

ИСПОЛЬЗОВАНИЕ СОВРЕМЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ ТЕХНОЛОГИЙ И МАТЕМАТИЧЕСКИХ МЕТОДОВ ОБРАБОТКИ СПУТНИКОВЫХ ДАННЫХ В ДАЛЬНЕВОСТОЧНОМ ЦЕНТРЕ «НИЦ «ПЛАНЕТА»

Кучма М.О. – м.н.с отдела разработки и внедрения НИОКР (ДЦ ФГБУ «НИЦ «Планета»), e-mail: m.kuchma@dvrspod.ru;

Крамарева Л.С. – директор, (ДЦ ФГБУ «НИЦ «Планета»), e-mail: kramareva@dvrspod.ru;

Филей А.А. – н.с отдела выходной продукции (ДЦ ФГБУ «НИЦ «Планета»), e-mail: andreym-61@mail.ru;

Амельченко Ю.А. – м.н.с отдела разработки и внедрения НИОКР (ДЦ ФГБУ «НИЦ «Планета»), e-mail: juliapkhv@gmail.com

В работе представлен опыт применения методик искусственного интеллекта, в частности нейронных сетей, для решения различных актуальных задач гидрометеорологии. Приводятся результаты исследований, проведенных в Научно-исследовательском центре космической гидрометеорологии «Планета».

This paper presents an experience of using artificial intelligence techniques, in particular neural networks, to solve various urgent problems of hydrometeorology. The results of studies conducted at the Scientific-research center for space hydrometeorology «Planeta».

Ключевые слова: графический ускоритель, распределенные вычисления, физическое моделирование, детектирование облачности, детектирование снега, типы облачности, профиль температуры, профиль влажности, нейронная сеть.

Введение. Разработка и оптимизация алгоритмов анализа и интерпретации данных, полученных в результате спутниковых наблюдений в гидрометеорологических целях, обеспечили значительное увеличение потенциала использования космической информации в процессе ежедневного анализа и прогнозирования погодных условий. Источником данного увеличения является фактор возможности постоянного глобального мониторинга, предоставляемый спутниковыми инструментами дистанционного зондирования Земли (ДЗЗ), а также их способностью к регулярному циклу наблюдений над поверхностью Земли и состоянием атмосферы. С учетом увеличения группировок гидрометеорологических космических аппаратов, эти данные позволяют более эффективно решать ряд прикладных научных проектов и задач оперативной метеорологии, включая прогнозирование различных атмосферных явлений и мониторинг климатических изменений. Но вместе с этим происходит и увеличение числа ограничений, обусловленных необходимостью обработки большого объема разнородных данных, которые не всегда обладают логически связанными характеристиками.

Физическое моделирование для спектрального анализа спутниковых данных, а также методы машинного обучения и нейронных сетей, требуют больших вычислительных ресурсов, поэтому на первый план выходит проблема оптимизации вычислений.

Аппаратные средства. Решение некоторых задач, таких как проведение физического моделирования параметров атмосферы, позволяющее учесть различную геометрию спутниковых наблюдений и влияние подстилающей поверхности, требует больших вычислительных ресурсов. Для уменьшения времени моделирования применяются технологии распределенного вычисления. В Дальневосточном центре НИЦ «Планета» решение этой задачи осуществляется с использованием 19 виртуальных машин на основе нескольких серверов с центральными процессорами Intel Xeon Gold 6130, Intel Xeon Silver 4114, Intel Xeon CPU E5-2660 v4 и Intel Xeon CPU E7-4820 v2, которые в совокупности имели 539 логических ядер. В результате такого подхода, генерирование справочной таблицы с

использованием модели переноса излучения 6S в размере 278419680 элементов составило 40 дней. Для сравнения, при использовании рабочей станции с центральным процессором Intel Core i7-6700K, имеющего 8 логических ядер, это заняло бы 7 лет.

Для автоматизации обработки больших потоков данных активно развивается тема применения в задаче обработки и интерпретации спутниковых данных методов искусственного интеллекта, таких как машинное обучение и нейронные сети. Для уменьшения времени обучения нейросетевых алгоритмов в Дальневосточном центре НИЦ «Планета» установлен и настроен высокопроизводительный сервер, предназначенный для ускорения построения спутниковой информационной продукции. Графические ускорители, установленные в сервере, позволяют выполнять сложные вычисления, производить быструю обработку больших массивов данных и хорошо подходят для обучения нейронных сетей и обработки данных уже обученными сетями.

Сервер позволяет одновременно запускать несколько ресурсоёмких задач по обучению нейронных сетей. Высокую производительность вычислений обеспечивают:

- два 10-ядерных центральных процессора Intel Xeon Silver 4210;
- 10-Гигабитные сетевые интерфейсы для быстрой передачи больших объёмов данных во внешнее хранилище;
- подсистема памяти, включающая 256 гигабайт ОЗУ, позволяющая производить операции с большими выборками данных без необходимости использования буфера на локальных дисках, 2 высокоскоростных твердотельных NVME-накопителя, а также 4 диска большого объёма для локального хранения данных;
- 8 высокопроизводительных графических ускорителей Nvidia RTX A5000 с тензорными ядрами 3 поколения, объединённых высокоскоростными интерфейсами NVLINK, благодаря которым обмен данными между видеокартами осуществляется напрямую с минимальными задержками (рис. 1).

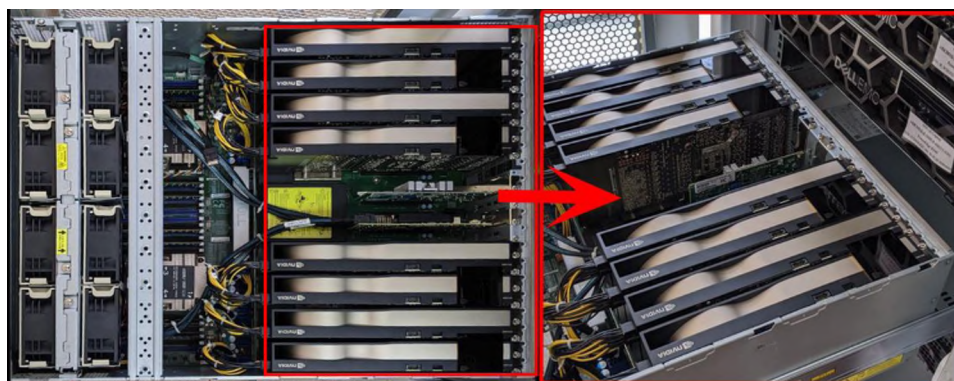


Рис. 1 Вычислительный сервер и кластер графических ускорителей Nvidia RTX A5000.

Использование графических ускорителей позволяет на порядок повысить скорость обработки в сравнении с обработкой на центральном процессоре и в несколько раз в сравнении с обработкой на рабочей станции с одним графическим ускорителем. Например, для сверточной нейронной сети, содержащей 4 млн. параметров, время полного обучения с использованием графического ускорителя Nvidia RTX 2080 на протяжении 300 итераций составляет приблизительно 20 дней. Использование сервера из 8 графических ускорителей Nvidia RTX A5000 дает возможность сократить время обучения до 3-х дней, т.е. процессы ускоряются, по меньшей мере, в 6 раз.

Для решения научных и исследовательских задач, требующих больших вычислительных ресурсов и обработки больших объемов архивных данных, специалисты центра используют ресурсы Центра коллективного пользования (ЦКП) «ИКИ-Мониторинг» [1] и ЦКП «Центр данных ДВО РАН» [2].

Методы обработки спутниковых данных. Использование представленных аппаратных средств и вычислительных технологий позволяет производить расчеты и использовать большое количество зависимых и независимых переменных для задач, требующих оперативного получения конечного результата.

Например, для высокоэллиптического космического аппарата (КА) Арктика-М № 1 на основе применения сверточной нейронной сети с вейвлет-преобразованиями был разработан алгоритм минимизации помех в виде полос на данных инфракрасных каналов спутникового прибора [3]. Внедрение алгоритма в оперативную работу позволило без ограничений получать весь перечень информационной продукции, предусмотренной техническим заданием.

Для получения тематической продукции разработаны и внедрены алгоритмы на основе проведения физического моделирования для расчета высоты верхней границы облачности, эффективного радиуса облачных частиц, оптической толщины в спектральном канале 0.6 мкм и фазового состава воды в облаке [4]. Также, на основе сверточных нейронных сетей разработаны алгоритмы для детектирования облачного [5], снежного [6] и ледяного покровов [7]. Они реализованы для приборов МСУ-МР, МСУ-ГС и АНІ, установленных на космических аппаратах (КА) серий Метеор-М, Арктика-М и Электро-Л, Himawari соответственно. Одной из актуальных разработок является расчет вертикальных профилей температуры и влажности в атмосфере [8] по измерениям прибора МТВЗА-ГЯ, установленного на КА серии Метеор-М. Особенностью разработки является то, что для обучения сети прямого распространения на основе многослойного персептрона используется обучающая выборка, которая содержит в себе практически все вариации атмосферных параметров за несколько лет. Дополнительно, по данным прибора АНІ рассчитываются типы облачности [9] и интенсивность осадков для территории Азиатско-Тихоокеанского региона с применением методов искусственного интеллекта. Эти параметры являются важным дополнением к аэросиноптическому материалу, особенно над территориями, где отсутствуют наземные наблюдения.

Заключение. Подходы к решению задач гидрометеорологии с использованием физического моделирования и нейросетевых алгоритмов зачастую требуют больших вычислительных ресурсов и учета большого количества различных зависимых и независимых величин. Использование высокопроизводительного сервера с графическими ускорителями и использование виртуальных машин для распределенного вычисления позволяет сократить время на разработку алгоритмов.

Список литературы:

1. Лупян Е. А., Прошин А. А., Бурцев М. А., Кашницкий А. В., Балашов И. В., Барталев С. А., Константинова А. М., Кобец Д. А., Мазуров А. А., Марченков В. В., Матвеев А. М., Радченко М. В., Сычугов И. Г., Толпин В. А., Уваров И. А. Опыт эксплуатации и развития центра коллективного пользования системами архивации, обработки и анализа спутниковых данных (ЦКП «ИКИ-Мониторинг») // Современные проблемы дистанционного зондирования Земли из космоса. 2019. Т. 16. № 3. С. 151-170. DOI: 10.21046/2070-7401-2019-16-3-151-170
2. Sorokin A.A., Makogonov S.I., Korolev S.P. The Information Infrastructure for Collective Scientific Work in the Far East of Russia // Scientific and Technical Information Processing Vol. 44, No. 4. P. 302-304. DOI: 10.3103/S0147688217040153.
3. Кучма М. О., Андреев А. И., Крамарева Л. С. Нейросетевой алгоритм минимизации помех в инфракрасных каналах прибора МСУ-ГС космического аппарата Арктика-М №1 // Материалы 20-й Всероссийской открытой конференции «Современные проблемы дистанционного зондирования Земли из космоса». Москва: ИКИ РАН. 2022. с. 45. DOI: 10.21046/20DZZconf-2022a
4. Филей А. А. Восстановление оптической толщины и эффективного радиуса частиц облачности по данным дневных измерений спутникового радиометра МСУ-МР // Оптика атмосферы и океана. 2019. Т. 32, № 8. с. 650-656. DOI: 10.15372/AOO20190807
5. Андреев А. И., Шамилова Ю. А. Детектирование облачности по данным КА Himawari-8 с применением сверточной нейронной сети // Исследование Земли из космоса. 2021. №2. с. 42-52. DOI: 10.31857/S0205961421010036
6. Bloschchinskiy V. D., Kuchma M. O., Andreev A. I., Sorokin A. A. Snow and cloud detection using a convolutional neural network and low-resolution data from the Electro-L No. 2 Satellite // J. Appl. Rem. Sens. 2020. Vol. 14, No. 3. pp. 034506. DOI: 10.1117/1.JRS.14.034506

7. Кучма М. О., Лотарева З. Н., Слесаренко Л. А. Определение маски ледяного покрова дальневосточных морей по данным прибора МСУ-МР спутника «Метеор-М» № 2 // Исследование Земли из космоса. 2021. № 2. с. 31-41. DOI: 10.31857/S0205961421020032

8. Филей А. А., Андреев А. И., Успенский А. Б. Использование искусственных нейронных сетей для восстановления температурно-влажностного состояния атмосферы по данным спутникового микроволнового радиометра МТВЗА-ГЯ КА Метеор-М № 2-2 // Исследование Земли из космоса. 2021. № 6. с. 83-95. DOI: 10.31857/S0205961421060087

9. Андреев А. И., Давиденко А. Н., Пустатинцев И. Н. Алгоритм классификации типов облачности на основе нейронной сети и метода опорных векторов // Материалы 17-й Всероссийской открытой конференции «Современные проблемы дистанционного зондирования Земли из космоса». Москва: ИКИ РАН. 2019. с. 14. DOI: 10.21046/17DZZconf-2019a