

УДК 004.93

КЛАССИФИКАЦИЯ ГИПЕРСПЕКТРАЛЬНЫХ ИЗОБРАЖЕНИЙ С ПОМОЩЬЮ СВЁРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

© В. И. Козик, Е. С. Нежевенко

*Институт автоматики и электрометрии СО РАН,
630090, г. Новосибирск, просп. Академика Коптюга, 1
E-mail: nejevenko@iae.nsk.su*

Показано, что для классификации фрагментов гиперспектрального изображения весьма эффективны предварительная трансформация его спектральных признаков к главным компонентам и последующее распознавание с применением свёрточной нейронной сети, которая обучена на выборке, составленной из фрагментов этого изображения. Получены высокие проценты правильной классификации при работе с крупноформатным гиперспектральным изображением, причём часть классов, на которые разбито гиперспектральное изображение, очень близки между собой и соответственно трудноразличимы по гиперспектрам. Исследована зависимость правильной классификации от изменения размеров фрагментов, из которых составлены обучающая и валидационная выборки, и от параметров свёрточной нейронной сети.

Ключевые слова: классификация, гиперспектральное изображение, главные компоненты, свёрточные нейронные сети, обучение.

DOI: 10.15372/AUT20210202

Введение. За последние годы всё большее распространение получает проблема классификации изображений местности по их гиперспектральным измерениям. В работах [1–4] эта проблема исследована довольно подробно. Главное — было показано, что существенное повышение процента правильной классификации гиперспектральных изображений (ГСИ) обеспечивает учёт не только спектральной характеристики, но и пространственной структуры ГСИ. Отметим, что в подавляющем числе работ в этой области используются классические алгоритмы классификации. Однако в последнее время можно считать доказанным, что наилучшие (если не уникальные) результаты в области распознавания изображений, к которой относится и классификация, были получены путём использования свёрточных нейронных сетей и глубокого обучения. Таким образом, применение этих подходов к ГСИ более чем обосновано. Работы в этом направлении уже проводились [5–7]. Наиболее обстоятельно в [7] исследуется зависимость точности классификации ГСИ от самых различных параметров свёрточных сетей, с помощью которых производится классификация. Методы, предложенные в этой работе, будем использовать с осторожностью: некоторые из них нам кажутся не совсем обоснованными.

Характеристика классифицируемого объекта. Исследуемое гиперспектральное изображение — это участок местности, полученный в рамках программы AVIRIS (Airborne Visible Infrared Imaging Spectrometer) на опытном полигоне Индиан Пайс (штат Индиана, США). Размер изображения 614×1408 пикселей, разрешение 20 м/пиксель, число каналов 220 в диапазоне 0,4–2,5 мкм. В RGB-представлении ГСИ показано на рис. 1.

На рис. 2 в псевдоцветах представлена разметка этого ГСИ на классы. Всего классов 57. Однако специфика избранного нами метода пространственной обработки такова, что в некоторых областях объекты классификации не могут быть сформированы из-за небольших размеров областей. Поэтому названия классов будут даны после обучения сети (при выделении действующих классов).



Рис. 1



Рис. 2

Структура свёрточной сети. Мы не будем описывать принципы функционирования свёрточных сетей, они достаточно хорошо отражены в литературе. Представим схему свёрточной сети (рис. 3) и рассмотрим её особенности.

Будем использовать 3D свёрточную нейронную сеть. Входной слой — 3D-куб размерами $M \times N \times F$; $M \times N$ — это размер фрагмента области, принадлежащей к одному классу, описывающему пространственные характеристики области; F — число признаков, представляющих спектральные характеристики области. Размер фрагмента $M \times N$ играет важнейшую роль. Слишком малый размер фрагмента не позволит выявить его пространственные особенности. При большом размере фрагментов уменьшается их количество в классе, поскольку области, принадлежащие к классам, имеют произвольную форму, а фрагменты — прямоугольную, вплоть до того, что на некоторые классы не выпадает ни одного фрагмента. Обсудим теперь третье измерение F . В [7] утверждается, что наиболее целесообразно в этом измерении использовать все спектральные компоненты без преобразования, поскольку последнее применяется только для того, чтобы путём уменьшения слоёв сократить вычислительные процедуры. В действительности это не совсем так. Спектральные компоненты, являющиеся составляющими третьего слоя, сильно коррелированы. А из теории распознавания известно, что использование коррелированных признаков уменьшает правильность распознавания, поэтому для эффективного распознавания осуществляется, как правило, декорреляция признаков. Поэтому была проведена предобработка спектральной информации путём преобразования её к главным компонентам (ГК). Количество ГК и соответственно число слоёв во входной плоскости определяются, например, вычислением «каменистой осыпи» — графика убывания собственных значений. Для

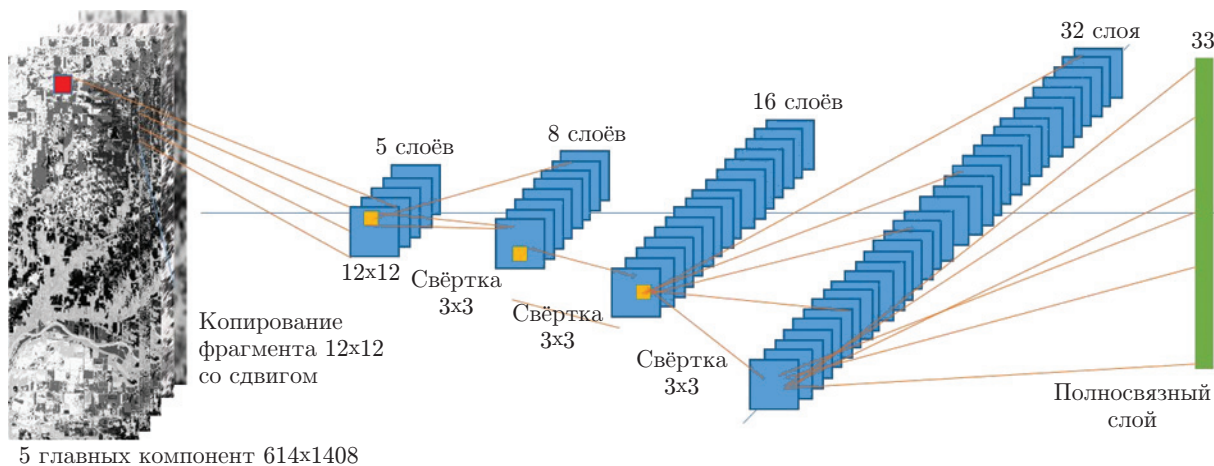


Рис. 3

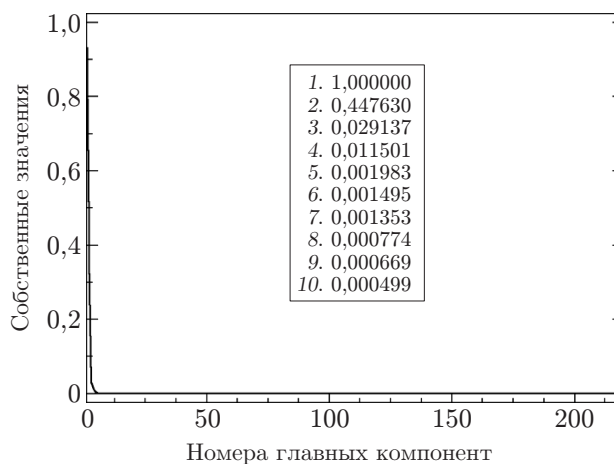


Рис. 4

исследуемого ГСИ график, характеризующий осыпь, представлен на рис. 4.

По оси x отложены номера главных компонент, по оси y — их нормированные собственные значения. Поскольку собственные значения убывают очень быстро, на графике показаны 10 первых числовых собственных значений. На графике видно, что уже пятое собственное значение составляет $1/500$ от первого значения, это обозначает, что на него приходится 0,2 % дисперсии спектральных компонент, поэтому большую часть экспериментов будем проводить при числе главных компонент, равных 5. Всего в представленной свёрточной нейронной сети 5 слоёв, размер ядер 3×3 . Уменьшение размеров карт признаков (Subsampling) в нашей сети не применяется, так как классифицируемые изображения и так достаточно малого размера. Размерность выходного слоя равна числу классов, идентифицированных на ГСИ с учётом размера фрагментов.

Важнейшим этапом при использовании свёрточных нейронных сетей и глубокого обучения для классификации является формирование обучающей выборки. В нашем случае объектами этой выборки являются фрагменты ГСИ.

Результаты экспериментов по классификации ГСИ. Перечислим все этапы классификации ГСИ (отметим, что производилась классификация в среде MATLAB, кроме нахождения главных компонент, вычисляемых в программе ENVI):

1. Вычисляются главные компоненты ГСИ.

2. Формируются директории классов от 1 до 57.

3. Из файла, содержащего разбику ГСИ на классы (см. рис. 2), с использованием плавающего окна размером $M \times N$ и сдвигов shift_M , shift_N выбираются фрагменты, все элементы которых принадлежат к одному классу.

4. Окно, все элементы которого относятся к одному классу, идентифицируется как объект, принадлежащий к этому классу, и его координаты определяются на изображении. Используя эти координаты, из файла, содержащего выбранные главные компоненты, берётся фрагмент размером $M \times N \times F$ и записывается в соответствующую директорию классов. В каждой директории регистрируются файлы по числу найденных фрагментов данного класса. По результатам их формирования определяется число классов и вносятся поправки в функцию обучения.

5. Корректируются параметры сети: число слоёв, размеры ядер, количество карт признаков.

6. Корректируются параметры процедуры обучения: число классов, количество эпох обучения; объекты каждого непустого класса разделяются на обучающую и валидационную выборки.

7. Запускается процедура обучения.

Рассмотрим результаты экспериментов. При этом отметим, что единственным критерием эффективности той или иной процедуры: генерации фрагментов, обучения, классификации — будем считать точность классификации, которая определяется как отношение числа правильно классифицированных объектов к общему числу объектов (термин «точность» применяется наряду с термином «вероятность правильной классификации»).

Для оценки точности классификации при формировании обучающей и валидационной (тестовой) выборок применяется кроссвалидация (валидация на отложенных данных (Hold-Out Validation)) [8].

Выборка разделена случайным образом на обучающую и валидационную в отношении 7 : 3, причём они не пересекаются. Метод Hold-Out применяется в случаях больших датасетов, что соответствует нашему случаю (общее количество объектов 34 596, в каждом классе не меньше 50).

Рассмотрим, как будет изменяться точность классификации при изменении некоторых параметров свёрточной сети (размера фрагмента, определяющего размерность входного слоя, количества слоёв нейронной сети, числа эпох обучения и числа ГК). Фрагменты выбраны квадратными, для обеспечения максимального количества фрагментов возьмём сдвиги $\text{shift}_M = \text{shift}_N = 1$. На рис. 5 приведена точность классификации в зависимости от размера фрагмента при фиксированных значениях числа слоёв сети 5, количества эпох обучения 50 при двух значениях количества ГК 5 и 10.

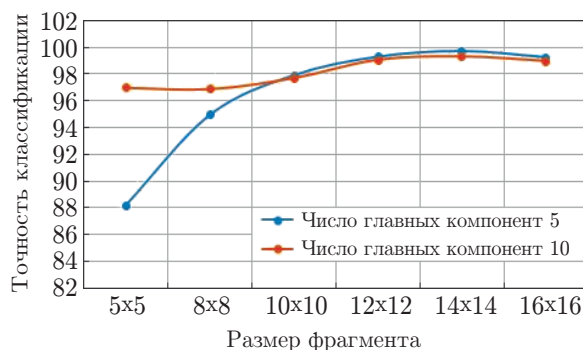


Рис. 5

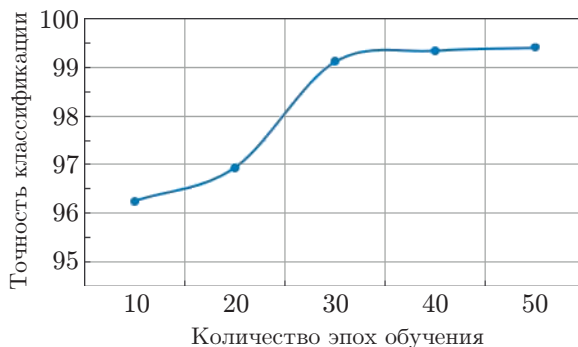


Рис. 6

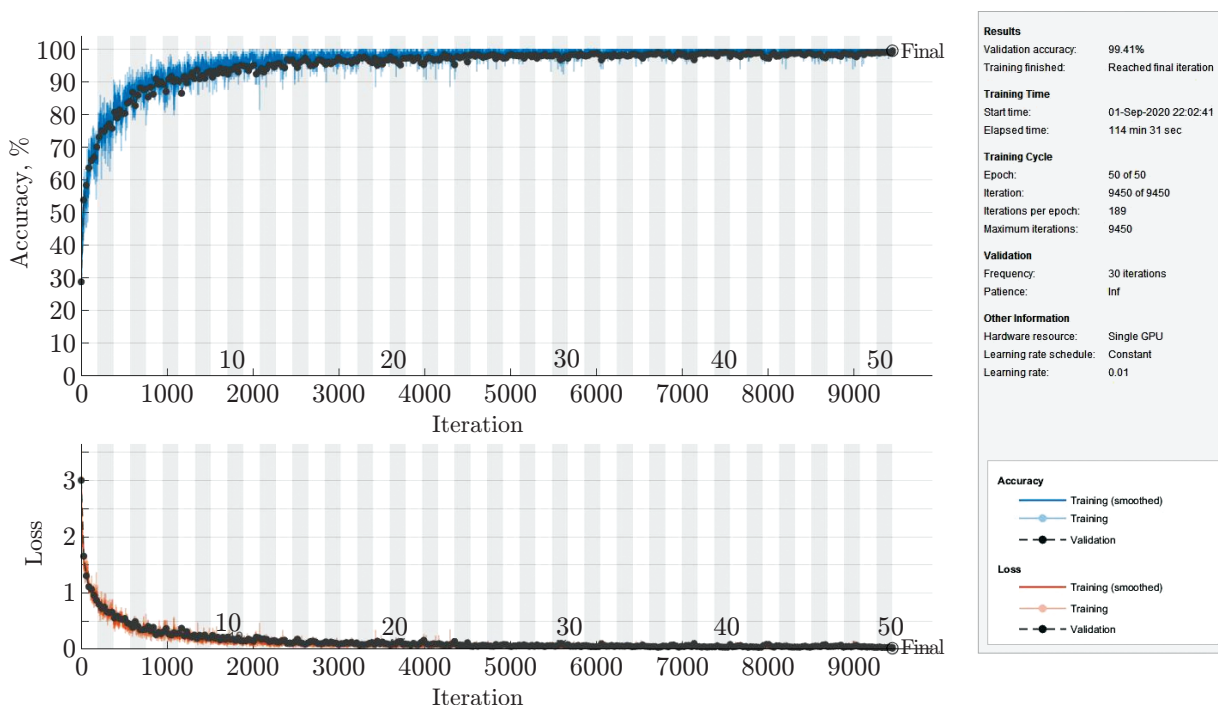


Рис. 7

Интересно, что при 10 ГК точность классификации значительно меньше зависит от размера фрагмента. Важным параметром любых нейронных сетей, в том числе и свёрточных, является скорость обучения `InitialLearnRate`. В нашем случае она составляет 0,01. Скорость не изменялась, поскольку задача быстрого обучения не ставилась. Другой важный параметр обучения — количество эпох обучения, которое определяет не только скорость, но и конечную точность классификации. На рис. 6 показана зависимость финальной точности классификации от количества эпох обучения при размере фрагмента 12×12 и числе ГК — 5. Видно, что точность классификации монотонно увеличивается с числом эпох, причём резкий скачок имеет место от 20 до 30 эпох. Однако эта зависимость в значительной степени определяется размерами фрагментов. На рис. 7 (верхняя часть) показано изменение точности классификации в процессе обучения для фрагментов размером 12×12 элементов, на рис. 8 — для фрагмента 5×5 . Для фрагментов 12×12 при 30 эпохах точность классификации фактически выходит в насыщение, тогда как для размера 5×5 точность продолжает расти при 50 эпохах.

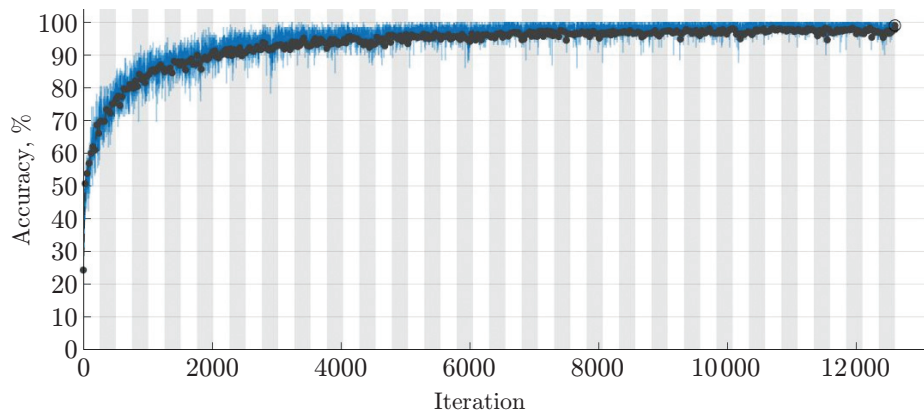


Рис. 8

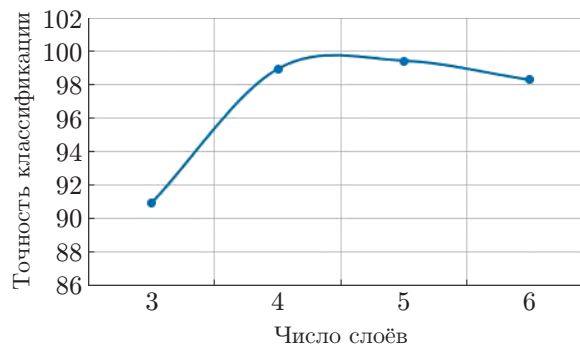


Рис. 9

На рис. 9 показана зависимость точности классификации от числа слоёв нейронной сети. Оптимальное число слоёв 5.

Параметр самого успешного эксперимента по классификации — размер фрагмента $M \times N = 12 \times 12$. В то же время следует заметить, чем больше фрагменты, тем меньше количество их в классе и количество самих классов. В табл. 1 показаны названия классов, число фрагментов в классе и точности классификации для двух фрагментов размерами 5×5 и 12×12 , сдвиги $\text{shift}_M = \text{shift}_N = 1$, число слоёв 5, число эпох обучения 50. Результаты: число полученных классов для фрагмента размером 5×5 — 45, при размере 12×12 — 33. Из названий классов видно, что мы не объединяли близкие классы в один (например, посе́вы кукурузы, посе́вы соевых бобов), как в [4]. Понятно, что гораздо легче отличить посе́вы кукурузы от леса, чем различить разные посе́вы той же кукурузы или соевых бобов. Различия близких классов показаны в [9], с ними и будем сравнивать результаты данной работы. Отметим, что при фрагментах размером 12×12 практически все трудноразличимые объекты (посе́вы кукурузы и соевых бобов) классифицируются с очень высокой (зачастую 100-процентной) вероятностью. Следует также указать, что результаты, полученные в данной работе, существенно превосходят результаты [9] с оговоркой: в последней нет проблемы покрытия области, принадлежащей к классу, с прямоугольными окнами, поэтому в ней можно классифицировать области со сложной конфигурацией. Что касается сравнения с публикациями [5–7], то в представленной работе значительно больше классов, в том числе трудноразличимых.

Высокие значения точности классификации вызывают сомнение в том, что имеет место переобучение нейронной сети. Согласно [10] переобучение, переподгонка (overtraining, overfitting) — нежелательное явление, возникающее при решении задач обучения по преце-

Таблица 1

№ класса	Название класса	Размер фрагмента 5 × 5		Размер фрагмента 12 × 12	
		Число фрагментов	Точность классификации	Число фрагментов	Точность классификации
1	Голая почва	168	0,7200	0	—
2	Здания	5541	0,8454	2621	0,9987
3	Кукуруза	6005	0,8440	2269	0,9927
4	Кукуруза, запад-восток	60	1,0000	0	—
5	Кукуруза, север-юг	648	0,8247	169	1,0000
6	Кукуруза, обычная вспашка	2541	0,7507	368	1,0000
7	Кукуруза, обычная вспашка, запад-восток	8571	0,7841	2481	0,9970
8	Кукуруза, обычная вспашка, север-юг	12803	0,8578	4241	0,9914
9	Кукуруза, обычная вспашка, север-юг, орошаемое	116	0,8286	0	—
10	Кукуруза, обычная вспашка - ?	307	0,4783	45	1,0000
11	Кукуруза, малоразрушающая вспашка	96	0,7586	0	—
12	Кукуруза, малоразрушающая вспашка, запад-восток	2006	0,8605	896	0,9963
13	Кукуруза, малоразрушающая вспашка, север-юг	3255	0,9037	1099	0,9970
14	Кукуруза без вспашки	879	0,8598	93	1,0000
15	Зерно, без вспашки, запад-восток	200	0,7833	30	1,0000
16	Кукуруза, без вспашки, север-юг	2705	0,8633	1304	1,000
17	Трава	32	0,8000	0	—
18	Трава / Деревья	561	0,9881	91	0,9630
19	Сено	48	0,7857	0	—
20	Сено?	964	0,9343	443	0,9925
21	Сено-люцерна	628	0,9894	191	1,0000
22	Не обрезается	180	0,9630	0	—
23	Овес	324	1,0000	70	1,0000
24	Выгон	3377	0,9704	1996	1,0000
25	Соевые бобы	1324	0,8363	326	1,0000
26	Соевые бобы?	152	0,9130	0	—
27	Соевые бобы, север-юг	72	0,9091	0	—
28	Соевые бобы, обычная вспашка	957	0,7526	227	1,0000
29	Соевые бобы, обычная вспашка?	792	0,6933	239	0,9722
30	Соевые бобы, обычная вспашка, запад-восток	4715	0,8112	2000	0,9900
31	Соевые бобы, обычная вспашка, север-юг	3830	0,6762	1057	0,9905
32	Соевые бобы, обычная вспашка, борозды	384	0,8174	40	0,6667
33	Соевые бобы, обычная вспашка, сорняки	116	0,8000	0	—
34	Соевые бобы, посаженные рядами	4680	0,8832	1046	0,9777
35	Соевые бобы, малоразрушающая вспашка	424	0,9449	50	1,0000
36	Соевые бобы, малоразрушающая вспашка, запад-восток	512	0,9221	65	1,0000
37	Соевые бобы, малоразрушающая вспашка, нарезные	2507	0,9109	689	1,0000
38	Соевые бобы, малоразрушающая вспашка, север-юг	2212	0,6717	721	1,0000
39	Соевые бобы, запад-восток	673	0,9356	185	1,0000
40	Соевые бобы, без вспашки, запад-восток	1054	0,9747	356	1,0000
41	Соевые бобы, без вспашки, север-юг	180	0,7222	0	—
42	Соевые бобы, без вспашки, посаженные рядами	2324	0,9813	436	1,0000
43	Деревья?	48	1,0000	0	—
44	Пшеница	1664	0,9880	636	1,0000
45	Лес	20324	0,9405	8115	0,9988

дентам, когда вероятность ошибки обученного алгоритма на объектах тестовой выборки оказывается существенно выше, чем средняя ошибка на обучающей выборке. Из рис. 7 (нижней части), характеризующего поведение ошибки в процессе обучения, следует, что ошибка на тестовой выборке весьма незначительно (доля процента) превышает ошибку на обучающей выборке, из чего следует, что переобучения в данном случае нет, а следовательно, никаких мер для его устранения предпринимать не надо.

Заключение. Таким образом, в предложенной работе экспериментально показано, что при классификации гиперспектральных изображений преобразование к главным компонентам спектральных составляющих и дальнейшее пространственное преобразование путём разбивки главных компонент на небольшие фрагменты, обучение на части этих фрагментов свёрточной нейронной сети и классификация с её помощью ГСИ обеспечивает очень высокий процент правильной классификации (99,43 %). При этом количество классов достаточно велико (33) и среди них есть очень близкие классы (8 классов кукурузных посевов и 13 классов посевов соевых бобов). При уменьшении размера фрагмента точность классификации несколько падает, однако число распознанных классов растёт. Для фрагмента размером 5×5 точность классификации числа главных компонент 5 составляет около 88 %, а для 10 — 97 %, число распознанных классов при этом — 45. Исследовано влияние изменения параметров свёрточной сети и числа главных компонент на точность классификации.

Следует отметить, что столь высокая точность классификации объясняется во многом способом формирования обучающей и валидационной выборок, характеризуемого очень тесным их перемешиванием. В то же время очевидно, что и такой способ классификации может применяться в некоторых случаях. В частности, этот способ формирования (случайное разделение на обучающую и валидационную выборки) хорошо работает, когда классы занимают и большие, и малые площади. Наши исследования, не вошедшие в данную публикацию, показывают, что в случае малых площадей высокая точность классификации получается и при пространственном разделении обучающей и валидационной выборок.

Финансирование. Работа выполнена при поддержке Министерства науки и высшего образования РФ (государственная регистрация № АААА-А17-117052410034-6).

СПИСОК ЛИТЕРАТУРЫ

1. **Borhani M., Ghassemian H.** Hyperspectral image classification based on non-uniform spatialspectral kernels // Proc. of the Iranian Conf. on Intelligent Systems. Bam, Iran, Febr. 4–6, 2014. DOI: 10.1109/IranianCIS.2014.6802579.
2. **Борзов С. М., Потатуркин О. И.** Исследование эффективности спектрально-пространственной классификации данных гиперспектральных наблюдений // Автометрия. 2017. **53**, № 1. С. 32–42. DOI: 10.15372/AUT20170105.
3. **Борзов С. М., Потатуркин О. И.** Классификация гиперспектральных данных при различных способах формирования обучающих выборок // Автометрия. 2018. **54**, № 1. С. 89–97. DOI: 10.15372/AUT20180112.
4. **Борзов С. М., Потатуркин О. И.** Спектрально-пространственные методы классификации гиперспектральных изображений. Обзор // Автометрия. 2018. **54**, № 6. С. 64–86. DOI: 10.15372/AUT20180607.
5. **Fang B., Li Y., Zhang H., Chan Ch.-W. J.** Hyperspectral images classification based on dense convolutional networks with spectral-wise attention mechanism // Remote Sens. 2019. **11**, Iss. 2. P. 159. DOI: 10.3390/rs11020159.

6. **Audebert N., Saux B., Lefèvre S.** Deep learning for classification of hyperspectral data: A comparative review // IEEE Geoscience and Remote Sens. Magazine. 2019. 7(2). P. 159–173.
7. **Li Y., Zhang H., Shen Q.** Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network // Remote Sens. 2017. 9, Iss. 1. P. 1–21. DOI: 10.3390/rs9010067.
8. **Кросс-валидация.** Университет ИТМО. 2020. URL: <http://neerc.ifmo.ru/wiki/index.php?title=Кросс-валидация> (дата обращения: 12.10.2020).
9. **Нежевенко Е. С.** Нейросетевая классификация трудноразличимых типов растительности по гиперспектральным признакам // Автометрия. 2019. 55, № 3. С. 62–70. DOI: 10.15372/AUT20190308.
10. **Переобучение.** URL: <http://www.machinelearning.ru/wiki/index.php?title=Переобучение> (дата обращения: 12.10.2020).

Поступила в редакцию 12.10.2020

После доработки 28.01.2021

Принята к публикации 04.02.2021
