

УДК 004.89

Автоматизированная обработка видовой информации с применением сверточных нейронных сетей

Software package for automated processing of specific information using neural network technologies

Митряев / Mitryev G.

Геннадий Андреевич

(vka@mil.ru)

кандидат технических наук.

ФГБВОУ ВО «Военно-космическая академия имени

А.Ф.Можайского» МО РФ

(ВКА им. А.Ф. Можайского),

начальник научно-исследовательской лаборатории.

г. Санкт-Петербург

Шленских / Shlyonskih D.

Денис Александрович

(vka@mil.ru)

ВКА им. А.Ф. Можайского,

младший научный сотрудник.

г. Санкт-Петербург

Ключевые слова: методы и технологии искусственного интеллекта – artificial intelligence methods and technologies; аэрокосмические изображения – aerospace images; сверточные нейронные сети – convolutional neural networks; детектирование – detection.

В научной статье рассматривается проблема обнаружения, оценки параметров и классификации пространственно-распределённых объектов. Проводится обоснование выбора сверточной нейронной сети для решения задачи обнаружения. Описывается создание классификатора разномасштабных изображений. Классификатор позволяет осуществить подготовку изображений, увеличить исходную выборку, разметить объекты интереса, обучить сверточную нейронную сеть и детектировать объекты как из набора изображений, так и из видеофайла. В качестве демонстрации работы классификатора были выбраны радиолокационные изображения. Точность архитектуры разработанного обнаружителя составляет 97,1 %, что является достаточно приемлемым результатом при детектировании любых объектов на изображениях.

The scientific article deals with the problem of detection, parameter estimation and classification of spatially distributed objects. The justification of the choice of a convolutional neural network for solving the detection problem is carried out. The creation of a classifier of multispectral images is described. The classifier allows you to prepare images, increase the initial sample, mark up objects of interest, train a convolutional neural network and detect objects from both a set of images and a video file. Radar images were selected as a demonstration of the classifier's operation. The accuracy of the architecture of the developed detector is 97.1%, which is a fairly acceptable result when detecting any objects in the images.

Введение

Проблема обнаружения, оценки параметров и классификации пространственно-распределённых объектов была и остаётся определяющей при разработке и создании современных воздушных и космических средств дистанционного зондирования космической, воздушной и наземной обстановки существующими и перспективными средствами наблюдения.

Кроме того, в соответствии с указом президента Российской Федерации от 1 декабря 2016 года «О Стратегии научно-технологического развития Российской Федерации» и рядом других нормативно-правовых документов продолжается качественное и количественное повышение возможностей отечественных средств дистанционного зондирования земной поверхности в различных диапазонах электромагнитного спектра.

Долгое время задача распознавания объектов не могла быть решена автоматически. Обученный оператор на основе визуального анализа на устройствах отображения информации отраженных от объекта сигналов определял класс обнаруженного объекта без учета его ракурса и степени освещенности.

В настоящее время для решения данной проблемы применяются различные методы и средства автоматизации процессов обнаружения и классификации объектов на получаемых аэрокосмических изображениях (АКИ). Одним из таких методов является использование нейросетевых технологий (искусственного интеллекта), применение которых позволяет значительно снизить нагрузку на аналитиков данных (дешифровщиков) при наблюдении протяженных участков земной поверхности, направляя их

внимание только на те области изображений, которые трудно поддаются автоматической интерпретации.

В связи с тем, что наземная обработка АКИ является наиболее ресурсоемким процессом, одним из основных направлений совершенствования является автоматизация дешифрирования объектов интереса. Применение алгоритмов, основанных на технологиях искусственного интеллекта, позволит существенно сократить временные затраты на процесс обработки АКИ. Это обеспечит минимизацию человеческого фактора и повысит достоверность результатов дешифрирования объектов интереса при распознавании их признаков, которые не могут быть выявлены с помощью визуально-интерактивных методов. Основным инструментом для решения подобного рода задач являются сверточные нейронные сети (СНС).

Подготовка обучающих (тестовых) выборок для обучения нейросетевых моделей, применяемых для распознавания различных типов (классов) объектов на фоне подстилающих поверхностей

Подготовка обучающих (тестовых) выборок предназначена для использования при тематической обработке АКИ с помощью нейросетевых моделей в интересах формирования базы данных по типовым объектам.

Особую важность в процессе работы с нейронными сетями представляет этап их обучения. Именно от него зависит получаемая точность работы алгоритма распознавания. Для качественного обучения нейросети необходимо иметь подготовленный набор исходных данных. Однако зачастую возникает проблема, связанная с трудностью поиска наборов данных, особенно при решении частных задач, таких как распознавание и классификация объектов. В этих случаях сбор изображений и формирование из них наборов исходных данных является подготовительным обеспечивающим этапом.

Первичный набор исходных данных принято называть генеральной совокупностью. Из генеральной совокупности в последующем формируются выборки – конечные подмножества элементов генеральной совокупности, изучив которые можно получить представление об особенностях исходного множества.

Вероятностная модель порождения данных предполагает, что выборка из генеральной совокупности формируется случайным образом. Простая выборка является математической моделью серии независимых опытов и, как правило, используется для машинного обучения. При этом для каждого этапа обучения необходим свой набор данных, включающий обучающую выборку, тестовую выборку и проверочную (валидационную) выборку.

Способы формирования обучающих и тестовых выборок зависят от класса задачи, решаемой с помощью машинного обучения:

– для задач распознавания (классификации) данные следует разделять так, чтобы в полученных наборах количественное соотношение объектов разных классов было таким же, как в исходной генеральной совокупности;

– для задач регрессионного анализа необходимо обеспечить одинаковое распределение целевой переменной в полученных наборах, которые будут использоваться для обучения и контроля качества.

В задачах распознавания подготовка обучающей (тестовой) выборки является достаточно трудоемким процессом, для выполнения которого предъявляются достаточно высокие требования к квалификации исполнителя. Как правило, подготовка выборки состоит из разметки данных, в процессе которой оператор работает с каждым отдельным изображением. Она заключается в выделении на нем области или точек интереса [1]. В результате получают обработанный набор очищенных данных, пригодный для обработки алгоритмами машинного обучения.

Алгоритм разметки изображений

Для осуществления разметки и создания базы данных прежде всего необходимо осуществить выбор разносектральных изображений для обработки, а также задать классы (типы) объектов для разметки.

Разметке подлежат изображения, на которых присутствуют интересующие типы объектов. При выборе изображений для разметки следует придерживаться нескольких правил:

- чем больше объектов на снимке, тем лучше;
- чем больше снимков с различными условиями съемки одного и того же объекта, тем лучше [2].

Подготовленная база исходных данных, содержащая размеченные изображения с заданными классами объектов, называется набором данных. При построении предсказательных моделей на основе аппарата нейронных сетей набор данных разбивается на обучающую выборку данных и тестовую выборку. Обучающая выборка используется собственно для обучения той или иной модели, то есть для построения математических отношений между некоторой переменной-откликом и предикторами, тогда как тестовая выборка служит для получения оценки прогнозных свойств модели на новых данных, которые не были использованы для обучения модели. Как правило, обучающая выборка составляет 75...80% от объема набора данных, хотя каких-то строгих правил в этом отношении не существует. Данное разбиение осуществляется соответствующей программой, входящей в программный комплекс автоматизированной обработки видовой информации с применением нейросетевых технологий [3].

Зачастую подготовка подобного рода набора исходных данных, в частности обучающей выборки, требует наличия и привлечения большого количе-

ства сил и материально-технических ресурсов. Кроме того, не всегда представляется возможным получение данных по всей номенклатуре объектов для всех возможных условий их наблюдения и параметров функционирования аппаратуры.

Для решения проблемы получения обучающей выборки достаточного объема, которая будет обладать высокой точностью и устойчивостью к условиям съемки, используют искусственное расширение обучающей выборки с помощью синтезированных изображений интересующих объектов – аугментацию реальных изображений с помощью вырезания объектов с дальнейшим наложением их на изображение фона со случайным поворотом, случайным зеркальным отражением или случайным масштабированием в заданных пределах. Этот прием использовался для создания дополнительных обучающих данных на основе уже имеющихся данных и позволил увеличить исходную обучающую выборку, которая содержала ограниченное количество изображений объектов заданного класса.

Наиболее распространенными вариантами аугментации с являются: отражение по горизонтали; случайное кадрирование; изменение цвета [4].

Существует еще один эффективный способ приумножения изображений для расширения обучающей выборки – это применение генеративно-состязательной нейронной сети (GAN), представляющей собой архитектуру, которая состоит из генератора и дискриминатора, которые также являются различными нейросетями. СНС-генератор создает случайные новые экземпляры данных, СНС-дискриминатор оценивает их на подлинность. Таким образом, СНС-дискриминатор принимает решение, относится ли экземпляр данных к набору обучающих данных или нет. Однако, при использовании сети GAN для генерации изображений существуют определенные трудности с её обуче-

нием. Так, при обучении дискриминатора необходимо удерживать значения генератора постоянными и наоборот. То есть каждая сеть должна тренироваться против статичного «противника».

Выбор СНС

Проведенный сравнительный анализ доступных СНС показывает, что наиболее подходящими для целей обнаружения, селекции и классификации объектов в радиолокационном диапазоне являются *SSD*, *YOLOv5* и *RFBNet* [5]. Однако семейство моделей *YOLO* обладает рядом преимуществ, таких как быстроедействие и малый размер весов (порогов) обнаружения. *YOLO* (*You Only Look Once*) – архитектура нейронных сетей, предназначенная для детектирования объектов на изображении. Данная нейросеть решает задачи обнаружения путем разбиения изображения на квадратные области, затем классификации этих областей на наличие объекта и классификация самого объекта.

В таблице 1 приведены характеристики моделей семейства *YOLOv5*.

В качестве архитектуры обнаружителя, обеспечивающей относительно высокую точность распознавания объектов, в настоящей работе использовалась модель СНС *YOLOV5x*. Данная модель реализована на фреймворке *PyTorch* и встроена в одноименный модуль для скриптового языка программирования Python, который совместим с *Windows* и *Linux*.

Отличительной особенностью *YOLO* является подход к решению задачи детектирования, который заключается в сжатии исходного изображения таким образом, чтобы получить квадратную матрицу размером 13 на 13, в каждой клетке которой записана информация о наличии объекта и классе этого объекта на соответствующей части изображения. Такая архитектура обнаружителя предоставляет очень хорошую

Таблица 1

Характеристики СНС семейства *YOLOv5*

Тип сети	Размер РЛИ, пкс	Средняя точность (mAP^{val})	Скорость распознавания на ЦП, мс	Скорость распознавания на видеокарте <i>NVIDIA TESLA V100</i> , мс	Прочность связей	Количество операций в секунду
<i>YOLOV5n</i>	640	46,0	45	6,3	1,9	4,5
<i>YOLOV5s</i>	640	56,0	98	6,4	7,2	16,5
<i>YOLOV5m</i>	640	63,9	224	8,2	21,2	49,0
<i>YOLOV5l</i>	640	67,2	430	10,1	46,5	109,1
<i>YOLOV5x</i>	640	68,9	766	12,1	86,7	205,7

инфраструктуру как для обучения, так и для тестирования модели [6].

Обучение одноэтапного автоматического обнаружителя на базе свёрточной нейронной сети типа YOLOV5x

В работе проводился процесс обучения одноэтапного автоматического обнаружителя, который имеет один класс – корабль. Для обучения использовался обновлённый набор данных *SAR Ship Dataset* [7], который сформирован из 102 РЛИ, полученных спутником *Gaofen-3*, и 108 РЛИ, полученных спутником *Sentinel-1*. Примеры данных РЛИ представлены на рис. 1.

На основе данных РЛИ было получено 39729 изображений кораблей разрешением 256 пикселей как по дальности, так и по азимуту, которые были размечены специалистами компании SAR (все объекты отнесены к одному классу – «корабль»). Изображения в наборе данных *SAR Ship Dataset* имеют различный масштаб и фон, что позволяет использовать его для обучения нейросетевых моделей с целью обнаружения как крупномасштабных, так и малоразмерных объектов [8].

Для формирования обучающей и тестовой выборки набор данных *SAR Ship Dataset* разбили в соотношении 80% на 20% соответственно.

Рассмотрим общий принцип функционирования СНС при решении данной задачи (задачи классификации). Пусть требуется создать автоматическую систему, которая различает два типа объектов – *A* и *B*. Системы технического зрения позволяют записать данные об объектах (признаки объекта) в цифровом виде. Охарактеризуем объект с помощью набора признаков (x_1, x_2, \dots, x_n) . Признаки могут быть выражены с помощью целых чисел или даже булевских переменных, то есть «есть признак» – «нет признака».

Совокупность признаков можно рассматривать как вектор с *n* компонентами, или точку *k*-мерного

евклидова пространства $X=(x_1, x_2, \dots, x_n)$. Тогда задача классификации сводится к следующей математической задаче: разделить два множества точек *A* и *B* *n*-мерного евклидова пространства некоторой гиперповерхностью размерности *n*-1.

Обучение нейронной сети в задачах классификации происходит на наборе обучающих примеров $X(1), X(2), \dots, X(P)$ для которых принадлежность объекта к классу *A* или классу *B* известна. Чтобы математически формализовать этот факт, определим индикатор:

$$D(X) = \begin{cases} 1, & X \in A \\ 0, & X \in B \end{cases}. \quad (1)$$

По накопленному в результате обучения «опыту» построим СНС, которая проводит разделяющую поверхность. Математически этот процесс может быть описан как поиск некоторой функции $Y=F(X, W)$, где *W* – набор параметров нейронной сети. Для СНС эти параметры, в частности, задают силу связи между нейронами и подбираются так, чтобы ошибка обучения (*error training*) была бы минимальной (как можно ближе к нулю). В качестве ошибки обучения обычно рассматривают функцию

$$E_{train}(W) = \sum |F(X(j), W) - D(X(j))|, \quad (2)$$

где $X(j), j = \overline{1, P}$ берутся из обучающего множества. Для проверки эффективности обучения нейронной сети возьмём тестовое множество объектов и вычислим

$$E_{test}(W) = \sum_j |F(X(j), W) - D(X(j))|, \quad (3)$$

где $X(j)$ взяты из тестового множества.

После того, как система обучена, она для любого поданного на вход системы объекта *X* автоматически решает, к какому классу он относится. [9].

Обучение СНС будем производить без учителя, т.е. обучение, при котором испытываемая система спонтанно



Рис. 1. Примеры РЛИ из набора SAR Ship Dataset

будет обучаться выполнять поставленную задачу без вмешательства со стороны экспериментатора. Такой тип обучения пригоден только для задач, в которых известны описания множества объектов (обучающей выборки), и требуется обнаружить внутренние взаимосвязи, зависимости и закономерности, существующие между объектами.

В процессе, а также после окончания всего процесса обучения выводится справочная информация о качестве обученной модели (рис. 2).

Для оценки качества обученной модели и работы алгоритма используются метрики: *precision*, *recall*, *mAP@0.5* и *F*-метрика [10].

Precision (точность) – мера того, сколько из сделанных положительных прогнозов являются верными. Точность метрики *Precision* рассчитывается в соответствии с выражением (4). Данная метрика описывается кривой *precision-confidence* (*P_curve*) и для выбранной нейросетевой модели имеет вид, представленный на рис. 3.

$$precision = \frac{TP}{TP + FP}, \tag{4}$$

где *TP* (*True Positive*) – правильное определение класса «корабль»,

FP (*False Positive*) – ошибка 1-го рода, (принятие неправильного объекта за «корабль»).

Recall (полнота) – мера того, сколько положительных случаев правильно спрогнозировал классификатор по всем положительным случаям. Точность метрики *Recall* рассчитывается в соответствии с выражением (5), описывается кривой *recall-confidence* (*R_curve*) и для выбранной нейросетевой модели имеет вид, представленный на рис. 4.

$$recall = \frac{TP}{TP + FN}, \tag{5}$$

где *FN* (*False Negative*) – ошибка 2-го рода (пропуск объекта «корабль»).

mAP@0.5 – средняя оценка точности работы классификатора, рассчитываемая нахождением площади под кривой *precision-recall* (*PR_curve*) (рис. 5).

F-метрика (в общем случае F_β) – метрика, позволяющая объединить *Precision* и *Recall* в агрегированный критерий качества, где β определяет вес точности в метрике. Данная метрика объединяет, и в то же время взвешивает два отношения (*Precision* и *Recall*) сбалансированным образом. Она достигает максимума в случае, когда *Precision* и *Recall* равны единице, и близка к нулю, если один из аргументов близок к нулю. *F*-метрика рассчитывается в соответствии с выражением (6), описывается кривой *F1-confidence* (*F1_curve*) и для выбранной модели имеет вид, представленный на рис. 6.

$$F_\beta = (1 + \beta^2) \frac{precision * recall}{(\beta^2 * precision) + recall}. \tag{6}$$

Значения вышеперечисленных метрик, характеризующих качество обученной нейросетевой модели, сведены в таблицу 2.

Таким образом, анализируя ресурсы, затраченные на процесс обучения (рис. 2) и предельные значения метрик, характеризующих качество обучения (таблица 2), можно сделать вывод о том, что нейросетевая модель одноэтапного обнаружителя на базе СНС YOLOV5x обладает достаточно высокой точностью обнаружения морских объектов на РЛИ при сравнительно малых затратах на время обучения.

После окончания процесса обучения СНС необходимо произвести валидацию – проверку правильности работы (предсказательной способности) аналитической модели. Она проводится на независимом (то есть на не использовавшемся для обучения и тестирования) валидационном множестве.

Детектирование изображений обученным одноэтапным автоматическим обнаружителем на базе свёрточной нейронной сети

Детектирование изображений позволяет использовать обученный одноэтапный автоматический обнаружитель на базе свёрточной нейронной сети следующим образом:

- выбрать и обработать одно изображение с последующим выводом на экран результатов детектирования;

```

15 epochs completed in 19.26 hours.
Optimizer stripped from runs/train/exp2/weights/last.pt, 173.1MB
Optimizer stripped from runs/train/exp2/weights/best.pt, 173.1MB

Validating runs/train/exp2/weights/best.pt...
Fusing layers...
Model Summary: 444 layers, 86173414 parameters, 0 gradients, 204.0 GFLOPs
  Class      Images  Labels    P      R   mAP@.5 mAP@.5: .95: 100%
   all         7593    9771  0.937  0.95   0.971   0.644
    
```

Рис. 2. Характеристики обучения СНС на наборе данных SAR Ship Dataset

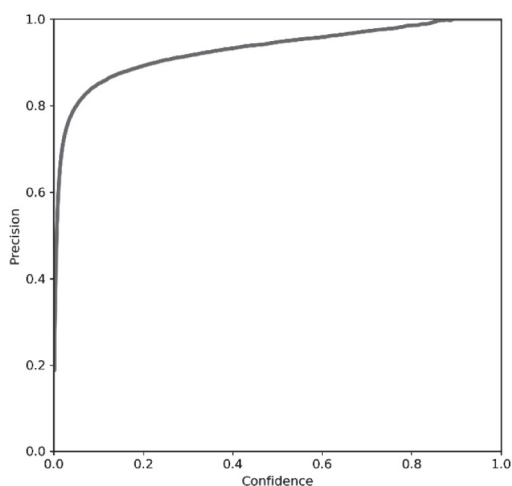


Рис.3. P_curve

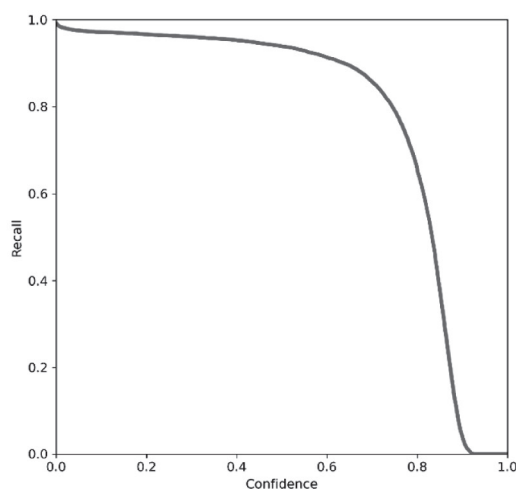


Рис. 4. R_curve

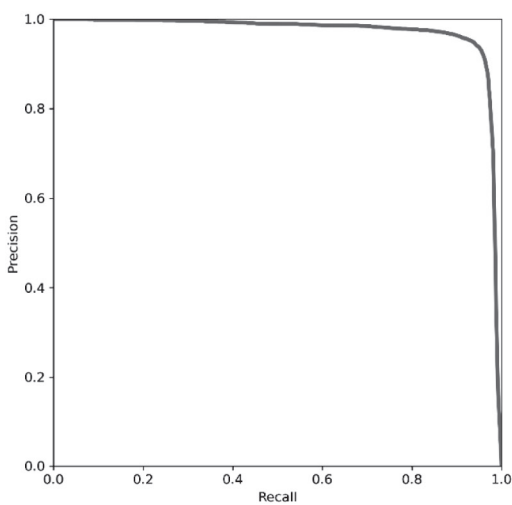


Рис. 5. PR_curve

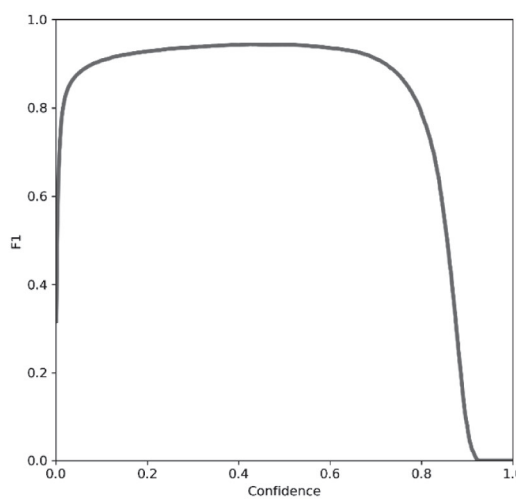


Рис.6. F1_curve



Рис. 8. РЛИ, обработанные СНС YOLOv5x

Оценка качества обученной СНС YOLOv5x

Метрика	<i>Precision</i>	<i>Recall</i>	<i>mAP@0.5</i>	<i>F1</i>
Значение точности	0.937	0.95	0.971	0.94

– выбрать и обработать массив изображений с последующим выводом на экран результатов детектирования;

– детектировать изображения из видеофайла с частотой десять кадров в секунду с последующим выводом на экран результатов детектирования;

– автоматически выбирать и вывести на экран изображения с детектированными объектами интереса из полного массива изображений.

Пример детектированного РЛИ представлен на рис. 8. Обнаруженные корабли помещены в ограничивающую рамку и подписаны согласно заданному классу изображений («*ship*»).

В случае выявления ошибок в работе автоматического обнаружителя (ложная тревога и пропуск цели) возможны несколько вариантов действия разработчика: увеличение количества эпох обучения, редактирование базы данных (обучающей выборки).

Заключение

Повышение степени автоматизации комплексной обработки АКИ, особенно на этапах обнаружения и распознавания объектов, является в настоящее время важной и актуальной задачей для повышения достоверности и скорости обработки изображений. Одним из перспективных направлений ее решения является реализация процессов обработки получаемых изображений с использованием методов и технологий искусственного интеллекта на основе СНС. При этом подготовка набора обучающих данных и выбор архитектуры обнаружителя является одной из главных задач при реализации нейросетевой модели.

В настоящей работе для достижения поставленной цели были решены следующие задачи:

– проведен сравнительный анализ доступных СНС, подходящих для целей обнаружения, селекции и классификации объектов интереса на АКИ;

– разработан программный комплекс автоматизированной обработки видовой информации с применением нейросетевых технологий, позволяющий решить целый ряд задач по автоматизации процесса наземной обработки изображений.

Для апробации работы нейросетевого обнаружителя нами был использован набор данных *SAR Ship Dataset*. Точность архитектуры разработанного обнаружителя составила 97,1 %, что является приемлемым

результатом при детектировании любых объектов на АКИ.

Научная новизна полученных результатов заключается в обеспечении решения актуальных задач:

– разметки разносектральных изображений для формирования банка обучающих (тестовых) выборок;

– автоматизированном формировании банка обучающих (тестовых) выборок из размеченных изображений;

– раскадровки видеофайла для дополнения банка обучающих (тестовых) выборок и детектирования АКИ;

– обучении СНС с использованием банка обучающих (тестовых) выборок;

– автоматическом обнаружении объектов интереса на АКИ.

Одним из путей дальнейшего исследования может быть увеличение размера исходного набора данных искусственным образом с помощью аугментации, генеративно-сопоставительных сетей и тестирования разработанного программного комплекса на реалистичных снимках, полученных в ходе лётно-экспериментальных исследований и дистанционного зондирования Земли.

Программный комплекс автоматизированной обработки видовой информации с применением нейросетевых технологий зарегистрирован Федеральной службой по интеллектуальной собственности свидетельство № 2023616221.

Литература

1. Вакуленко, С. А. Практический курс по нейронным сетям / С.А. Вакуленко, А.А. Жихарева. – Санкт-Петербург : Университет ИТМО, 2018. – 71 с.

2. Траск, Э. Грокаем глубокое обучение / Э. Траск. – Санкт-Петербург : Питер, 2019. – 352 с.

3. Гиляев, Р. А. Активное обучение и краудсорсинг: обзор методов оптимизации разметки данных / Р.А. Гиляев, Д.Ю. Турдаков // Труды Института системного программирования РАН. – 2018. – Т. 30, № 2. – С. 215–250.

4. Малыгина, Г. Ф. Элементы статистической концепции обучения нейронной сети и прогнозирование ее функционирования / Г.Ф. Малыгина, А.В. Меркушева // Научное приборостроение. – 2005. – Т. 15, № 1. – С. 29–45.

5. Кириченко, А. А. Нейропакеты – современный интеллектуальный инструмент исследователя. Сетевое электронное издание учебного пособия / А.А. Кириченко. –

Москва : Национальный исследовательский университет «Высшая школа экономики», 2013. – 297 с.

6. Чорбаа, Н. А. Сравнительный анализ методов детектирования объектов на радиолокационных изображениях при помощи нейронных сетей / Н.А. Чорбаа, Ле Ань Ту, И.М. Толстой // Научный результат. Информационные технологии. – 2020. – Т. 5, № 4. – С. 15–25.

7. SAR Ship Detection Dataset (SSDD): Official Release and Comprehensive Data Analysis / T. Zhang, X. Zhang, Jianwei Li [et al.] // Remote Sensing MDPI. – 2021. – Vol. 13 (18). – P. 3690.

8. Окунев, С. В. Рассмотрение способов формирования наборов данных для обучения нейронных сетей / С.В. Окунев // Вестник науки и образования. – 2020. – № 2-3 (80). – С. 16–19.

9. Мониторинг объектов морского судоходства по аэрокосмическим данным дистанционного зондирования в СВЧ диапазоне с применением нейросетевых технологий / М.Л. Белокопытов, Д.А. Шленских, С.В. Морозов, С.В. Сирота // Журнал радиоэлектроники. – Электронный журнал. – 2022. – № 4. – URL: <http://jre.cplire.ru/jre/apr22/index.html> (дата обращения : 01.11.2022).

10. Искусственный интеллект в космической технике. Состояние. Перспективы применения. Монография / Под ред. А.Н. Балухто. – Москва : Радиотехника, 2021. – 440 с.