

Интеллектуальная технология распознавания подстилающей поверхности Земли

СЕРГЕЙ КРУГЛИКОВ,
заместитель генерального
директора Объединенного
института проблем информатики
Национальной академии наук
Республики Беларусь (ОИПИ НАН
Беларуси) по научной
и инновационной
работе, доктор военных
наук, доцент, член-
корреспондент Академии
военных наук
Российской Федерации

ВАСИЛИЙ КОВАЛЕВ,
заведующий
Лабораторией анализа
биомедицинских
изображений ОИПИ
НАН Беларуси, кандидат
технических наук

ДМИТРИЙ ПАВЛЕНКО,
инженер-программист
ОИПИ НАН Беларуси

ЭДУАРД СНЕЖКО,
ведущий научный
сотрудник ОИПИ НАН
Беларуси, кандидат
технических наук

ВИТАЛИЙ ЛЕВЧУК,
научный сотрудник
ОИПИ НАН Беларуси

Автоматический поиск и распознавание объектов и подстилающей поверхности (ПП) определенного вида, автоматически осуществляемые по видео- и фотоизображениям на борту малых космических и беспилотных летательных аппаратов (БЛА) в полете, входят в перечень важных задач, решение которых обеспечивает целесообразность их применения. Для ведения поиска и распознавания в полетные задания БЛА включаются и используются образы ПП – специально сформированные изображения участков местности с характерными типовыми и надежно определяемыми признаками, наблюдаемыми на определенных курсах полета и углах места наблюдения.

Автоматическое распознавание местности с использованием образов ПП исключает длительный и утомительный просмотр видеоматериалов операторами наземных пунктов управления и возможные при этом ошибки выбора участков/объектов интереса, повышая достоверность выбора, сокращает время подготовки необходимых данных, уменьшает потребление энергии от бортовых источников.

В ходе разработки микро модуля автоматического предварительного распознавания и отбора изображений подстилающих поверхностей, способного к функциональному взаимодействию с бортовыми системами управления БЛА, в Объединенном институте проблем информатики Национальной академии наук (ОИПИ НАН) Беларуси проводится комплекс вычислительных экспериментов по проверке программно-аппаратного обеспечения коммерчески доступных вычислительных микроплатформ.

Основные функции микро модуля и ключевые технические решения:

1. Микро модуль является автономной системой автоматического распознавания изображений подстилающих поверхностей различных типов: леса различных видов, промышленные сооружения, водоемы, дороги, поля, кустарники, дачная застройка, сельхозугодия различных видов с растительностью, находящейся на различных стадиях роста и так далее.

2. Конкретные целевые классы задаются в полетном задании в виде типовых образцов соответствующих поверхностей для распознавания по традиционной схеме «изображение-признаки-классификатор» или с помощью сверточной нейронной сети, подготовленной по технологии глубокого обучения (Deep Learning).

3. Встроенная видекамера микро модуля задействуется только для получения изображения поверхности. Кадры видеопоследовательности – изображения низкого разрешения целевых классов полетного задания могут сохраняться во внешней памяти микро модуля для целей последующего камерального анализа.

4. Объем взаимодействия и обмена информацией микро модуля с основным бортовым компьютером минимизируется и сводится к передаче в реальном времени выявленного номера класса текущего изображения подстилающей поверхности и значения вероятности его достоверного определения. Класс «0» означает, что текущее изображение не относится ни к одному из целевых классов, заданных в полетном задании. Решение о выполнении высококачественной съемки, записи и передачи изображения на пункт управления вырабатывается бортовым компьютером или на наземном пункте управления. Соответствующая команда передается на видекамеру БЛА.

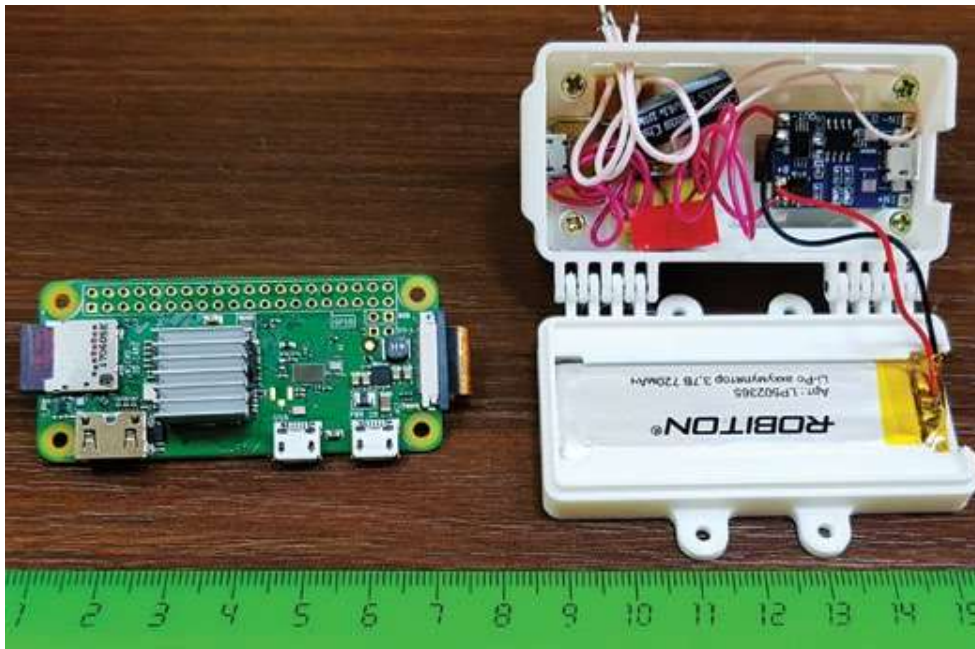


Рис. 1. Комплектация макета микромодуля

5. Компьютер микромодуля должен быть вычислительной системой со всем необходимым: операционной системой, средствами разработки программ, интерфейсным оборудованием. Схемотехнические решения по комплексированию микромодуля из отдельных специализированных разнородных цифровых сигнальных процессоров (DSP), микроконтроллеров, чипсетов, систем на кристалле и других компонентов не разрабатываются и не рассматриваются.

Эволюционно электронно-вычислительные системы стали многократно мощнее, дешевле, доступнее и уменьшились в размерах. Образовался рынок одноплатных микрокомпьютеров для решения разнообразных задач. Современные микрокомпьютеры и персональные компьютеры имеют совместимые операционные системы и специальное программное обеспечение. Полетные задания, образы местности и объектов можно разрабатывать на автоматизированных рабочих местах уже в требуемых для микрокомпьютера форматах и переносить на микроплатформу. Упрощается и ускоряется работа, обеспечивается требуемое качество полетного задания и последующей обработки информации, накопленной в полете.

При изготовлении макета микромодуля спроектирован и отпечатан на 3D-принтере корпус микромодуля (рис. 1 и 2), в котором раз-

мещаются: микрокомпьютер Raspberry Pi Zero Wireless с камерой модели Raspberry Pi Zero V1.3 Mini Camera (в сумме 13 граммов), Li-Po аккумулятор типа ROBITON емкостью 720 мАч (14 граммов), плата зарядки АКБ, сигнальные светодиоды, соединительные провода.

Размеры макета микромодуля 7,3×3,6×2,5 см, масса в собранном состоянии – 57 граммов (при ограничении 300 граммов), включая автономный источник питания микромодуля, сред-

Рис. 2. Макет микромодуля

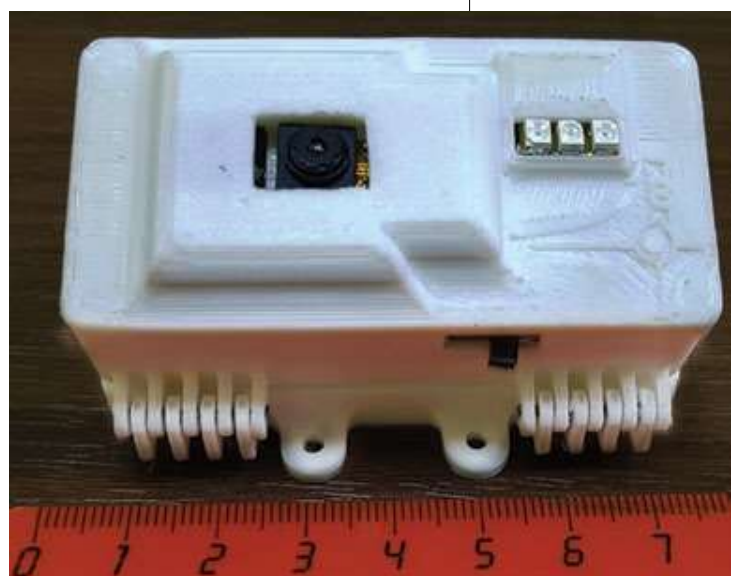




Рис. 3. Выборка изображений подстилающей поверхности различных типов

ства его зарядки, переключатель на внешний источник питания от сети в горячем режиме, что весьма необходимо и удобно при разработке и отладке программного обеспечения. Это соответствует конструктивным характеристикам бортовых систем управления, начиная с легких бюджетных БЛА, и создает конкурентные преимущества.

В лабораторных испытаниях макета микро модуля с квадрокоптером Parrot AR.Drone 2.0 распознанный класс текущего изображения в реальном времени выводился на цветной светодиодный экран непосредственно на корпусе. Копия текущего кадра изображения и результаты его распознавания передавались на ноутбук по Wi-Fi для наблюдения, контроля и отладки программного обеспечения.

Программное обеспечение микро модуля состоит из наземной и полетной частей. Основная задача наземной части – выбор для формируемого полетного задания образцов изображений целевых классов для традиционной схемы распознавания либо загрузка в микро модуль предобученной нейронной сети.

Полетная часть программного обеспечения осуществляет анализ текущих кадров видеопоследовательности, оперативное распознавание

изображений ПП и отнесение их к одному из указанных классов.

НАБОР ДАННЫХ ДЛЯ ТЕСТИРОВАНИЯ НЕЙРОСЕТЕВЫХ АЛГОРИТМОВ

Для тестирования и сравнения между собой алгоритмов распознавания типа ПП, основанных на применении сверточных нейронных сетей, использован набор изображений восьми различных типов ПП размером по 256×256 пикселей: поле, лес, городская застройка, частный сектор, шоссе, проселочная дорога, озеро, река.

Изображения выделены из цветных спутниковых снимков Google Maps с уровнем увеличения Z18. Для каждого из рассматриваемых типов подстилающей поверхности выделено 5000 изображений. Из них 2500 относились к обучающей выборке и 2500 – к тестовой. Всего в тестировании задействовалось 40 000 изображений. Во избежание эффекта переобучения спутниковые снимки для обучающей и тестовой выборки выбирались из разных регионов. Обучающая выборка представлена снимками окрестностей города Минска, а данные для тестирования выбраны из окрестностей города Бреста и озера Нарочь. Выборка фотоизображений ПП различного типа из массивов, предназначенных для обучения нейросетей, – в блоке рисунков 3.

АРХИТЕКТУРА ИСПОЛЬЗУЕМОЙ НЕЙРОННОЙ СЕТИ

Для оперативного распознавания типов подстилающей поверхности на микро модуле с учетом низкой производительности его процессора разработана авторская архитектура сверточной нейронной сети. Используемая нейронная сеть подобно многим популярным глубоким сетям составлена из набора идущих друг за другом слоев. Данные с выхода одного слоя подаются на вход следующего, сигналы входных данных последовательно проходят все слои, пока не достигнут выхода нейронной сети.

Глубина нейронной сети (количество слоев) сравнительно небольшая, размер входных данных также уменьшен, чтобы обеспечивалась высокая скорость ее работы. Перед подачей на вход нейронной сети цветное RGB-изображение приводится к размеру 64×64 пикселей. Умень-

Таблица 1. Классификация изображений тестовой выборки данных при помощи разработанной нейронной сети

класс	озеро	город	поле	лес	шоссе	дорога	река	дачи
озеро	0,81	0,00	0,05	0,13	0,00	0,00	0,01	0,00
город	0,03	0,60	0,04	0,01	0,07	0,01	0,01	0,23
поле	0,04	0,00	0,95	0,00	0,00	0,01	0,00	0,00
лес	0,60	0,00	0,02	0,32	0,00	0,00	0,05	0,00
шоссе	0,00	0,01	0,04	0,01	0,73	0,11	0,10	0,01
дорога	0,00	0,00	0,01	0,00	0,05	0,91	0,04	0,00
река	0,03	0,00	0,09	0,02	0,05	0,29	0,47	0,04
дачи	0,00	0,00	0,01	0,00	0,00	0,06	0,00	0,92

шенное изображение подается на вход первого сверточного слоя.

Первая часть сети состоит из двух последовательно расположенных блоков, содержащих сверточные слои. В сверточном слое первого блока 12 ядер свертки размером 3×3 пикселей. Во втором блоке сверточный слой содержит 24 ядра. Все сверточные слои чередуются со слоями с активационной функцией ReLU: $f(x)=\max(0, x)$. За активационными слоями расположены субдискретизирующие слои, выполняющие нелинейное уплотнение карты признаков. После упомянутых выше блоков расположена полносвязная сеть, выходом которой является еще один активационный слой с функцией Softmax. Выходом сверточной сети для входного изображения является вектор размером 8 элементов, что соответствует количеству распознаваемых классов подстилающей поверхности. Каждый элемент вектора представляет собой вероятность принадлежности входного изображения к соответствующему классу. Изображение относится к некоторому классу, если указанная вероятность превышает порог 0,99.

ОЦЕНКА КАЧЕСТВА РАСПОЗНАВАНИЯ И СРАВНЕНИЕ

Для сравнения качества распознавания типов подстилающей поверхности использованы популярные архитектуры нейронных сетей: LeNet, AlexNet и GoogLeNet. У них разная глубина и различное количество обучаемых весов.

Так, сеть LeNet имеет два сверточных и два полносвязных слоя, и при стандартном для этой сети размере цветного изображения 28×28 пикселей количество обучаемых параметров сети составляет 431 078. У сети AlexNet пять сверточных слоев, а общее количество обучаемых параметров при обработке цветного изображения размером 28×28 пикселей – 56 901 000. Текущая реализация разработанной архитектуры сверточной сети соответствует 64 490 обучаемым параметрам благодаря относительно небольшому количеству узлов в полносвязных слоях.

Перед подачей на вход сети LeNet все изображения приводились к размеру 28×28 пикселей. Для предложенной сети размер входных изображений составлял 64×64 пикселей. В сетях AlexNet и GoogLeNet использовались изображения исходного размера 256×256 пикселей. Во всех случаях цветные изображения представлены в формате RGB. Все сети обучены и тестированы на одинаковых наборах изображений. Для сети LeNet в расчет брались только изображения, для которых вероятность принадлежности к одному из классов была не менее 0,5. Для AlexNet и GoogLeNet значение аналогичного порогового параметра составляло 0,75. Эти значения выбирались таким образом, чтобы количество данных, участвующих в оценке точности распознавания, для различных архитектур сети было приблизительно одинаковым.

В таблице 1 приведены относительные количества изображений, отнесенных к одному

Таблица 2. Классификация изображений тестовой выборки данных при помощи сети GoogLeNet

класс	озеро	город	поле	лес	шоссе	дорога	река	дачи
озеро	0,91	0	0,01	0,05	0,01	0	0,02	0
город	0,1	0,6	0	0	0,15	0,01	0,01	0,13
поле	0,04	0	0,94	0	0	0,01	0,01	0
лес	0,55	0	0	0,44	0	0	0,01	0
шоссе	0,01	0	0,05	0	0,73	0,07	0,13	0
дорога	0,01	0	0	0	0,01	0,95	0,03	0
река	0,01	0	0	0	0,01	0,11	0,87	0
дачи	0	0	0	0	0	0,08	0	0,91

из восьми классов при помощи разработанной нейронной сети отдельно для каждого класса изображений. Доля изображений, вероятность принадлежности к одному из классов для которых превышала пороговое значение для этой сети, составила 0,86. Средняя точность классификации составила 0,71. Видно, что некоторые типы подстилающей поверхности, такие как «озеро», «поле», «проселочная дорога», «частный сектор» и «дачи» распознаются с достаточно высокой точностью. Но изображения класса «лес» распознаются сетью достаточно плохо и зачастую путаются с изображениями класса «озеро». Это можно объяснить как некоторыми особенностями выделения изображений класса «озеро» (выделение берегов озер), так и недостаточным разнообразием изображений этого класса в выборках данных.

В таблице 2 приведены относительные количества изображений, отнесенных к одному из восьми классов при помощи сети GoogLeNet. Их доля, вероятность принадлежности к одному из классов для которых превышала пороговое значение для этой сети, в данном случае составила 0,88. Средняя по типам подстилающей поверхности точность распознавания составила 0,79.

При низкой точности распознавания определенных классов с помощью разработанной сети, отмечалось большое количество ошибок и в случае использования сети GoogLeNet. Таким образом, следует вывод, что для значительного улучшения результатов использования в полете образов ПП «проблемных» классов необходима

подготовка более массивного и качественного набора данных для обучения и тестирования нейронных сетей.

Средняя точность распознавания классов изображений на тестовой выборке данных при помощи сети AlexNet составила 0,78 при доле изображений с приемлемой вероятностью принадлежности к одному из классов, равной 0,82. Для сети LeNet аналогичные показатели соответственно 0,48 и 0,92. Таким образом, разработанная архитектура нейронной сети обеспечивает точность распознавания типов ПП, сравнимую с популярными глубокими нейронными сетями, имеющими десятки миллионов обучаемых параметров.

*

Проведенные экспериментальные исследования показывают, что разработанная в ОИПИ НАН Беларуси архитектура сверточной нейронной сети соответствует требованиям по распознаванию типов ПП на аэрокосмических снимках. Надежность распознавания типов ПП сопоставима с той, что достигается при использовании популярных архитектур глубоких сверточных сетей с множеством слоев.

Малое время обработки одного кадра дает возможность для высококачественного распознавания с применением мобильных микроплатформ с предложенной архитектурой и малой вычислительной мощностью с помощью образов ПП, подготовленных на базе качественного и представительного набора данных различных типов ПП. ◆