самолетного типа, в том числе «летающее крыло», и вертолетного типа), пилотируемые летательные аппараты (самолеты, вертолеты, дельтапланы и планеры). Несмотря на то, что задачи обнаружения и распознавания подвижных объектов на изображениях изучаются достаточно давно, исследования на эту тему и на сегодняшний день являются актуальными [1].

Для решения таких задач создаются системы компьютерного зрения (СКЗ). Особое практическое значение имеют мобильные (возимые или носимые, включая встраиваемые) СКЗ на основе современных моделей сверточных нейронных сетей (СНС). Вычислительные возможности мобильных СКЗ обычно весьма ограничены, что накладывает жесткие требования на скорость вычисления моделей СНС. Это указывает на актуальность разработки и исследования новых моделей СНС и выбора из них тех, что отвечают требованиям по скорости вычислений и точности распознавания объектов на изображениях, выдвигаемым при проектировании конкретных мобильных СКЗ [2, 3].

Данная статья посвящена созданию новой модели СНС путем модификации известной архитектуры СНС LeNet5 и анализу ее эффективности при обнаружении и классификации объектов воздушного пространства трех классов на изображениях.

Постановка задачи исследования

Современным направлением в развитии СКЗ различного назначения является создание мобильных (возимых

и носимых, включая встраиваемые) вариантов СКЗ [3]. Мобильные СКЗ востребованы также при решении задач обнаружения и классификации летающих объектов в воздушном пространстве. На изображениях и видео, получаемых при мониторинге воздушного пространства с помощью соответствующего фото или видеоаппаратуры в составе таких СКЗ, могут появляться один и более летающих объектов, причём в случае нескольких объектов они могут принадлежать к разным классам. Мобильные СКЗ на основе моделей СНС должны обнаруживать, локализовать и классифицировать такие объекты.

Современные мобильные СКЗ на основе моделей СНС должны позволять решать два класса задач распознавания объектов на изображениях. Первый из них – распознавание объекта (англ. Object Recognition). При решении этих задач модель СНС получает на вход изображение с одним объектом, а на её выходе появляются координаты содержащего объект прямоугольника и его класс. Таким образом, решается задача локализации и классификации одного объекта на изображении. Второй класс задач – задачи объектного детектирования (англ. Object Detection), часто говорят, детектирования объектов на изображении. В этих случаях имеется несколько объектов на изображении и модель СНС должна обвести прямоугольником и предсказать класс каждого из объектов. Эти задачи являются фундаментальными задачами компьютерного зрения и заключаются в одновременном обнаружении, локализации (определении местоположения) и классификации всех объектов на изображении. Местоположение объекта определяется координатами ограничивающего его прямоугольника (англ. bounding box).

Задачей данного исследования является создание новой модели СНС путем модификации известной архитектуры СНС LeNet5 и анализ ее эффективности при реше-

нии указанных двух классов задач: распознавания одиночного объекта и объектного детектировании нескольких объектов воздушного пространства на изображениях. В качестве таких объектов рассматриваются объекты трех классов: БПЛА самолетного типа, в том числе «летающее крыло», БПЛА вертолетного типа и птицы (одиночные и стаи птиц). По изображениям таких объектов должны быть сформированы два датасета, причем первый из них предназначен для исследования предложенной модели СНС при решении задач из первого класса, а второй датасет – при решении задач из второго класса. Поэтому анализ эффективности этой модели СНС должен проводиться в два этапа:

- 1) обучение и исследование (тестирование) модели СНС на первом датасете, который содержит только одиночные летающие объекты на изображениях;
- 2) обучение и исследование (тестирование) модели СНС на втором датасете, содержащем два и более таких объектов на изображениях.

Формирование датасетов

При создании датасетов для обучения, валидации и тестирования предложенной модели СНС осуществлялся поиск изображений с летающими объектами трех классов: БПЛА самолетного типа, включая «летающее крыло», БПЛА вертолетного типа и Птица (птицы). Подбор изображений выполнялся в соответствии с правилом, которое гласит, что изображения должны быть максимально приближены к реальным условиям (наличие на изображении одного, двух и более летающих объектов перечисленных классов, наличие фона в виде участков земной поверхности или облаков и т. д.), в которых будет задействована предлагаемая модель СНС. Также часть полученных изображений была подвергнута аугментации (создание дополнительных изображений из имеющихся данных) для увеличения объема выборок.

Все изображения были приведены к единому размеру (416×416 пикселей) и размечены с помощью средств системы Roboflow [4]. Сформированы два датасета. В первом датасете содержатся изображения только с одиноч-

ными летающими объектами. Часть изображений из набора для формирования первого датасета в качестве примера представлена на рис. 1. Можно видеть, что на изображениях присутствуют только одиночные летающие объекты рассматриваемых классов и для них используется различный фон.

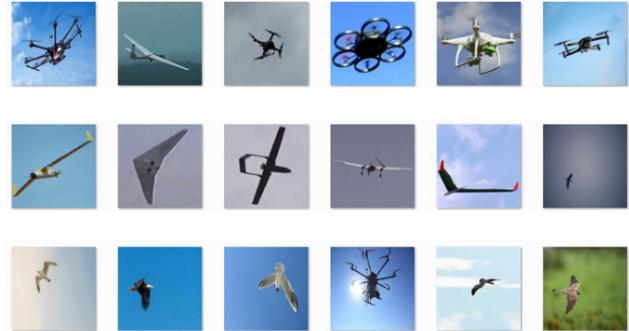


Рис. 1. Примеры изображений из набора для формирования первого датасета

Fig. 1. Examples of images from the set for forming the first dataset

Далее все размеченные изображения первого датасета объемом 2892 изображения были разделены физически на три выборки: обучающая (75 % объема датасета), валидационная (15 %) и тестовая (10 %).

Второй датасет был сформирован путем равномерного добавления по классам размеченных изображений, содержащих два и более летающих объекта, к изображениям первого датасета. После добавления таких изображений объем второго датасета составил 3927 размеченных изображений. Часть изображений второго сформированного датасета в качестве примера представлена на рис. 2. Можно видеть, что на изображениях присутствуют как одиночные объекты, так и несколько объектов рассматриваемых классов. Используется различный фон изображений. Из этого датасета также были получены три выборки, их формирование по объемам происходило в тех же пропорциях, что и в случае первого датасета.

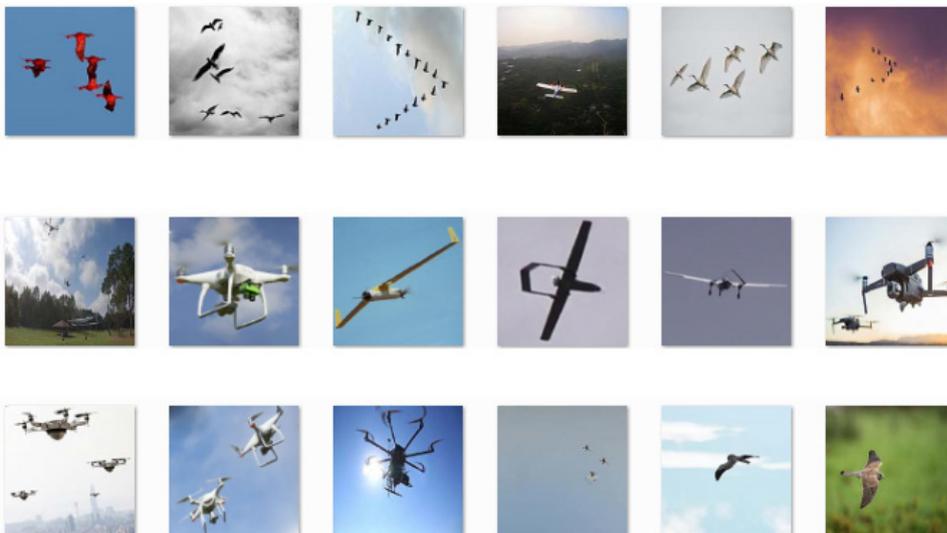


Рис. 2. Примеры изображений из набора для формирования второго датасета

Fig. 2. Examples of images from the set for forming the second dataset

На рис. 3, 4 представлены количественные характеристики датасетов по классам летающих объектов.



Рис. 3. Количество изображений по классам объектов в первом датасете

Fig. 3. Number of images by object class in the first dataset

Предложенная модель СНС

В 1998 г. Ян ЛеКун с коллегами предложил новую нейросеть, получившую название СНС. Архитектура такой сети называется LeNet5 и сегодня считается классической [5]. Эта архитектура приведена на рис. 5 и включает в себя: два сверточных слоя conv2D [6], два слоя подвыборки по максимальному значению MaxPooling2D [7] и три выходных слоя Dense [8] (адаптирована к нашей задаче классификации объектов трех классов).

Для выполнения поставленной задачи была предложена новая модель СНС с архитектурой, являющейся модификацией классической LeNet5 (рис. 6).

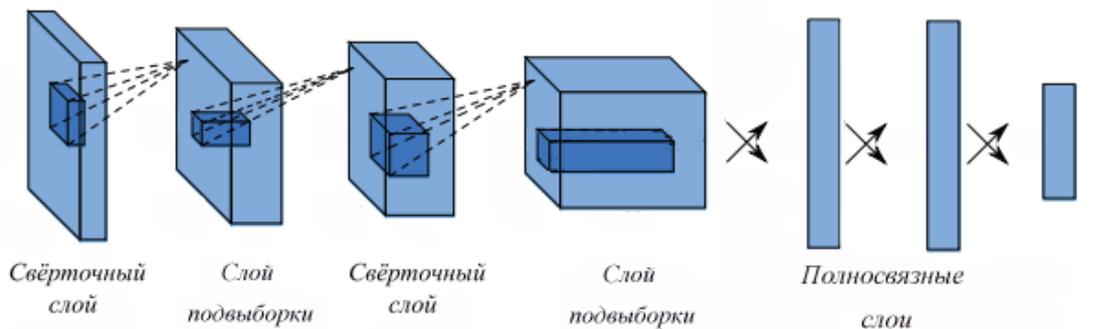


Рис. 5. Классическая архитектура LeNet5

Fig. 5. Classic LeNet5 architecture

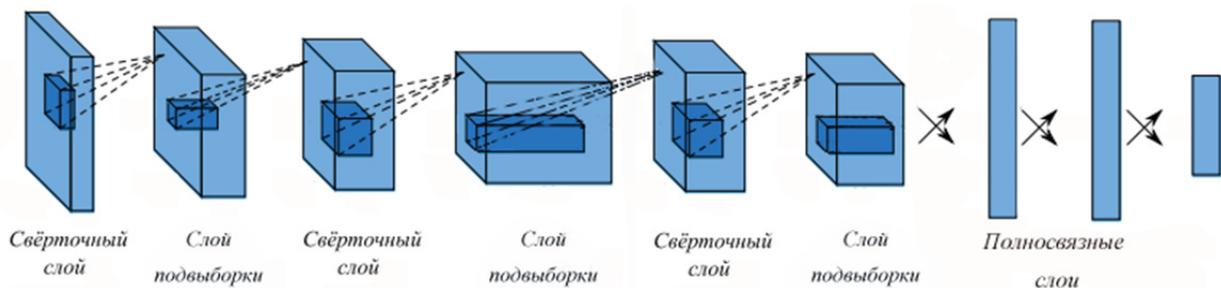


Рис. 6. Модифицированная архитектура LeNet5

Fig. 6. Modified LeNet5 architecture

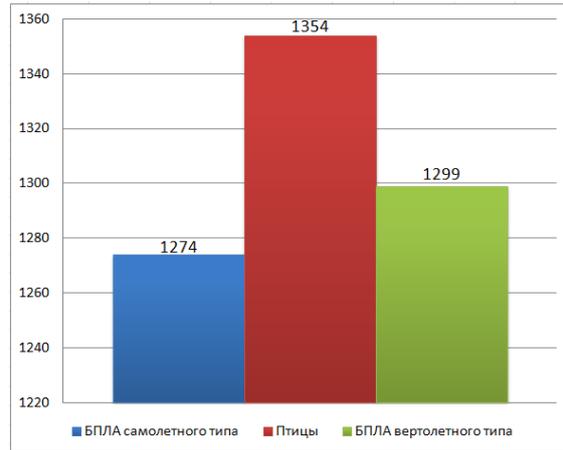


Рис. 4. Количество изображений по классам объектов во втором датасете

Fig. 4. Number of images by object class in the second dataset

Из рис. 5, 6 следует, что архитектура предложенной модели СНС отличается от классической архитектуры LeNet5 наличием трёх сверточных слоёв (conv2D) вместо двух сверточных слоёв у классической архитектуры, а также дополнительным слоем подвыборки MaxPooling2D. Новая модель СНС также отличается от классической модели LeNet5 параметрами сверточных слоёв, подобранных экспериментальным путем (количество фильтров, шаг, размер ядер свертки). Такие изменения должны привести к увеличению количества карт признаков и способствовать увеличению точности обнаружения и классификации объектов на изображениях с помощью предложенной модели.

Сравнение предложенной модели СНС с другими известными СНС класса LeNet5 позволяет считать её по ряду архитектурных признаков оригинальной.

Обучение и исследование эффективности предложенной модели СНС

Программная реализация предложенной модели СНС осуществлялась на языке программирования Python версии 3.6 с использованием библиотек TensorFlow+Keras [9].

Для оценки качества моделей СНС (точности классификации объектов) обычно используются метрики, основными из которых являются Accuracy, Average Precision, Recall, F1-score [10].

На этапе обучения предложенной модели СНС экспериментальным путем были определены следующие параметры ее настройки: оптимизатор (англ. optimizer) Adam с шагом обучения (англ. learning rate) 0,001, в качестве функции потерь (англ. loss function) использовалась функция categorical_crossentropy.

Измерение точности классификации объектов с помощью предложенной модели СНС при проведении ее обучения с использованием соответствующей выборки из первого датасета производилось с помощью метрики Accuracy, так как классы в обучающей выборке этого датасета примерно равны по количеству изображений, имеется только малый разброс. Также использовалась метрика Recall, в первую очередь, для выявления точности классификации объектов в рамках отдельных классов.

На рис. 7 в качестве примера результатов обучения и валидации модели СНС показаны графики зависимости точности классификации объектов по метрике Accuracy от количества эпох обучения и валидационной выборках первого датасета от количества эпох обучения. Результаты приведены для входных изображений размером 256×256 пикселей.

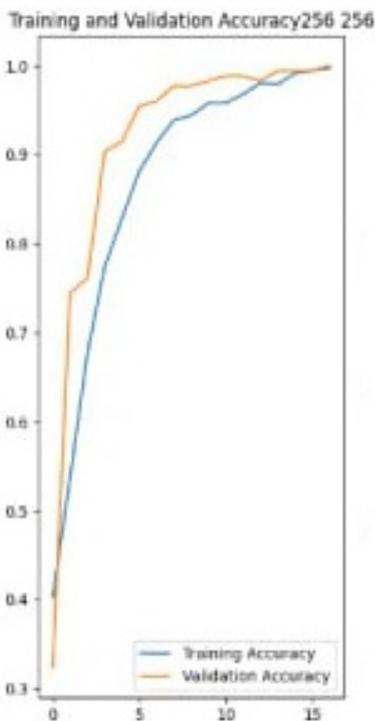


Рис. 7. Зависимость точности классификации объектов по метрике Accuracy от числа эпох

Fig. 7. Dependence of the accuracy of object classification using the Accuracy metric on the number of epochs

Результаты исследования точности классификации на тестовой выборке первого датасета в зависимости от изменения размера входных изображений (128×128; 180×180; 256×256 пикселей) представлены в табл. 1. Анализируя их, можно сделать вывод, что предложенная модель СНС весьма успешно справляется с классификацией объектов на изображениях тестовой выборки в случае одиночного объекта на изображении.

Таблица 1. Результаты исследования модели СНС на тестовой выборке первого датасета

Table 1. Results of the study of the convolutional neural networks (CNN) model on the test sample of the first dataset

Метрика Metrics	Класс объекта Object class	Размер входного изображения (в пикселях) Input image size (pixels)		
		128×128	180×180	256×256
Recall	БПЛА самолетного типа Aircraft-type UAV	0,9589	0,9387	0,9654
	БПЛА вертолетного типа Helicopter-type UAV	0,9552	0,9388	0,9323
	Птица/Bird	0,8748	0,8924	0,8933
Accuracy	Все классы All classes	0,8943	0,9105	0,9336

UAV – unmanned aerial vehicle.

При этом размер входного изображения влияет на точность классификации объектов: чем больше изображение, тем выше точность классификации объектов каждого из классов. Не очень высокий результат по точности классификации для объектов класса Птица объясняется тем, что довольно большое число изображений имеет объект Птица малых размеров. Это затрудняет их классификацию. Также можно убедиться, что по метрике Accuracy модель СНС успешно определяет и относит изображения к конкретному классу, когда в тестовой выборке используются все изображения с объектами трех классов.

Второй этап исследований эффективности предложенной модели СНС, когда изображения на входе модели содержат два и более летающих объекта, проводился с использованием второго датасета. Сначала модель обучалась и валидировалась с использованием обучающей и валидационной выборок этого датасета, а затем на его тестовой выборке проводилось собственно исследование модели СНС.

Для оценки точности классификации объектов использовались метрики Accuracy, Recall (в первую очередь, для выявления точности классификации объектов в рамках отдельных классов), Average Precision (средневзвешенное значение точности классификации объектов на изображении) и F1-score (гармоническое среднее между Precision и Recall). Применение метрики Accuracy оправдано, поскольку объем изображений в обучающей выборке второго датасета примерно одинаков для каждого из классов летающих объектов. Так как на первом этапе исследований было установлено, что размер изображения влияет на точность классификации объектов (чем больше размер изображения, тем выше точность классификации объектов), было решено использовать для исследования на втором этапе изображения размером 256×256 пикселей.

Результаты исследования модели СНС по точности классификации на тестовой выборке второго датасета представлены в табл. 2. Анализируя их и сравнивая с результатами по точности классификации летающих объектов из табл. 1, можно сказать, что они значительно хуже результатов, полученных на первом этапе исследований. Это позволяет сделать вывод о том, что даже предложенная модель СНС, архитектура которой улучшена по сравнению с классической архитектурой LeNet5, не обеспечивает высокую точность классификации изображений, когда на них присутствуют два и более летающих объекта. В этой связи исследования точности классификации объектов по отдельным классам на втором датасете не проводились. Все это означает, что предложенная модель СНС показывает высокие результаты при решении задач только первого класса при распознавании летающих объектов на изображениях, когда на каждом анализируемом изображении имеется только один объект. Модель не следует применять при решении задач второго класса, когда на изображениях находятся два и более летающих объекта.

Таблица 2. Результаты исследования модели СНС на тестовой выборке второго датасета

Table 2. Results of the study of the CNN model on the test sample of the second dataset

Метрика/Metrics	Класс объекта/Object class	Результат/Result
Recall	Все классы All classes	0,7773
Average Precision		0,7835
F1-score		0,7804
Accuracy		0,7522

СПИСОК ЛИТЕРАТУРЫ

1. Nebaba S.G., Markov N.G. Effectiveness of moving objects detecting and tracking in airspace by images in nearinfrared // *Light & Engineering*. – 2022. – V. 30. – № 2. – P. 62–69.
2. Tan M., Pang R., Le Q.V. EfficientDet: scalable and efficient object detection // *CVPR 2020*. – V. 6. – № 5. URL: <https://arxiv.org/abs/1911.09070> (дата обращения: 08.09.2023).
3. Зоев И.В., Марков Н.Г., Рыжова С.Е. Интеллектуальная система компьютерного зрения беспилотных летательных аппаратов для мониторинга технологических объектов предприятий нефтегазовой отрасли // *Известия Томского политехнического университета. Инжиниринг георесурсов*. – 2019. – Т. 330. – № 11. – С. 34–49. DOI: <https://doi.org/10.18799/24131830/2019/11/2346>
4. Roboflow. URL: <https://roboflow.com/> (дата обращения: 08.09.2023).
5. Gradient-based learning applied to document recognition / Y. Lecun, L. Bottou, Y. Bengio, P. Haffner // *Proc. of the IEEE*. – 1998. – V. 86. – Iss. 11. – P. 2278–2324. DOI: 10.1109/5.726791

Информация об авторах

Клековкин В.А., аспирант отделения информационных технологий Инженерной школы информационных технологий и робототехники Национального исследовательского Томского политехнического университета.

Марков Н.Г., доктор технических наук, профессор отделения информационных технологий Инженерной школы информационных технологий и робототехники Национального исследовательского Томского политехнического университета.

Заключение

Задача мониторинга объектов воздушного пространства на сегодняшний день является актуальной и чаще всего решается с помощью СКЗ, создаваемых на основе современных моделей СНС. Для решения задачи распознавания летающих объектов на изображениях создана новая модель СНС, архитектура которой является модификацией известной архитектуры LeNet5.

Для обучения, валидации и исследования эффективности этой модели разработаны два датасета. Первый из них содержит размеченные изображения с одиночными летающими объектами трех классов: БПЛА самолетного типа, включая «летающее крыло», БПЛА вертолетного типа и Птица (птицы). Второй датасет был сформирован путем добавления размеченных изображений, содержащих два и более летающих объекта, к изображениям первого датасета. Проведены исследования предложенной модели СНС по точности классификации летающих объектов на изображениях тестовых выборок первого и второго датасетов. Результаты исследований показали, что модель успешно справляется с обнаружением и классификацией одиночных летающих объектов на изображениях, но показывает весьма невысокую точность классификации объектов, когда на изображении находятся два и более летающих объекта. Все это позволяет рекомендовать к применению в мобильных СКЗ разработанную модель СНС только для решения задач первого класса – задач распознавания объекта (англ. Object Recognition), находящегося в единственном числе на анализируемом изображении.

6. Tf.keras.layers.Conv2D // TensorFlow. URL: https://www.tensorflow.org/api_docs/python/tf/keras/layers/Conv2D (дата обращения: 08.09.2023).
7. MaxPooling2D layer // Keras. URL: https://keras.io/api/layers/pooling_layers/max_pooling2d/ (дата обращения: 08.09.2023).
8. Dense layer // Keras. URL: https://keras.io/api/layers/core_layers/dense/ (дата обращения: 08.09.2023).
9. Module: tf.keras // TensorFlow. URL: https://www.tensorflow.org/api_docs/python/tf/keras (дата обращения: 07.09.2023).
10. Метрики в задачах машинного обучения. URL: <https://habr.com/ru/company/ods/blog/328372/> (дата обращения: 07.09.2023).

Поступила: 10.09.2023.

Принята после рецензирования: 02.10.2023.

УДК 004.032.26
DOI: 10.18799/29495407/2023/2/17

MODEL OF CONVOLUTIONAL NEURAL NETWORK LENET5 FOR DETECTION AND CLASSIFICATION OF AIR SPACE OBJECTS IN IMAGES

Vadim A. Klekovkin¹,
vak37@tpu.ru

Nikolay G. Markov¹,
markovng@tpu.ru

¹ National Research Tomsk Polytechnic University,
30, Lenin avenue, Tomsk, 634050, Russia.

To solve the problem of detecting and classifying flying objects of three classes in images, a new convolutional neural network model has been developed, the architecture of which is a modification of the well-known LeNet5 convolutional neural network architecture. Two datasets were developed for its training, validation and effectiveness research. The first of them contains labeled images with single flying objects of three classes: aircraft-type unmanned aerial vehicle, including «flying wing», helicopter-type unmanned aerial vehicle, and Bird. The second dataset, along with the images of the first dataset, includes labeled images containing two or more flying objects of the same classes. Studies of the proposed convolutional neural network model on the accuracy of classification of flying objects in images of test samples of these datasets shown that the model gives good results only when solving problems of recognizing flying objects in images, when there is one such object in each analyzed image.

Key words: convolutional neural networks, LeNet5 convolutional neural networks architecture, classification of flying objects, aircraft-type unmanned aerial vehicle.

REFERENCES

1. Nebaba S.G., Markov N.G. Effectiveness of moving objects detecting and tracking in airspace by images in nearinfrared. *Light & Engineering*, 2022, vol. 30, no. 2, pp. 62–69.
2. Tan M., Pang R., Le Q.V. EfficientDet: scalable and efficient object detection. *CVPR*, 2020, vol. 6, no 5. Available at: <https://arxiv.org/abs/1911.09070> (accessed: 8 September 2023).
3. Zhev I.V., Markov N.G., Ryzhova S.E. Intelligent computer vision system for unmanned aerial vehicles for monitoring technological objects of oil and gas industry. *Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*, 2019, vol. 330, no. 11, pp. 34–49. DOI: <https://doi.org/10.18799/24131830/2019/11/2346>
4. *Roboflow*. Available at: <https://roboflow.com/> (accessed: 8 September 2023).
5. Lecun Y., Bottou L., Bengio Y., Haffner P. Gradient-based learning applied to document recognition. *Proc. of the IEEE*, 1998, vol. 86, Iss. 11, pp. 2278–2324. DOI: 10.1109/5.726791
6. Tf.keras.layers.Conv2D. *TensorFlow*. Available at: https://www.tensorflow.org/api_docs/python/tf/keras/layers/Conv2D (accessed: 8 September 2023).
7. MaxPooling2D layer. *Keras*. Available at: https://keras.io/api/layers/pooling_layers/max_pooling2d/ (accessed: 8 September 2023).
8. Dense layer. *Keras*. Available at: https://keras.io/api/layers/core_layers/dense/ (accessed: 8 September 2023).
9. Module: tf.keras. *TensorFlow*. Available at: https://www.tensorflow.org/api_docs/python/tf/keras (accessed: 7 September 2023).
10. *Metriki v zadachakh mashinnogo obucheniya* [Metrics in machine learning problems]. Available at: <https://habr.com/ru/company/ods/blog/328372/> (accessed: 7 September 2023).

Received: 10 September 2023.

Reviewed: 2 October 2023

Information about the authors

Vadim A. Klekovkin, graduate student, National Research Tomsk Polytechnic University.

Nikolay G. Markov, Dr. Sc., professor, National Research Tomsk Polytechnic University.