

УДК 681.518.5

© 2021 г. М.О. Кучма^{1,2},

В.В. Воронин¹, д-р техн. наук,

В.Д. Блощинский^{1,2}

¹(Тихоокеанский государственный университет, Хабаровск),

²(Дальневосточный центр ФГБУ "НИЦ "Планета", Хабаровск)

НАСТРОЙКА, ОБУЧЕНИЕ И ТЕСТИРОВАНИЕ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ В ЗАДАЧЕ ТЕМАТИЧЕСКОЙ ОБРАБОТКИ СПУТНИКОВЫХ ИЗОБРАЖЕНИЙ

Рассматривается алгоритм, основанный на сверточной нейронной сети, которая обнаруживает облачные образования и снежный покров на спутниковых снимках с помощью текстур. При этом в качестве входных данных используются мультиспектральные спутниковые изображения, полученные с многозонного сканирующего прибора, используемого для гидрометеорологического обеспечения и установленного на российском спутнике «*Электро-Л № 2*». Разработанный алгоритм позволяет создавать маски облачности и снежного покрова для области, ограниченной значениями зенитного угла Солнца в диапазоне от 0 до 80 градусов для дневного времени. Точность алгоритма оценивалась с помощью показателей машинного обучения. Предлагаемый алгоритм полностью автоматический и работает в любое время года в дневное время.

Ключевые слова: нейронные сети, обучение, тестирование, текстура, дистанционное зондирование, «*Электро-Л*», «*МСУ-ГС*».

DOI: 10.22250/isu.2021.67.90-100

Введение

В настоящее время в области дистанционного зондирования Земли все большее распространение получают методы текстурного анализа данных, включающие применение нейронных сетей [1, 2]. Последние исследования по данному вопросу показывают, что машинное обучение может успешно применяться для задач классификации подстилающей поверхности на спут-

никовых изображениях, а также для детектирования облачности [3, 4].

Для решения задачи [5] детектирования снежного покрова и облачных образований с помощью нейронных сетей по данным прибора МСУ-ГС космического аппарата (КА) "Электро-Л" №2 используется текстурный метод [6]. Для этого метода была создана обучающая выборка [7] по спутниковым данным в виде изображений. Данная спутниковая информация доступна в Дальневосточном центре ФГБУ «Научно-исследовательский центр космической гидрометеорологии «Планета» (далее – «НИЦ «Планета»).

Обучающая выборка (ОВ) для классификации набиралась с использованием данных КА «Электро-Л» № 2 в период с марта 2018 г. по март 2019 г., с пространственным разрешением пикселя 4 км. В рамках решения поставленной задачи специалистами вручную было отобрано около 112 тысяч текстур [8], каждая из которых относилась к одному из трех классов: облачность, снег и чистая земля/чистая вода. Пример текстур показан на рис. 1.

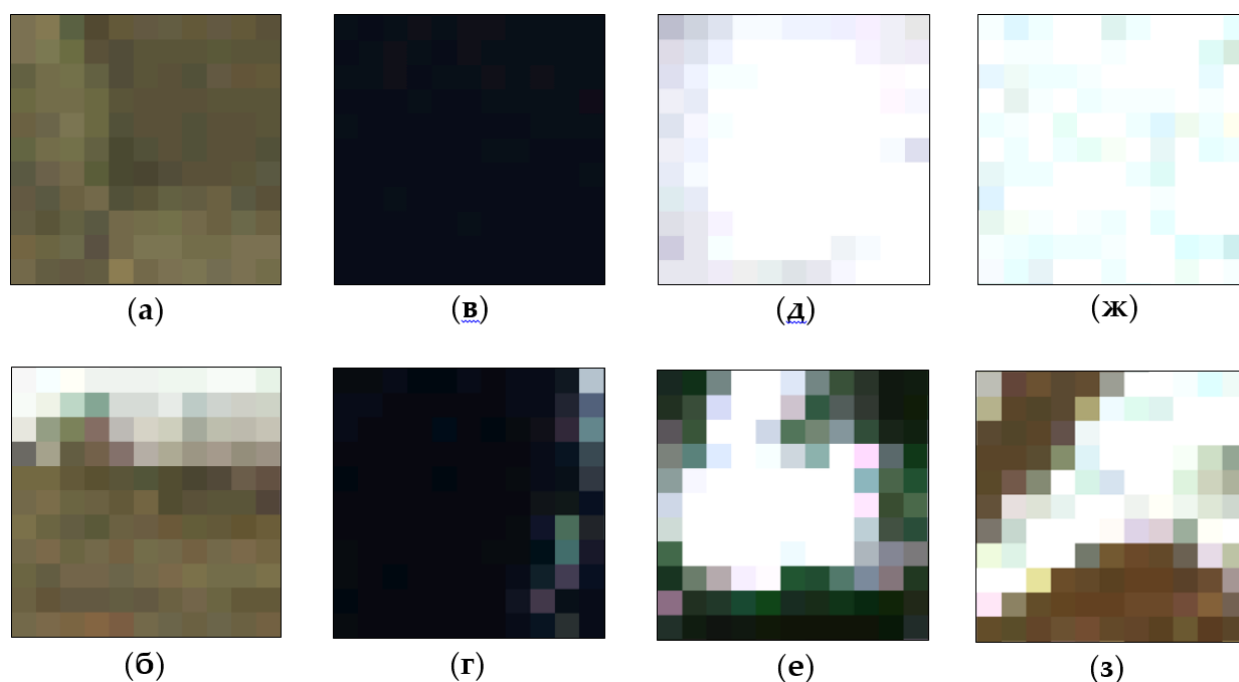


Рис. 1. Пример текстур: (а, б) земля; (в, г) вода; (д, е) облачность; (ж, з) снег.

Основные характеристики прибора МСУ-ГС КА "Электро-Л" №2 показаны в табл. 1.

Таблица 1

№ канала	Спектральный диапазон, мкм	Разрешение	№ канала	Спектральный диапазон, мкм	Разрешение
1	0.5 – 0.65	1 км	6	7.5 – 8.5	4 км
2	0.65 – 0.8		7	8.2 – 9.2	
3	0.8 – 0.9		8	9.2 – 10.2	
4	3.5 – 4.01	9	10.2 – 11.2		
5	5.7 – 7.0	4 км	10	11.2 – 12.5	

После того, как *ОВ* была набрана, сохраненные пиксели выбирались в качестве центра текстуры для обучения нейронной сети. Каждая текстура состояла из данных 5 спектральных каналов спутникового прибора: три в видимом спектре и два в инфракрасном. Данные по трем каналам видимого были скорректированы с использованием зенитного угла Солнца путем деления значений излучения (количества солнечной энергии, отраженной от поверхности Земли в определенной области) в каналах на косинус угла для каждого пикселя. Зенитный угол был добавлен ввиду того, что значения яркости в спутниковых каналах физически зависят от угла, под которым солнце освещает рассматриваемый участок [9]. Эта информация позволяет нейронной сети сводить на нет различия в значениях яркости при изменении положения Солнца. Инфракрасные каналы остались без изменений, поскольку они содержат информацию о температуре в кельвинах.

Последующие операции аугментации – зеркальное отражение и поворот каждой текстуры на 30° – позволили увеличить объем выборки до 890 тысяч. Для обучения и тестирования нейронной сети все данные были линейно нормализованы в диапазоне от 0 до 1. Нормализация проводилась отдельно по каждому каналу данных со спутника, в качестве 0 и 1 выбиралось соответствующее минимальное и максимальное значение по каждому параметру из всей обучающей выборки.

Обучение нейронной сети

Процесс обучения нейронной сети заключается в итерационной перенастройке весовых коэффициентов у нейронов заданным алгоритмом оптимизации. Обучение проходит по алгоритму обратного распространения ошибки, согласно которому на каждом шаге с зафиксированными весами нейронов вычисляется реакция сети на поданные данные, от входного слоя к выходному. Далее по полученной реакции определяется ошибка выходного слоя, называемая функцией потерь, которая при обратном проходе сети изменяет весовые коэффициенты каждого нейрона по методу градиентного спуска. В качестве функции потерь выбрана кросс-энтропия, так как она хорошо зарекомендовала себя при многоклассовой классификации [10]. Данная функция задается следующим выражением:

$$J(p, q) = - \sum_C p(C) \cdot \log q(C), \quad (1)$$

где C – количество прогнозируемых классов; $p(C)$ – желаемая вероятность для каждого класса; $q(C)$ – фактическая вероятность для каждого класса.

В качестве алгоритма оптимизации был выбран алгоритм Adam (*adaptive moment estimation*), в котором сочетаются идеи сохранения инерции

и масштабирования градиента. Суть инерции заключается в использовании значений градиентов предыдущих итераций обучения для частичного «смягчения» быстрого изменения градиента на каждой итерации. Идея масштабирования заключается в выборе шага обучения каждого параметра в отдельности, основываясь на «истории» прошлых градиентов для данного параметра. Это реализуется за счет деления каждого элемента в градиенте на квадратный корень суммы квадратов прошлых соответствующих элементов. По сравнению с другими методами оптимизации такое сочетание позволяет достичь более высокого быстродействия в поиске оптимальных значений параметров [11].

Обучение нейронных сетей происходит до обнаружения минимальных изменений показателей метрик машинного обучения для валидационной выборки по результатам нескольких, идущих подряд, итераций обучения. После чего обученная модель проходит проверку на тестовой выборке для определения пригодности ее классифицирующих свойств.

Архитектура нейронной сети

В качестве решения поставленной задачи рассматривались уже существующие архитектуры нейронных сетей, модифицированные для работы с отобранными текстурами, которые хорошо зарекомендовали себя в области компьютерного зрения, в частности, *VGG-16* и *AlexNet* [12,13]. В ходе экспериментов было выявлено, что такие архитектуры либо не могли обучиться различать представленные классы (*VGG-16*), либо имели малые показатели метрик машинного обучения. В качестве метрик машинного обучения в данной задаче рассматриваются следующие: точность (*precision*) (2), полнота (*recall*) (3) и *f-мера(F)* (4), которая является гармоническим средним между точностью и полнотой.

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (3)$$

$$F = 2 \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (4)$$

где *TP*, *FP* и *FN* – количество соответственно истинно-положительных, ложно-положительных и ложно-отрицательных решений.

Стоит также отметить, что идея *transfer learning* и *fine-tuning* предварительно обученных нейросетей *VGG-16* и *AlexNet* применительно к данным дистанционного зондирования не рассматривалась в этой работе, поскольку

сети имеют входные данные в формате *RGB*. Это означает, что на вход можно подавать данные только по трем каналам, что не соответствует нашей идее обнаружения облаков и снежного покрова с использованием данных пяти спутниковых каналов. Поэтому был рассмотрен подход к полноценному обучению этих сетей, были немного изменены их слои, чтобы можно было обучить их на подготовленных данных с пятью каналами. Однако даже в этом случае точность обнаружения в результате обучения *AlexNet* и *VGG-16* оказалась довольно низкой (показатель *F* имел значения меньше 84%). Авторами было решено разработать собственную архитектуру для решения проблемы.

Первая версия архитектуры содержит следующие слои: *Input*; *Conv2D* [14]; *MaxPooling* [15]; *FC*; *ReLU*, *tanh* и *Softmax*; *Dropout* [16]. Эта модель предназначена для работы с текстурами размером 31×31 пиксель, и ее структура показана на рис. 2. Такой размер текстуры был выбран потому, что он показал хорошие результаты при работе над аналогичной задачей [17].

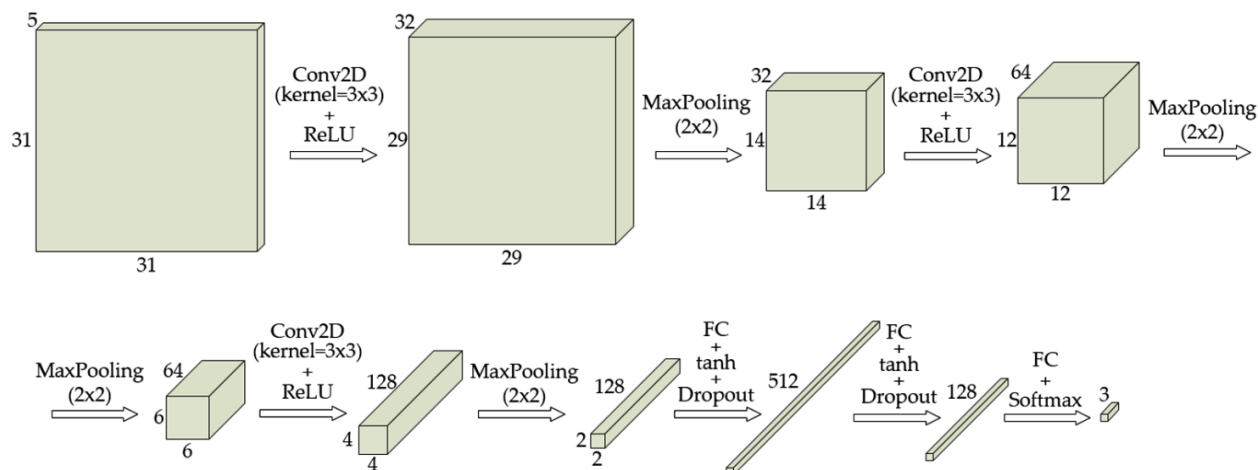


Рис. 2. Первая версия архитектуры рассматриваемого нейросетевого классификатора.

Рассмотрим каждый слой подробнее. *Input* – входные данные (текстуры) с указанием размерности. *Conv2D* – сверточный слой, который имеет ядро свертки – матрицу весов. Ядро, «двигаясь» по изображению, поэлементно выполняет операцию умножения с частью данных, над которой находится, формируя сигнал активации. Такая процедура выполняется для каждой локации, над которой проходит ядро, формируя таким образом карту признаков, передающуюся следующим слоям. *MaxPooling* – сверточный слой, который по принципу действия очень похож на *Conv2D*, но с тем отличием, что при «движении» происходит не операция умножения, а из части данных выбирается только максимальное значение. Этот слой был разработан для сжатия карт объектов в результате сверточной операции. Кроме того, данный слой позволяет уменьшить зависимость от смещения объектов на изображении.

ражении. *Dropout* – слой регуляризации, который предназначен для снижения вероятности переобучения нейронных сетей. Алгоритм работы состоит в прореживании, суть которого заключается в том, что при обучении часть нейронов исключается из сети, тем самым обнуляя несущие ими признаки. *FC (Fully connected)* – полносвязный слой, который представляется как слой обычного многослойного перцептрона. Он состоит из нейронов, каждый из которых соединен со всеми нейронами на предыдущем слое. Основной целью слоя является классификация при помощи создания модели сложной нелинейной функции. *ReLU*, *tanh* и *Softmax* – функции активации, необходимые для формирования выходного сигнала нейронов. Он получается при преобразовании функцией активации взвешенной суммы входных сигналов нейрона. Для активации выходных нейронов сверточных слоев применяется функция *ReLU*. Ее основные преимущества – отсутствие ресурсоемких операций, отсутствие затухание градиента при обучении, быстрое обучение. Для активации выходных нейронов полносвязных слоев используется функция активации – гиперболический тангенс (*tanh*), имеющая следующие преимущества: область значений лежит в диапазоне $[-1; 1]$, простое вычисление производной. Для последних слоев, дающих результат классификации в виде прогноза класса, используется функция активации *Softmax*.

Для этой архитектуры оценка F на тестовой выборке составила 93,56%. Для дополнительной оценки качества нейронной сети были построены маски облаков и снежного покрова для нескольких полигонов.

На следующем этапе проводился эксперимент по определению размера входных текстур. Этот эксперимент продемонстрировал, что самые высокие значения метрик в тестовом наборе данных достигаются, когда входные текстуры имеют размер 11×11 пикселей. С уменьшением размера входной текстуры архитектура нейронной сети была немного изменена. Это необходимо, потому что после трех-четырёх операций свертки размер последующих слоев уменьшился до $1 \times 1 \times 5$ и дальнейшая свертка не имела смысла. Принимая это во внимание, архитектура, представленная ранее, была переработана для нового размера входных данных, как показано на рис. 3.

При использовании новой архитектуры оценка F для набора данных проверки составила 98,52%. По сравнению с предыдущей версией количество ложных отнесений облаков к снегу и земли к облакам существенно уменьшилось.

Помимо того, скорость расчета масок резко снизилась из-за уменьшения размерности входных данных и сокращения количества слоев нейронной сети.

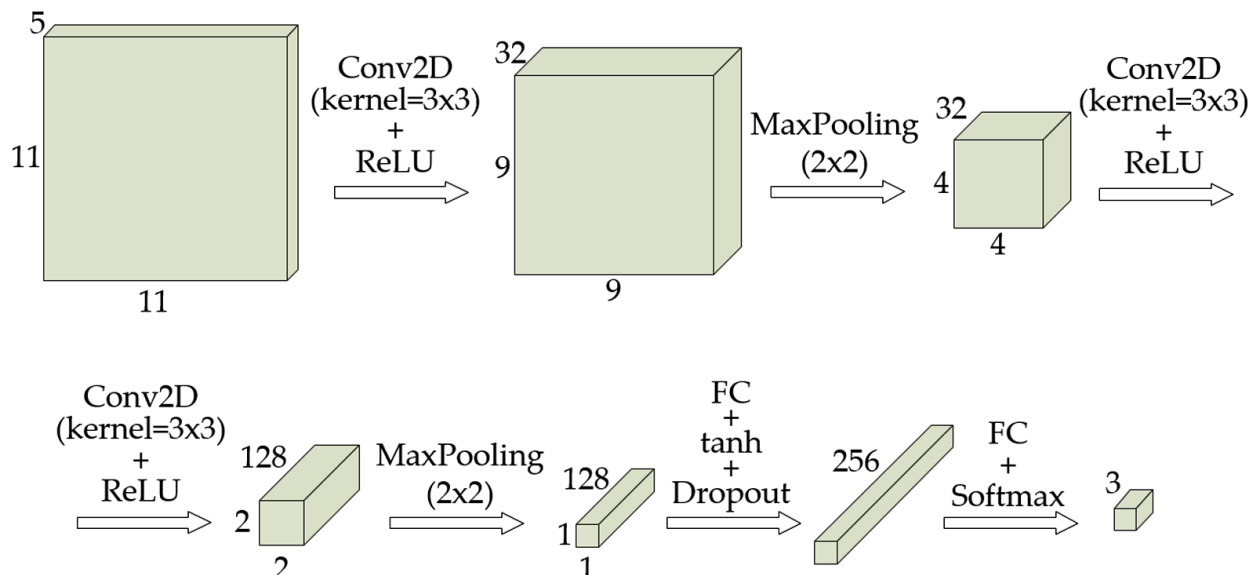


Рис. 3. Вторая версия архитектуры рассматриваемого нейросетевого классификатора.

Затем был добавлен слой *Batch normalization* (нормализация пакетов), который основан на нормализации вывода уровня активации и делает обучение нейронной сети более устойчивым [18]. Также были удалены слои *MaxPooling*, поскольку слой сжимает текстуру, а в данной работе классификация всегда выполняется только для центрального пикселя текстуры и текстуры изначально имеют небольшую размерность. По этим причинам данный слой был удален из окончательной версии архитектуры. На основе результатов вычисленных статистических коэффициентов была выбрана архитектура классификатора нейронной сети, представленная на рис. 4, поскольку у нее были самые высокие значения показателей.

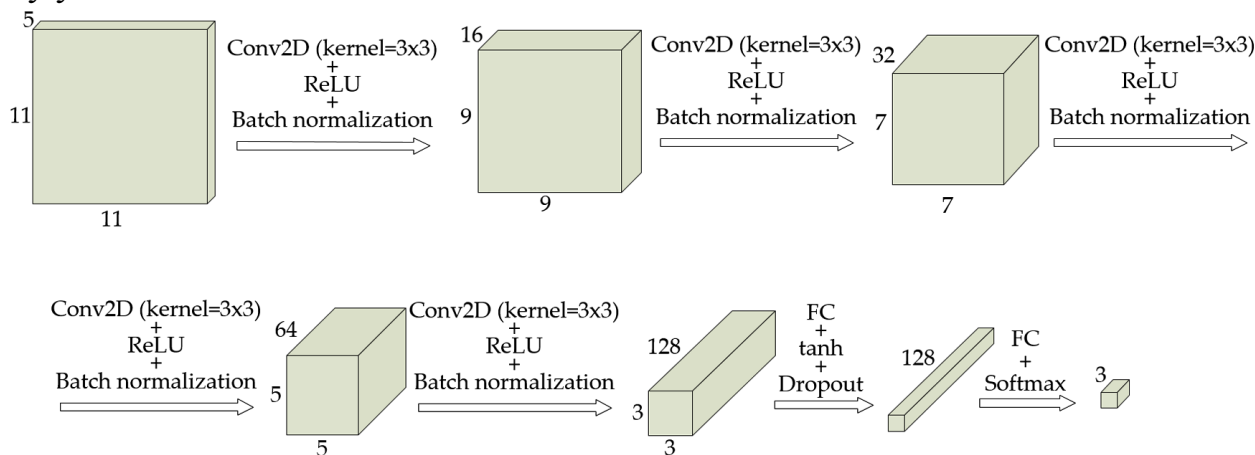


Рис. 4. Третья версия архитектуры рассматриваемого нейросетевого классификатора.

Архитектура содержит четыре блока свертки (*Conv2D*), что связано с размером входного сетевого уровня 11×11 . Каждый блок включает сверточный уровень, уровень активации (*ReLU*) и *Batch normalization*. Впоследствии два *FC* слоя разделяются слоем регуляризации (*dropout*), среди которых размер выходного слоя последнего слоя соответствует количеству клас-

сов. Оценка F этой архитектуры для тестовой выборки составила 99,67%. Кроме того, для этой нейронной сети была проведена *fine-tuning*. Были заморожены первые два уровня, а скорость обучения снижена в 100 раз. Оценка F на тестовой выборке немного увеличилась – до 99,68%. Количество ложных классификаций холодной земли, как облаков или снега и облаков у береговой линии, уменьшилось. Примеры визуального изменения результатов классификации спутниковых снимков по трем вариантам нейронных сетей представлены на рис. 5.

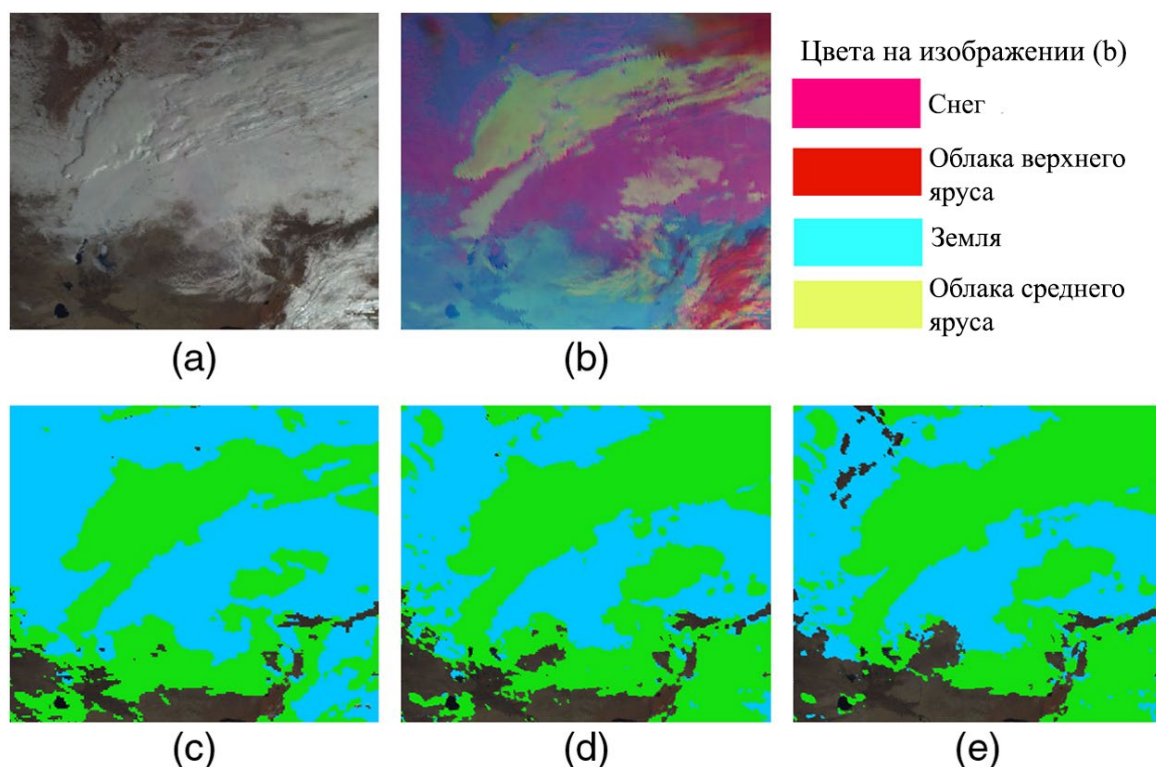


Рис. 5. Сравнение масок облачности и снежного покрова в трех вариантах нейросетевых классификаторов.

На рис. 5 представлены следующие фрагменты: (a) – мультиспектральное изображение каналов 0,9, 0,7 и 0,6 мкм (RGB) прибора $MCU-ГС$; (b) – мультиспектральное изображение каналов 0,9, 3,8 и 10,7 мкм (RGB) прибора $MCU-ГС$; (c) – RGB -изображение спектральных каналов 0,9, 0,7 и 0,6 мкм с маской, полученной с помощью первой версии классификатора (зеленый: облака; синий: снег); (d) – RGB -изображение спектральных каналов 0,9, 0,7 и 0,6 мкм с маской, полученной второй версией классификатора (зеленый: облака; синий: снег); (e) – RGB -изображение спектральных каналов 0,9, 0,7 и 0,6 мкм с маской, полученной с помощью третьей версии классификатора (зеленый: облака; синий: снег).

Все обученные модели были протестированы на основе статистических коэффициентов с использованием тестовой выборки. *Precision*, *Recall* и

показатель F были получены таким же образом, как и во время обучения. В табл. 2 представлены полученные значения статистических коэффициентов для обученных моделей нейронных сетей для трех основных этапов экспериментов. Для окончательной архитектуры значения статистических коэффициентов для тестовой выборки до и после *fine-tuning* были практически идентичны (различия в тысячных долях 1%).

Таблица 2

Архитектура	<i>Precision</i>	<i>Recall</i>	<i>F</i>
С входным слоем размерностью 31x31 пикселей	92,84	92,8	92,82
С входным слоем размерностью 11x11 пикселей и слоем <i>MaxPooling</i>	98,33	98,32	98,32
С входным слоем размерностью 11x11 пикселей и без слоя <i>MaxPooling</i>	99,45	99,46	99,45

Разработанный классификатор позволяет обрабатывать спутниковые снимки всего диска Земли с прибора *МСУ-ГС* (разрешение 2764x2764 пикселей), поступающие с частотой два снимка в час. Работа нейронной сети занимает около 6 минут, что позволяет получать необходимые продукты практически в реальном времени. Измерения времени проводились для программы, написанной на *Python 3.7* (*ОС Windows 10*, *Keras 2.2.4*, *Tensorflow 1.13.1*) с видеокартой *NVIDIA GTX 1060*.

Заключение

В данной работе описан алгоритм на основе *CNN*, который может классифицировать поверхность на три класса, а именно – земля/вода, снег и облачные образования на изображениях дистанционного зондирования, полученных с прибора *МСУ-ГС КА «Электро-Л №2»*. Алгоритм полностью автоматический – нет необходимости вручную корректировать какие-либо значения в описанном классификаторе. Работает в любое время года днем. Для обеспечения точной работы нужно использовать пять спектральных каналов с длинами волн 0,6, 0,7, 0,9, 3,8 и 10,7 мкм в случае прибора *МСУ-ГС*, а также информацию о зенитном угле Солнца. Еще одно преимущество алгоритма в том, что нет необходимости использовать дополнительные сторонние данные, требуются только данные с прибора *МСУ-ГС*.

Однако возможны некоторые ошибки в классификации снега во время облачности из-за специфики используемых данных и ошибок на краях изображений вследствие геометрического искажения. Разработанный алгоритм может быть адаптирован для обнаружения снежного покрова и облачных образований на основе данных, полученных с других спутников.

В настоящее время алгоритм используется в оперативной практике Дальневосточного центра НИЦ «Планета», в составе программы восстановления параметров облачности.

ЛИТЕРАТУРА

1. *Liangpei Z., Lefei Z., Bo D.* Deep Learning for Remote Sensing Data: A technical tutorial on the state of the art // IEEE Geoscience and remote sensing magazine, 2016.
2. *Zhu X.X., Tuia D., Mou L., Xia G., Zhang L., Xu F., Fraundorfer F.* Deep Learning in Remote Sensing: A Review // IEEE Geoscience and remote sensing magazine, 2017.
3. *Yang C., Rongshuang F., Muhammad B., Xiucheng Y., Jingxue W., Li W.* Multilevel Cloud Detection for High-Resolution Remote Sensing Imagery Using Multiple Convolutional Neural Networks // ISPRS Int. J. Geo-Inf. – 2018. – Vol. 7(5). – P. 181.
4. *Yun Y.* Temporal Data Mining via Unsupervised Ensemble Learning. – Elsevier, 2016. – P. 172.
5. *Блощинский В.Д., Кучма М.О., Андреев А.И.* Применение сверточных нейронных сетей в задачах детектирования облачности и снежного покрова по данным прибора МСУ-ГС КА Электро-Л №2 // Материалы 17-й Всероссийской открытой конференции «Современные проблемы дистанционного зондирования Земли из космоса». – М.: ИКИ РАН. – 2019. – С. 18.
6. *Кучма М.О., Воронин В.В., Блощинский В.Д.* Анализ возможности использования сверточной нейронной сети для тематической обработки спутниковых изображений // Материалы Восьмой Всероссийской научно-практической конференции «Решение». – Березники. Изд-во Пермского национального политехнического ун-та. – 2019. – С. 70-72.
7. *Кучма М. О., Воронин В. В.* Создание обучающей выборки для сверточной нейронной сети в задаче тематической обработки спутниковых изображений // Сб. науч. трудов «Информационные технологии XXI века». – Хабаровск: Изд-во Тихоокеан. гос. ун-та. – 2020. – С. 187-191.
8. *Kuchma M.O., Bloschchinskiy V.D.* Electro Textures Dataset. figshare. Dataset, 2019. URL: <https://doi.org/10.6084/m9.figshare.9722723.v3>.
9. *Cracknell A.P.* The Advanced Very High Resolution Radiometer. – London: Taylor and Francis Ltd, 1997.
10. *Mueller A., Guido S.* An introduction to Machine Learning with Python. O'Reilly, 2017.
11. *Diederik P.K., Jimmy L.B.* Adam: a method for stochastic optimization // Proceedings of International Conference on Learning Representations, 2015.
12. *Yang C., Rongshuang F., Muhammad B., Xiucheng Y., Jingxue W., Li W.* Multilevel Cloud Detection for High-Resolution Remote Sensing Imagery Using Multiple Convolutional Neural Networks // ISPRS Int. J. Geo-Inf. – 2018. – Vol. 7(5). – P. 181.
13. *Simonyan K., Zisserman A.* Very Deep Convolutional Networks For Large-Scale Image Recognition, 2015.
14. *Jay Kuo C.-C.* Understanding Convolutional Neural Networks with A Mathematical Model, 2016.
15. *Boureau Y., Ponce J., LeCun Y.* A theoretical analysis of feature pooling in visual recognition // In 27th Int. Conf. Mach. Learn (ICML'10). Madison, Omnipress, 2010. – P. 111-118.

16. *Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R.* Dropout: a simple way to prevent neural networks from overfitting // J. Mach. Learn. Res. – 2014. – Vol. 15. – P. 1929-1958.
17. *Kramareva L.S., Andreev A.I., Simonenko E.V., Sorokin A.A.* The use of a convolutional neural network for detecting snow according to the data of the multichannel satellite device of meteor-M No. 2 spacecraft // Procedia Comput. Sci. – 2019. – Vol. 150. – P. 368-375.
18. *Ioffe S., Szegedy C.* Batch normalization: accelerating deep network training by reducing internal covariate shift, 2015.

Статъа представлена к публикации членом редколлегии Чъе Ен Уном.

E-mail:

Кучма Михаил Олегович – m.kuchma@dvrspod.ru;

Воронин Владимир Викторович – 004183@pnu.edu.ru;

Блощинский Владислав Дмитриевич – v.bloshchinsky@dvrspod.ru.