

УДК 004.89

Разработка модели распознавания разнородных объектов на основе вложенных методов машинного обучения

Новикова К. Н., Новикова С. В.

Казанский национальный исследовательский технический университет им. А.Н.
Туполева-КАИ

Аннотация: В статье предложен метод идентификации объектов в случае, когда их тип (природа, происхождение) заранее не известны. В существующих системах идентификации или распознавания, как правило происхождение или общий тип объектов заранее известен. К примеру, в систему подаются зашифрованные цифровые кардиограммы, и система должна выявить патологию. Или на вход системы подаются изображения отпечатков пальцев, и задача модели состоит в определении их владельцев. Ситуации, когда тип идентифицируемого объекта не известен, практически не встречаются. В данной работе предложена двухуровневая интеллектуальная модель, позволяющая на первом этапе своего функционирования выявить тип объекта (например, является ли объект изображением, зашифрованной кардиограммой, показаниями телеметрии и пр.), а на втором определить непосредственно его вид (если объект – кардиограмма, необходимо определить патологию). Первый уровень включает кластеризующую модель, работающую по алгоритму k-средних. Второй уровень образует ансамбль моделей типа сверточных нейронных сетей CNN. Структура сверточных сетей определяется индивидуально для каждого выделенного на предыдущем уровне кластера (типа объекта).

Ключевые слова: разнотипные объекты, идентификация объектов, кластеризация объектов, модели машинного обучения, вложенные модели.

1. Введение

В настоящее время задача распознавания (идентификации) объектов по входному набору признаков встречается во множестве областей науки и техники. В первую очередь это касается распознавания изображений, когда набором определяющих тип (класс) объекта характеристик является трехмерная матрица пикселей RGB [1]. Также в литературе описаны методы распознавания состояния двигателей внутреннего сгорания по вектору выходных характеристик [2], идентификации состояния пациентов по набору биометрических показателей [3], идентификация состояния окружающей среды по комплексу мониторинговых и метеорологических показателей [4] и др.

Однако в описанных методах, как правило, все объекты и классы, подлежащие распознаванию, изначально относятся к одной, заранее заданной области. Например, в задаче распознавания дорожных знаков заранее известно, что на входы модели будут подаваться изображения дорожного полотна и прилегающих непосредственно к дороге изображений, но мы не ожидаем увидеть, например, изображения отпечатков пальцев. В задачах распознавания состояния пациента заранее известен набор входных характеристик биометрических показателей человека с областями их воз-

можных значений и т. п.

2. Задача исследования

В данной работе предлагается метод разработки модели, способной идентифицировать конкретные объекты (назовем их *виды объектов*) произвольной природы с высокой точностью, при этом заранее не известны ни возможные области происхождения объектов (назовем их *типы объектов*), ни количество таких областей-типов. Единственным требованием к данным, поступающим на вход разрабатываемой модели, является единая размерность векторов характеристик идентифицируемых элементов, что может быть достигнуто применением специальных методов предобработки и масштабирования данных [5].

3. Разработка метода решения задачи идентификации разнородных объектов

Для решения задачи предлагается строить двухуровневую вложенную интеллектуальную модель, на первом уровне которой осуществляется «грубая обработка» данных, позволяющая выявить конкретные наборы типов объектов, а на втором уровне происходит точное распознавание класса объекта внутри выделенных на предыдущем этапе типов.

Так как на первом этапе количество и вид определяемых типов не известны, а сами типы будут выявляться из набора подаваемого на вход смешанного набора данных, задачу первого этапа можно рассматривать как задачу кластеризации. В качестве модели кластеризации предлагается использовать. В качестве модели кластеризации предлагается использовать метод k-means, который относится к методам машинного обучения (ML – machine learning). В отличие от методов иерархической кластеризации, он позволяет проводить обучение и дообучение на вновь поступивших в систему данных без полной перестройки всей модели, в то время как иерархические кластеризующие деревья при появлении даже одного нового вектора требуют проводить всю кластеризацию заново. Нейронные сети Кохонена являются также хорошими кандидатами на роль кластеризующего алгоритма первого уровня вложенной модели, однако они проигрывают алгоритму k-средних в быстродействии без значительного выигрыша в точности. В литературе описано, что сети Кохонена плохо справляются с кластеризацией векторов большой размерности, как, например, графические изображения [6]. Тогда как алгоритм k-means не чувствителен к размерности кластеризуемых объектов [7].

На втором этапе в распоряжении исследователей уже существуют выделенные группы – типы объектов. Поэтому для распознавания объекта внутри выделенного типа целесообразно использовать модели, хорошо зарекомендовавшие себя в задачах идентификации и классификации. Такими моделями, безусловно, являются сверточные нейронные сети (Convolutional Neural Network – CNN), также, как и алгоритм k-средних, относящиеся к группе методов ML. Несмотря на то, что подавляющее число областей их применения так или иначе касается распознавания графических изображений [8], сети данной топологии широко применяются и в других областях. Например, модели машинного обучения CNN используются для распознавания аудиосигналов [9], состояния объектов окружающей среды [10], и даже прогнозирования успешности шахматных позиций [11]. Подобная универсальность позволяет рекомен-

довать данную модель для второго уровня разрабатываемой системы при условии изначальной неопределенности природы входных векторов для идентификации.

Таким образом, модель глубокого машинного обучения для распознавания разнородных объектов можно представить как вложенную двухуровневую систему следующего вида (рис. 1).

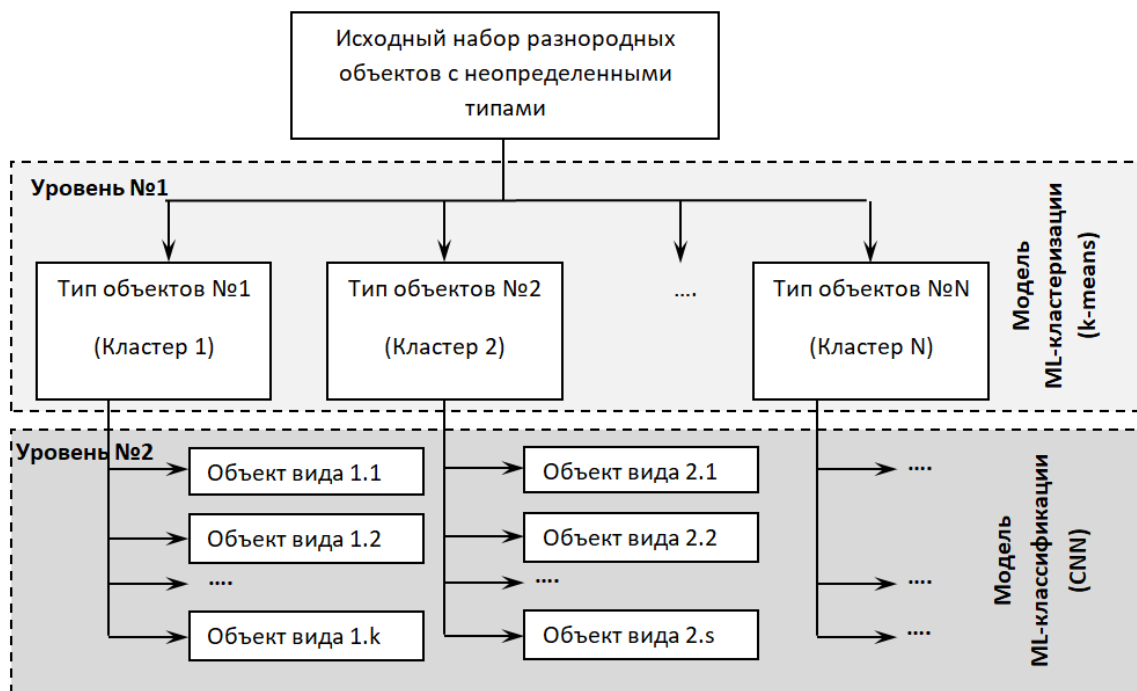


Рис. 1. Структура двухуровневой вложенной модели глубокого машинного обучения для распознавания объектов разнородных типов.

В соответствии с теорией машинного обучения, модель проходит два этапа жизненного цикла – обучение и использование.

На этапе обучения следует выполнить следующие шаги:

Для Уровня №1:

Шаг 1.1. Подготовка исходных данных для обучения: предобработка и масштабирование для приведения к единой размерности.

Шаг 1.2. Добавление к исходным данным для обучения меток типов данных, если в исходном наборе данных таковые отсутствуют.

Шаг 1.3. Соккрытие меток типов данных для исключения их влияния на процесс обучения – кластеризации исходного набора данных.

Шаг 1.4. Самообучение модели глубокого машинного обучения первого уровня (Этапа №1) по алгоритму k-means на имеющемся обучающем наборе данных (без разметки типов).

Шаг 1.5. Добавление скрытых ранее меток типов данных к данным каждого выделенного кластера.

Шаг 1.6. Анализ выделенных в результате кластеризации групп (кластеров). Определение меток классов – типов групп данных на основе меток типов данных. Кластеру присваивается тип данных, метка которого преобладает в присутствующих в кластере данных.

Для Уровня №2:

Шаг 2.1. Для каждого выделенного типа данных построить собственную модель глубокого машинного обучения типа CNN.

Шаг 2.2. Разметить данные каждого кластера (каждого типа), приписав каждому набору метку конкретного вида объекта.

Шаг 2.3. Обучить каждую модель CNN на данных соответствующего типа с предоставленными метками видов объектов внутри типа.

Схематично разработанный алгоритм может быть представлен в виде (рис. 2):

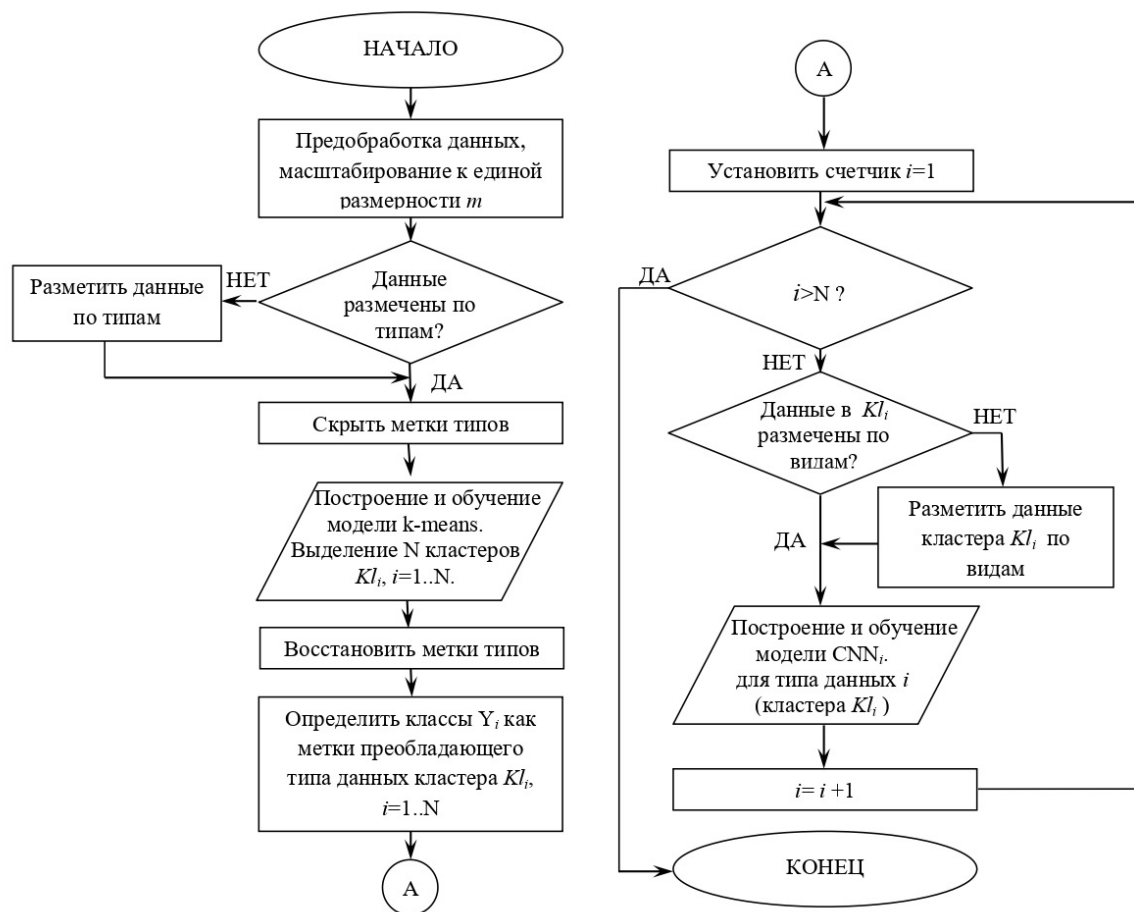


Рис. 2. Структурная схема построения и обучения двухуровневой вложенной модели идентификации разнотипных объектов.

На этапе использования вложенной модели распознавания разнородных объектов:

Шаг 1. Получить входной вектор данных – параметров для идентификации. При необходимости привести его к размерности, предусмотренной в модели, методами предобработки и масштабирования.

Шаг 2. Подать данные на вход обученной модели k-means первого уровня. Проанализировать кластер, в который его распределила модель k-means - определить тип объекта.

Шаг 3. Подать данные на вход модели CNN, обученной для определенного на предыдущем шаге типа объекта.

Шаг 4. Определить вид объекта согласно выходу нейронной сети CNN.

Разработанный алгоритм схематично может быть представлен в виде рис. 3.

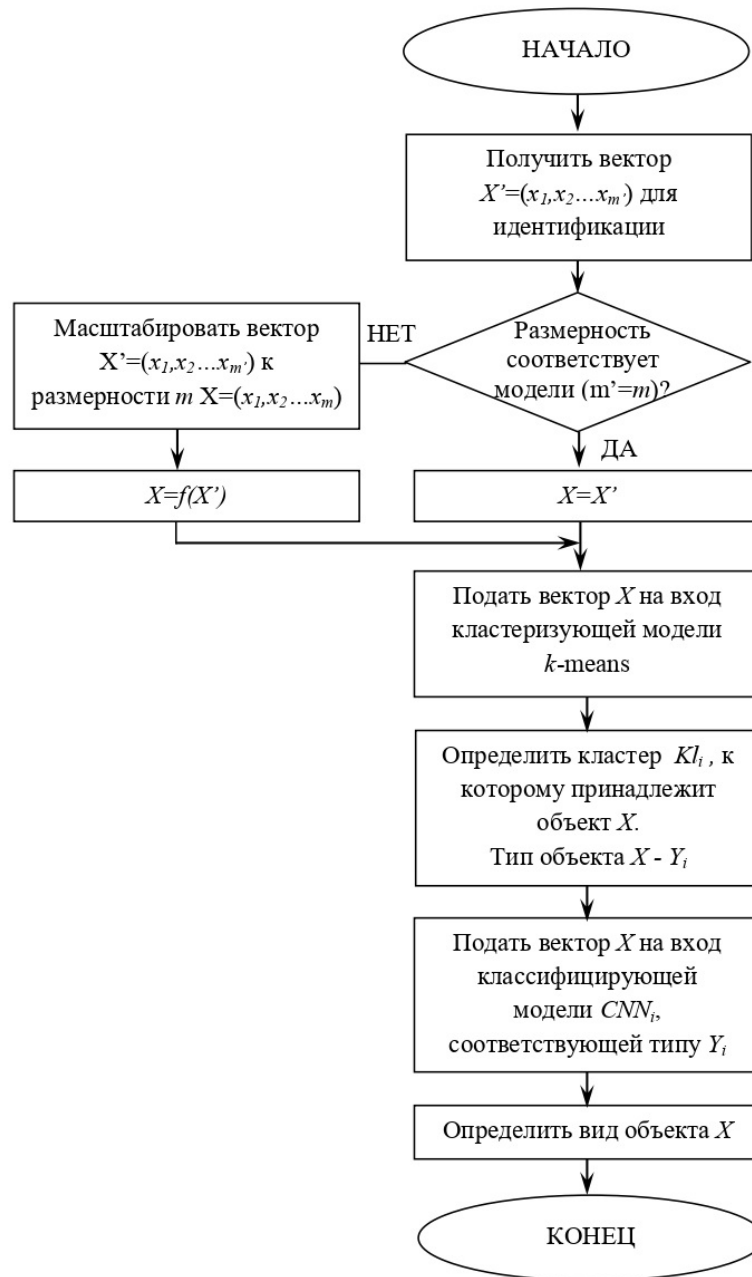


Рис. 3. Структурная схема использования двухуровневой вложенной модели идентификации разнотипных объектов.

4. Практическая реализация разработанной модели

В качестве практически используемых алгоритмов реализации модели для этапа кластеризации разработан вариант алгоритма k-means структуры, представленной на рис. 4.

В качестве алгоритма классификации отобрана нейронная сеть CNN парадигмы AlexNet. Для увеличения точности распознавания на этапе построения модели в стан-

дартный алгоритм были внесены модификации путем добавления слоёв нормализации BatchNormalization. Итоговая структура модифицированной модели приведена на рис. 5.

В процессе разработки второго уровня вложенной интеллектуальной модели сеть данной конфигурации будет строиться для каждого выделенного на первом уровне кластера. При этом параметры сети будут подбираться индивидуально на основании свойств данных, образующих кластер (тип). В процессе функционирования сеть будет определять вид объекта для выделенного типа.

В качестве языка программирования на основании проведенного исследования сравнительных характеристик и анализа альтернатив был выбран язык Python.

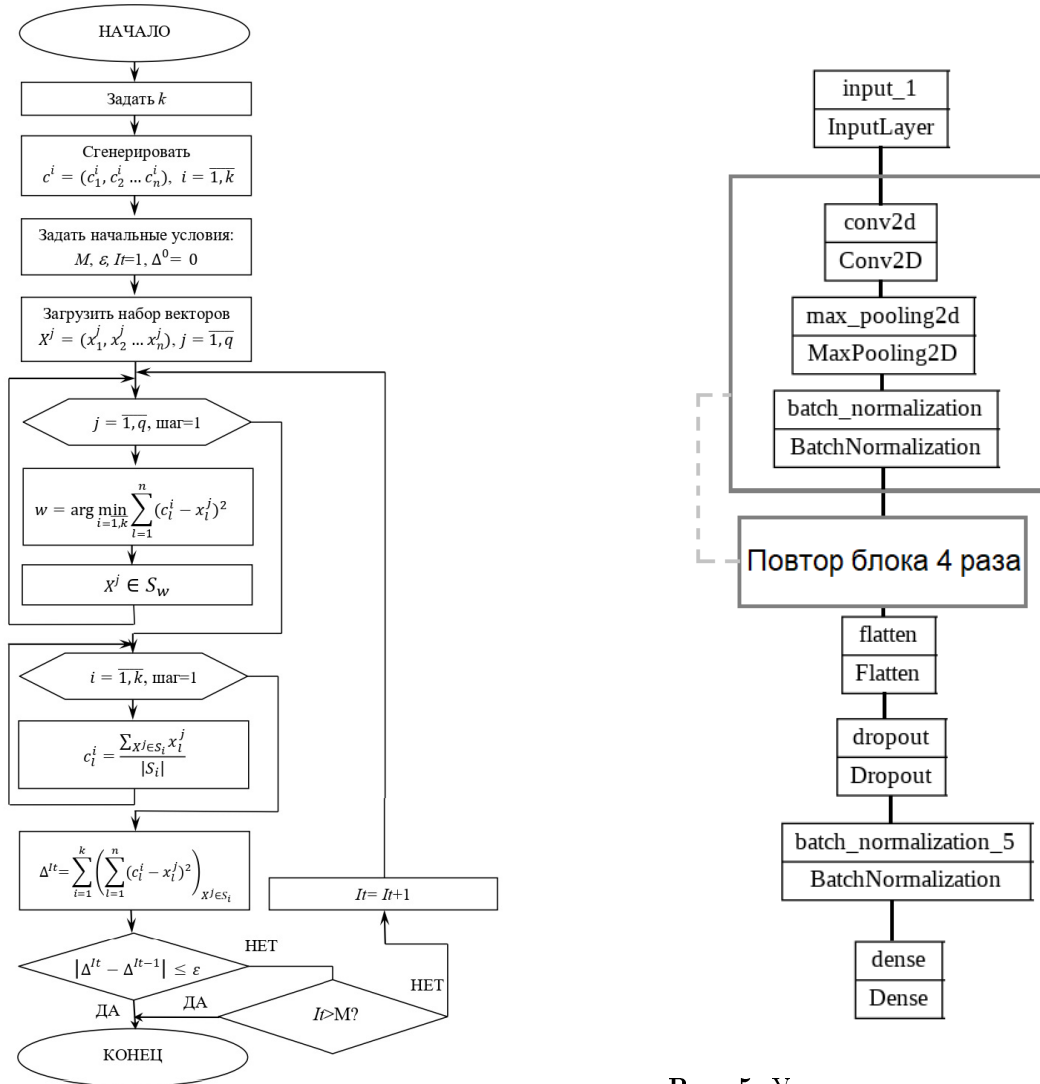


Рис. 4. Структурная схема алгоритма кластеризации на основе метода k-средних.

Рис. 5. Упрощенная структурная схема модели классификации.

5. Вычислительный эксперимент

Для вычислительных экспериментов использовался смешанный набор данных изображений объектов различной природы: средств наземного транспорта (автомобили различных марок, мотоциклы различных видов и т. п.), воздушного транспорта (самолеты различных моделей), флоры (цветы различных сортов) и фауны (собаки, кошки и птицы различных пород), фотографий людей и продуктов питания. В ходе экспериментов модель должна на первом этапе выявить *тип* объекта на изображении, а на втором - его конкретный *вид*. Например, на этапе кластеризации модель определяет изображение как фотографию птицы, а на втором, что это соловей.

5.1. Эксперименты по определению типов данных (кластеризация)

Для проведения эксперимента в программный модуль алгоритма процесса кластеризации было загружено 500 изображений смешанного исходного датасета с объектами 5 различных типов по 100 изображений на каждый тип (метки типов скрыты): «Люди», «Птицы», «Цветы», «Кошки», «Собаки».

В результате работы алгоритма в течение 6 итераций точность выявления типов составила 99% (ошибочно отнесены не к своим типам 4 изображения из 500). Наилучший результат кластеризации алгоритм показал для группы изображений типа «Люди» (рис. 6). Все 100 изображений попали в один кластер, т. е. были отнесены к данному типу. Для типа изображений «Птицы» в кластер попало 99 изображений из 100, недостающее изображение птицы было ошибочно отнесено к типу «Цветы» (рис. 7-8). Для изображений типа «Кошка» и «Собака» 3 изображения типа «Кошка» ошибочно отнесены в кластер «Собака» (рис. 9).

В целом работу алгоритма следует признать удовлетворительной.



Рис. 6. Результат определения типа «Люди».

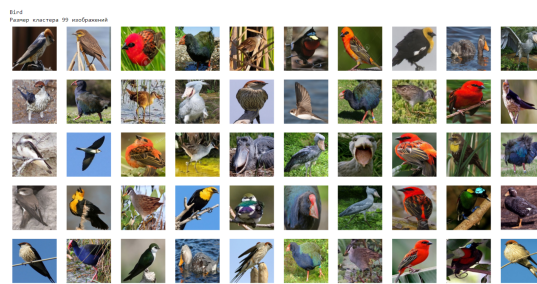


Рис. 7. Результат определения типа «Птицы»: потеряно 1 изображение.

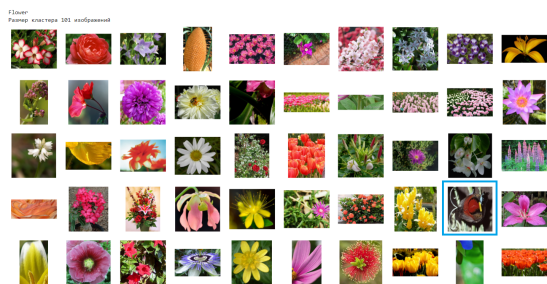


Рис. 8. Результат определения типа «Цветы»: 1 лишнее изображение (в рамке).

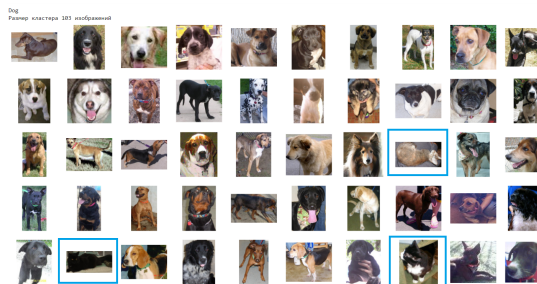


Рис. 9. Результат определения типа «Собака»: 3 лишних изображения (в рамке).

Литература

1. Andreyanov N.V., Sytnik A.S., Shleymovich M.P. Object detection in images using deep neural networks for agricultural machinery. В сборнике: IOP Conference Series: Earth and Environmental Science. Сер. "International Science and Technology Conference "Earth Science ISTC EarthScience 2022 - Chapter 2." 2022. С. 032002.
2. Савченко О.Ф., Добролюбов И.П. Моделирование процесса идентификации состояния тракторных двигателей. Проблемы вычислительной и прикладной математики. 2016. 4(6). С. 4-12.
3. Емалетдинова Л.Ю., Катасёв А.С. Нечетко-продукционная каскадная модель диагностики состояния сложного объекта // Программные системы и вычислительные методы. 2013. №1. С. 69-81.
4. Новикова С.В., Тунакова Ю.А., Шагидуллин А.Р., Кремлева Э.Ш., Валиев В.С., Габдрахманова Г.Н., Кузнецова О.Н. Использование нейросетевых технологий для зонирования территории на примере г. Казани // Вестник Технологического университета. 2019. Т. 22, № 5. С. 128-131.
5. Машинное обучение с поддержкой масштабирования: техническая документация Microsoft. [Электронный ресурс]. 03.04.2022. URL: <https://docs.microsoft.com/ru-ru/azure/architecture/data-guide/big-data/machine-learning-at-scale> (дата обращения 16.06.2022)
6. Сосулин Ю. Г., Фам Чунг Зунг Нейросетевое распознавание двумерных изображений // Радиотехника и электроника. 2003. Т. 48, № 8. С. 969-978.
7. Котелина Н.О., Матвийчук Б.Р. Кластеризация изображения методом k-средних // Вестник Сыктывкарского университета. Серия 1. Математика. Механика. Информатика. 2019. №32.
8. Голубинский А.Н., Толстых А.А. Распознавание объектов на телевизионных изображениях с использованием аппарата сверточных нейронных сетей // Вестник ВИ МВД России. 2017. №1.
9. Марьяна В.В. Методика выбора нейросетевой структуры для решения задач классификации аудиосигналов. В сборнике: Наука в движении: от отражения к созданию реальности // Материалы III Всероссийской научно-практической конференции студентов и учащихся с международным участием. Под общей редакцией С.В. Юдиной. Москва. 2020. С. 180-182.
10. Тарасов А.С., Никифоров М.Б., Бакамбис Н.И. Применение сверточных сегментационных нейронных сетей для экологического мониторинга земной поверхности // Известия Тульского государственного университета. Технические науки. 2021. № 6. С. 3-10. DOI 10.24412/2071-6168-2021-6-3-10.
11. Сальманов И.Р. Анализ и синтез методов оценки шахматных позиций в системах компьютерного моделирования партий на базе сверточных нейронных сетей Образование в России и актуальные вопросы современной науки // Сборник статей V Всероссийской научно-практической конференции. Министерство науки и высшего образования РФ. Пензенский государственный

университет и [др]. Под ред. Гагаева П.А., Белозерцева Е.П. Пенза: Пензен. гос. аграр. ун-т. 2022. С. 391-394

MSC 97R40

Development of a heterogeneous object recognition model based on nested machine learning methods

K. N. Novikova, S. V. Novikova

Kazan National Research Technical University named after A. N. Tupolev

Abstract: The article proposes a method for identifying objects in the case when their type (nature, origin) is not known in advance. In existing systems of identification, or recognition, as a rule, the origin, or general type, of objects is known in advance. For instance, encrypted digital cardiograms are fed into the system, and the system must detect pathology. Either fingerprint images are fed into the system, and the task of the model is to determine their owners. The situation when the type of the identified object is not known is practically never encountered. This paper proposes a two-level intellectual model that allows at the first stage of its operation to identify the type of object (for example, whether the object is an image, an encrypted cardiogram, telemetry readings, etc.), and at the second stage to determine its type directly (if the object is a cardiogram, it is necessary to determine pathology). The first level includes a clustering model based on the k-means algorithm. The second level forms an ensemble of CNN type models. The structure of convolutional networks is determined individually for each cluster (object type) selected at the previous level.

Keywords: heterogeneous objects, object identification, object clustering, machine learning models, nested models.

References

1. N.V. Andreyanov, A.S. Sytnik, M.P. Shleymovich, Object detection in images using deep neural networks for agricultural machinery, *IOP Conference Series: Earth and Environmental Science*, "International Science and Technology Conference "Earth Science", *ISTC EarthScience 2022, Chapter 2*", 2022, 032002.
2. O.F. Savchenko, I.P. Dobrolyubov, Modelirovanie processa identifikatsii sostoyaniya traktornykh dvigatelej, *Problemy vychislitel'noj i prikladnoj matematiki*, 2016, 4(6), P. 4-12.
3. L.YU. Emaletdinova, A.S. Katasyov, Nechetko-produkcionnaya kaskadnaya model' diagnostiki sostoyaniya slozhnogo ob'ekta, *Programmnye sistemy i vychislitel'nye metody*, 2013, №1, P. 69-81.
4. S.V. Novikova, YU.A. Tunakova, A.R. SHagidullin, E.SH. Kremleva, V.S. Valiev, G.N. Gabdrahmanova, O.N. Kuznecova, Ispol'zovanie nejrosetevykh tekhnologij dlya zonirovaniya territorii na primere g. Kazani, *Vestnik Tekhnologicheskogo universiteta*, 2019, v. 22, № 5, P. 128-131.
5. Mashinnoe obuchenie s podderzhkoj masshtabirovaniya: tekhnicheskaya dokumentatsiya Microsoft. [Electronic resource]. 03.04.2022. URL:

- <https://docs.microsoft.com/ru-ru/azure/architecture/data-guide/big-data/machine-learning-at-scale> (date of the application: 16.06.2022)
6. YU.G. Sosulin , Fam CHung Zung, Nejrosetevoe raspoznavanie dvumernyh izobrazhenij, *Radiotekhnika i elektronika*, 2003, v. 48, № 8, P. 969-978.
 7. N.O. Kotelina, B.R. Matvijchuk, Klasterizaciya izobrazheniya metodom k-srednih, *Vestnik Syktyvkarского университета. Seriya 1. Matematika. Mekhanika. Informatika*, 2019, №32.
 8. A.N. Golubinskij, A.A. Tolstyh, Raspoznavanie ob'ektov na televizionnyh izobrazheniyah s ispol'zovaniem apparata svertochnyh nejronnyh setej *Vestnik VI MVD Rossii*, 2017, №1.
 9. V.V. Mar'ina, Metodika vybora nejrosetevoj struktury dlya resheniya zadach klassifikacii audiosignalov, *Nauka v dvizhenii: ot otrazheniya k sozdaniyu real'nosti. materialy III Vserossijskoj nauchno-prakticheskoy konferencii studentov i uchashchihsya s mezhdunarodnym uchastiem. Pod obshchej redakciej S.V. YUdinoj. Moskva*, 2020, P. 180-182.
 10. A.S. Tarasov, M.B. Nikiforov, N.I. Bakambis, Primenenie svertochnyh segmentacionnyh nejronnyh setej dlya ekologicheskogo monitoringa zemnoj poverhnosti, *Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki*, 2021, № 6, P. 3-10. DOI 10.24412/2071-6168-2021-6-3-10.
 11. I.R. Sal'manov, Analiz i sintez metodov ocenki shahmatnyh pozicij v sistemah komp'yuternogo modelirovaniya partij na baze svetochnyh nejronnyh setej, *Obrazovanie v Rossii i aktual'nye voprosy sovremennoj nauki: sbornik statej V Vserossijskoj nauchno-prakticheskoy konferencii. Ministerstvo nauki i vysshego obrazovaniya RF, Penzenskij gosudarstvennyj universitet i [dr]. Pod red. Gagaeva P.A., Belozerceva E.P.*, Penza, Penzen. gos. agrar. un-t, 2022, P. 391-394