

МОДЕЛЬ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ ДЛЯ ДЕТЕКТИРОВАНИЯ ОБЪЕКТОВ ВОЗДУШНОГО ПРОСТРАНСТВА НА ИЗОБРАЖЕНИЯХ

Клековкин В.А.

Томский политехнический университет, ИШИТР, студент гр. А2-39, e-mail: vak37@tpu.ru

Введение

В настоящее время практически важными являются задачи мониторинга воздушного пространства с целью выявления в нем летающих объектов. 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].

Задачей данного исследования является создание модели сверточной нейронной сети (СНС) путем модификации известной архитектуры LeNet5 и анализ эффективности этой модели при детектировании (обнаружении и классификации) объектов воздушного пространства на изображениях. В качестве таких объектов рассматриваются объекты трех классов: беспилотные летательные аппараты (БПЛА) самолетного типа, в том числе «летающее крыло», БПЛА вертолетного типа и птицы (одиночные и стаи птиц).

Формирование датасета

При создании датасета для обучения, валидации и тестирования модели СНС осуществлялся поиск изображений с летающими объектами трех классов: БПЛА самолетного типа, включая «летающее крыло», БПЛА вертолетного типа и 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 объектов перечисленных классов, наличие фона в виде участков земной поверхности или облаков и т.д.), в которых будет задействована предлагаемая модель СНС. Также изображения частично были подвергнуты аугментации (создание дополнительных изображений из имеющихся данных) для увеличения объема выборки.

Все изображения были приведены к единому размеру (416x416 пикселей) и размечены с помощью средств системы Roboflow [2]. Часть изображений сформированного датасета в качестве примера представлена на рисунке 1. Видим, что на изображениях присутствуют как одиночные объекты, так и группы объектов рассматриваемых классов и для них используется различный фон. Далее все изображения датасета объемом 2892 изображения были разделены физически на три выборки: обучающая (70%), валидационная (20%) и тестовая (10%).

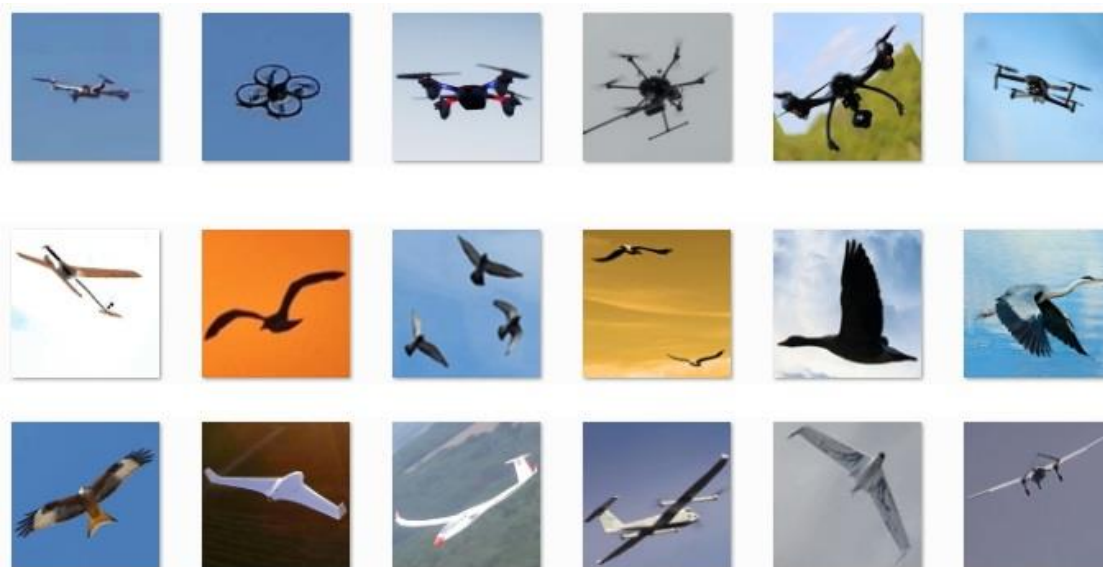


Рис. 1. Примеры изображений из сформированного датасета

Модификация архитектуры LeNet5

В 1998 г. Ян ЛеКун с коллегами предложил новую нейросеть, получившую название СНС. Архитектура такой сети называется LeNet5 и сегодня считается классической [3]. Эта архитектура включает в себя: rescaling (слой, который масштабирует входное изображение до требуемого размера); два сверточных слоя conv2D; два слоя подвыборки по максимальному значению MaxPooling2D; flatten (конвертация изображения в меньшую размерность) и три выходных слоя dense (адаптировано к нашей задаче классификации объектов трех классов) размером 1x1 пиксель.

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

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 maxpooling и различными параметрами сверточных слоёв. Такие изменения ведут к увеличению количества карт признаков и должны способствовать улучшению точности обнаружения и классификации объектов на изображениях с помощью предложенной модели.

Сравнение предложенной модели с другими известными СНС подкласса 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 [4].

Для оценки качества моделей используются метрики, основными из которых являются Accuracy, Precision, Recall [5]. Оценка точности классификации предложенной модели СНС производилась с помощью метрики Accuracy, так как классы в сформированном датасете имеют равное количество изображений. Также использовалась метрика Recall, в первую очередь, для выявления точности классификации объектов в рамках отдельных классов.

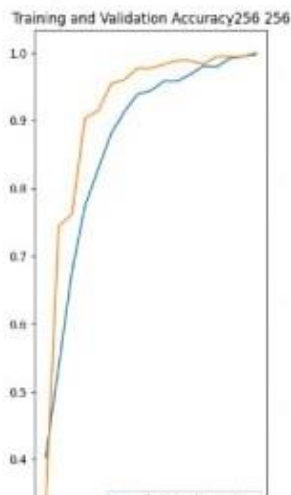


Рис. 2. Графики точности модели СНС

На этапе обучения предложенной модели СНС экспериментальным путем были определены следующие параметры ее настройки: оптимизатор (optimizer) Adam с шагом обучения (learning rate) 0.001, в качестве функции потерь (loss function) использовалась categorical_crossentropy.

На рисунке 2 показаны графики зависимости точности модели СНС на обучающей и валидационной выборках по метрике Accuracy от количества эпох обучения. Результаты приведены для входных изображений размером 256x256 пикселей.

Результаты исследования точности классификации на тестовой выборке в зависимости от изменения масштаба входных изображений (128x128; 180x180; 256x256 пикселей) представлены в таблице 1. Анализируя их, можно сделать вывод, что предложенная модель СНС весьма успешно справляется с классификацией объектов на изображениях тестовой выборки. Размер изображения влияет на точность классификации объектов: чем больше изображение, тем выше точность классификации объектов по метрике Accuracy.

Таблица 1

Результаты исследования модели СНС

Метрика	Класс объекта	Размер изображения (пиксели)		
		128x128	180x180	256x256
Recall	БПЛА самолетного типа	0,9589	0,9387	0,9654
	БПЛА вертолетного типа	0,9552	0,9388	0,9323
	птицы	0,8748	0,9024	0,8633
Accuracy	все классы	0,8943	0,9105	0,9336

Из таблицы 1 также следует, что значения метрики Recall зависят весьма сложным образом от класса объектов и размера входного изображения.

Заключение

Задача мониторинга воздушного пространства и выявления в нем различных объектов сегодня является весьма актуальной. Для решения задачи детектирования таких объектов была предложена модель СНС, архитектура которой является модификацией классической архитектуры LeNet5. Для обучения и исследования модели СНС сформирован датасет. Исследования модели показали, что она успешно справляется с классификацией изображений.

Список использованных источников

1. Stepan G. 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. – P. 62–69.
2. Roboflow [Электронный ресурс]. – URL: <https://roboflow.com/> (дата обращения: 22.01.2023).
3. Le Cun, Y. Gradient-Based Learning Applied to Document Recognition // Proc of the IEEE. – 1998. – Vol. 86. – № 11. – P. 2278–2324.
4. Module: Keras, Tensorflow [Электронный ресурс]. – URL: https://www.tensorflow.org/api_docs/python/tf/keras (дата обращения: 22.01.2023).
5. Метрики в задачах машинного обучения [Электронный ресурс]. – URL: <https://habr.com/ru/company/ods/blog/328372/> (дата обращения: 12.02.2023)