

КЛАССИФИКАЦИЯ ГИПЕРСПЕКТРАЛЬНЫХ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ С ИСПОЛЬЗОВАНИЕМ КОМБИНИРОВАННЫХ 3D–2D СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Л.Т. Ньян

А.И. Гаврилов

М.Ч. До

nyanlin54@gmail.com

alexgavrilov@bmmstu.ru

dominhtrieuvhp@gmail.com

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация

Аннотация

Классификацию гиперспектральных изображений используют для анализа данных дистанционного зондирования Земли. Сверточная нейронная сеть — один из наиболее часто используемых методов обработки визуальных данных на основе глубокого обучения. Предложена комбинированная спектральная сверточная нейронная сеть для классификации гиперспектральных изображений. На начальном этапе предложена простая комбинированная обученная модель глубокого обучения, которая строится путем объединения 2D и 3D сверточных нейронных сетей для извлечения более глубоких пространственно-спектральных объектов с меньшим числом 3D–2D-сверток. 3D-сеть облегчает совместное пространственно-спектральное представление объектов из стека спектральных полос. Для классификации гиперспектральных изображений использованы функции 3D–2D сверточных нейронных сетей. Для уменьшения размерности применен алгоритм метода главных компонент. Эксперименты по классификации изображений гиперспектральных изображений проведены над наборами данных дистанционного зондирования индийских сосен (Indian Pines), университета Павии (University of Pavia) и сцены Салинаса (Salinas Scene). Первый слой карты объектов использован в качестве входных данных для последующих слоев при прогнозировании конечных меток для каждого гиперспектрального пикселя. Предлагаемый метод не только включает в себя преимущества расширен-

Ключевые слова

Сверточные нейронные сети, гиперспектральное изображение, обучение, дистанционное зондирование Земли

ного извлечения признаков из сверточных нейронных сетей, но и полностью использует спектральную и пространственную информацию. Эффективность предложенного метода проверена на трех эталонных наборах данных. Результаты показывают, что основанная на таких сетях многофункциональная система обучения значительно повышает точность классификации (более 99 %)

Поступила 19.04.2021

Принята 04.08.2021

© Автор(ы), 2022

Введение. Компьютерная классификация изображений дистанционного зондирования заключается в выявлении и классификации информации о земной поверхности и окружающей ее среде на снимках дистанционного зондирования в целях выявления характерной информации, соответствующей изображению, и ее извлечения [1]. Гиперспектральные изображения (ГСИ) снимают специализированными дистанционными датчиками на летательном аппарате и собирают из спектральных данных, отраженных наземными объектами в определенном районе Земли. Гиперспектральная визуализация приводит к появлению многочисленных полос изображений, которые затрудняют анализ изображения вследствие увеличенного объема данных [2]. Гиперспектральное изображение — трехмерный куб данных, содержащий двумерную пространственную информацию (объект изображения признаков) и одномерную спектральную информацию (спектральные полосы). Как правило, спектральные полосы встречаются в виде тонких длин волн, в то время как особенности наземного покрова и формы демонстрируют неравенство и ассоциацию между соседними пикселями с различных направлений на определенной длине волны [3].

Иерархическая архитектура сверточных нейронных сетей (СНС) — наиболее эффективный способ изучения визуальных представлений. Такая сеть предлагает потенциал для описания структурных характеристик на высоких уровнях в соответствии с процедурой извлечения иерархических признаков [4, 5]. Фундаментальной проблемой в этих визуальных задачах является моделирование внутриклассового внешнего вида и изменения формы объектов. В настоящей работе предложен новый подход, основанный на глубоком обучении, который использует 3D–2D СНС для эффективного использования спектральной и пространственной информации из гиперспектральных данных (рис. 1).

Наборы данных и обучение. Научному сообществу доступно несколько наборов данных, полученных с использованием гиперспектрального датчика. Обычно данные приходят вместе с аннотациями для оценки эффективности классификации. Здесь использованы три набора данных ди-

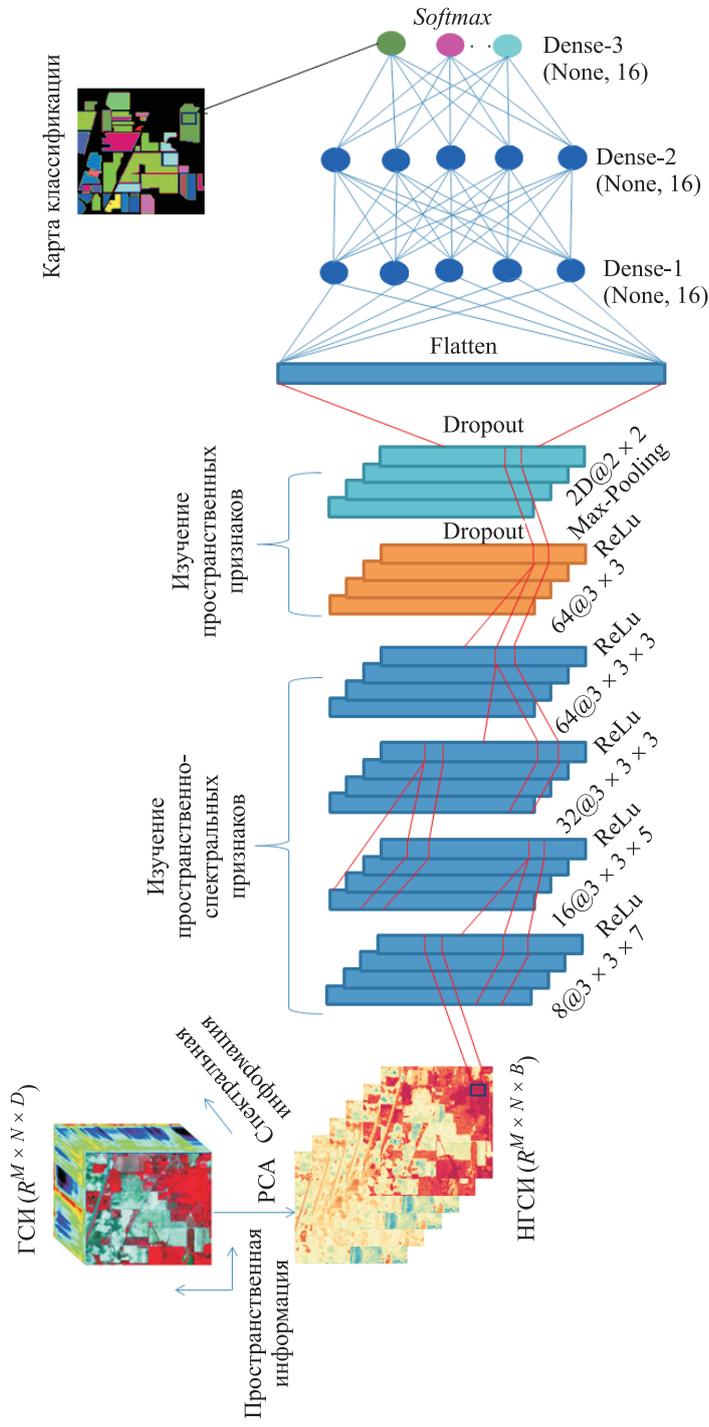


Рис. 1. Структура предлагаемой комбинированной 3D–2D-модели

станционного зондирования: 1) индийские сосны (Indian Pines, IP); 2) университет Павии (University of Pavia, PU); 3) сцены Салинаса (Salinas Scene, SA) [6]. Набор данных IP содержит изображения с пространственными размерами 145×145 пикселей и 224 спектральными полосами в диапазоне значений длин волн $400 \dots 2500$ нм, из которых 24 спектральные полосы, охватывающие область поглощения воды, исключены. Набор данных IP имеет 16 классов растительности. Размерность пространственной размерной информации набора данных 610×340 пикселей, 103 спектральные полосы в диапазоне значений длин волн $430 \dots 8600$ нм. Набор данных PU разделен на девять классов городского типа подстилающей поверхности [7–9]. Функция изображения набора данных SA содержит 512×217 пикселей и 224 спектральные полосы в диапазоне значений длин волн $360 \dots 2500$ нм. Из набора данных SA удалены 20 полос, связанных с водой. Эти наборы данных дистанционного зондирования имеют 3D-куб данных, который состоит из 2D пространственной информации (функции изображения) и 1D спектральной информации (спектральные полосы).

Уменьшение размерности 3D кубических данных. Большие наборы данных и высокие размерности изображения все более распространены и часто трудны для интерпретации. Здесь использован алгоритм анализа главных компонент (PCA) для уменьшения размерности данных 3D-куба. Данные гиперспектральных изображений имеют высокую размерность. Для устранения спектральной избыточности сначала применяют традиционный анализ главных компонент (PCA) по исходным данным ГСИ (I) наряду со спектральными полосами. Гиперспектральные пиксели демонстрируют смешанные классы почвенного покрова, внося высокую внутриклассовую изменчивость и межклассовое сходство в исходные входные данные. Пусть спектрально-пространственный гиперспектральный куб данных обозначен через $I \in R^{M \times N \times D}$, где I — исходные входные данные; M — ширина; N — высота; D — число спектральных полос/глубина. Каждый пиксель ГСИ в данных I содержит D спектральных мер и образует вектор метки $Y = (y_1, y_2, \dots, y_c \in R^{1 \times 1 \times C})$, где C — категория растительного покрова. Алгоритм анализа главных компонент уменьшает число спектральных полос от D до B , сохраняя при этом те же пространственные размеры (т. е. ширину M и высоту N). Сокращены только спектральные полосы так, что сохраняется пространственная информация, которая очень важна для распознавания любого объекта. Представление уменьшенного куба данных PCA на $X \in \text{HGSI} (R^{M \times N \times B})$, где X — модифици-

рованный входной сигнал после PCA; B — число спектральных полос после PCA.

Для использования методов классификации изображений куб данных ГСИ разделен на небольшие перекрывающиеся 3D-патчи, метки истинности которых определяются меткой центрального пикселя. Созданы 3D соседние патчи из куба X , $P \in R^{S \times S \times B}$, центрированного в пространственном положении (α, β) , патчи покрывают окно $S \times S$ или пространственный экстенд и все спектральные группы B . Общее число генерированных 3D-патчей n из X задают через $(M-S+1)(N-S+1)$. Таким образом, 3D-патч на месте (α, β) обозначен как $P_{\alpha, \beta}$, охватывает ширину от $\alpha - (S-1)/2$ до $\alpha + (S-1)/2$, высоту от $\beta - (S-1)/2$ к $\beta + (S-1)/2$ и по всем B спектральным полосам PCA уменьшает куб данных X .

После уменьшения размерности раздела PCA сети выдают входные данные с четырех сверточных слоев для получения карты объектов, которая генерируется с использованием 3D-свертки по нескольким смежным полосам во входном слое, что захватывает спектральную информацию [10–12]. За 2D сверточным слоем следует 3D сверточный слой для выполнения извлечения объектов из 3D-слоя, чтобы создать лучшую карту объектов без преобразования знаков данных ГСИ. 2D сверточный слой только занимает пространственную информацию. 2D-сети не могут обрабатывать спектральную информацию:

$$v_{i,j}^{x,y} = G \left(b_{i,j} + \sum_{\tau=1}^{d_{l-1}} \sum_{\rho=-\gamma}^{\gamma} \sum_{\sigma=-\delta}^{\delta} \omega_{i,j,\tau}^{\sigma,\rho} v_{i-1,\tau}^{x+\sigma, y+\rho} \right). \quad (1)$$

Уравнение (1) можно рассматривать как 2D-свертку, которая генерирует $v_{i,j}^{x,y}$ карту объектов на особенность признака j в пространственном положении (x, y) на i -м слое путем вычисления суммы точечного произведения входа и 2D-фильтра:

$$v_{i,j}^{x,y} = G \left(b_{i,j} + \sum_{\tau=1}^{d_{l-1}} \sum_{\lambda=-a}^a \sum_{\rho=-\gamma}^{\gamma} \sum_{\sigma=-\delta}^{\delta} \omega_{i,j,\tau}^{\sigma,\rho,\lambda} v_{i-1,\tau}^{x+\sigma, y+\rho, z+\lambda} \right). \quad (2)$$

Уравнение (2) можно рассматривать как 3D-свертку, которая может извлекать спектральные и пространственные характеристики представления одновременно из данных ГСИ. Структура предлагаемой комбинированной сети (см. рис. 1) состоит из одной 2D-свертки (уравнение (1)), трехмерных сверток (уравнение (2)) и трех полносвязных слоев.

В (1) и (2) введены следующие обозначения: G — функция активации; $b_{i,j}$ — параметр смещения (bias) для j -й карты пространственных объектов

i -го слоя; d_{l-1} — число объектов на карте на $(l - 1)$ -го слоя; $2\gamma + 1$, $2\delta + 1$, $2a + 1$ — ширина, высота и глубина ядра; $\omega_{i,j}$ — значение весового параметра для j -й карты пространственных объектов i -го слоя. В предлагаемой модели использованы размеры трех сверточных слоев с 3D-ядрами свертки. Для сверточных ядер первого слоя 3D-размеры содержат восемь узлов, размер ядра $3 \times 3 \times 3$, для второго слоя — 16 узлов, размер ядра $3 \times 3 \times 5$, третьего слоя — 32 узла, размер ядра $3 \times 3 \times 3$, четвертого слоя — 64 узла, размер ядра $3 \times 3 \times 3$. После проведения 3D-сверток применяют один сверточный слой с 2D-ядром, содержащим 64 узла, размер ядра 3×3 . 2D-слой различает пространственную информацию в различных спектральных диапазонах без значительной потери спектральной информации, что очень важно для данных ГСИ. Слой объединения (max-pooling layer) выполняется после 2D сверточного слоя для фильтрации и уменьшения числа выходов. В предлагаемой модели общее число узлов конечного выходного слоя или плотного слоя зависит от числа классов трех эталонных наборов данных. Например, число узлов в последнем плотном слое равно 16, что соответствует числу классов в наборе данных IP (рис. 2).

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	(None, 21, 21, 30, 1)	0
conv3d_5 (Conv3D)	(None, 19, 19, 24, 8)	512
conv3d_6 (Conv3D)	(None, 17, 17, 20, 16)	5776
conv3d_7 (Conv3D)	(None, 15, 15, 18, 32)	13856
conv3d_8 (Conv3D)	(None, 13, 13, 16, 64)	55360
reshape_2 (Reshape)	(None, 13, 13, 1024)	0
dropout_5 (Dropout)	(None, 13, 13, 1024)	0
conv2d_2 (Conv2D)	(None, 11, 11, 64)	589888
max_pooling2d_2 (MaxPooling2D)	(None, 5, 5, 64)	0
dropout_6 (Dropout)	(None, 5, 5, 64)	0
flatten_2 (Flatten)	(None, 1600)	0
dense_4 (Dense)	(None, 256)	409856
dropout_7 (Dropout)	(None, 256)	0
dense_5 (Dense)	(None, 128)	32896
dropout_8 (Dropout)	(None, 128)	0
dense_6 (Dense)	(None, 16)	2064
Total params: 1,110,208.0		
Trainable params: 1,110,208.0		
Non-trainable params: 0.0		

Рис. 2. Набор параметров комбинированной модели 3D–2D CNN для набора данных IP и окна размером 21×21 пикселей и 30 полос

Постановка эксперимента. Как уже было отмечено выше, использованы три набора ГСИ. После выбора наборов данных выполнено уменьшение размерности исходных данных алгоритмом PCA, затем — уменьшение размерности из предыдущего состояния. Предлагаемая модель содержит четыре сверточных слоя (четыре 3D сверточных слоя и один 2D сверточный слой), слой объединения, выравнивающий слой (flatten layer), два полностью связанных слоя и выходной слой (или плотный). Кроме выходного слоя, во всех слоях добавлена функция активации ReLU. Нелинейность ReLU применена к выходу каждого сверточного и двух полностью связанных слоев. Принятая в работе функция ReLU представляет собой простую нелинейную функцию, которая производит значения 0 или 1, соответствующие положительному или отрицательному входу нейрона. Подтверждено, что функция ReLU может повысить производительность сетей во многих случаях [13]. Классификационная модель *Softmax* добавлена к выходу последнего полностью подключенного слоя. Для выполнения классификации с изученными репрезентативными признаками к верхнему слою предлагаемой сети применен оператор *Softmax*. Этот оператор — одна из вероятностных классификационных моделей, которая измеряет корреляцию выходного и эталонного значений с помощью оценки вероятности. Следует отметить, что в конструкции СНС *Softmax* может быть применен по всем спектральным каналам для всех пространственных местоположений сверточным способом [14]. Функция потерь (categorical_cross entropy) использована для определения пригодности предложенной модели при достижении заданной цели. Оптимизатор *Adam* применен для минимизации функции потерь (целей) путем обновления весов с помощью градиентов [15]. По результатам классификации определена оптимальная скорость обучения 0,001. Размер входного окна каждого объекта для трех наборов данных IP, UP и SA, пиксель: 11×11 ; 13×13 ; 15×15 ; 17×17 ; 19×19 ; 21×21 ; 25×25 (рис. 3, табл. 1). Чтобы выполнить сравнение, извлекли одно и то же пространственное измерение в 3D-патчах входного объема для различных наборов данных: $11 \times 11 \times 30$ для IP; $11 \times 11 \times 15$ для PU и SA.

Решение задачи. Для оценки эффективности классификации ГСИ использованы общая (OT) и средняя (СТ) точности, а также коэффициент Каппа (КК). Общая точность (рис. 4) представляет собой число правильно классифицированных выборок из общего числа тестовых выборок, СТ — среднее значение точности классификации по классам. Коэффициентом Каппа оценивают результаты деления изображений (учитывает число образцов точек, которые правильно разделены, и число неправильных делений для оценки результатов классификации) [15, 16]. Наборы данных раз-

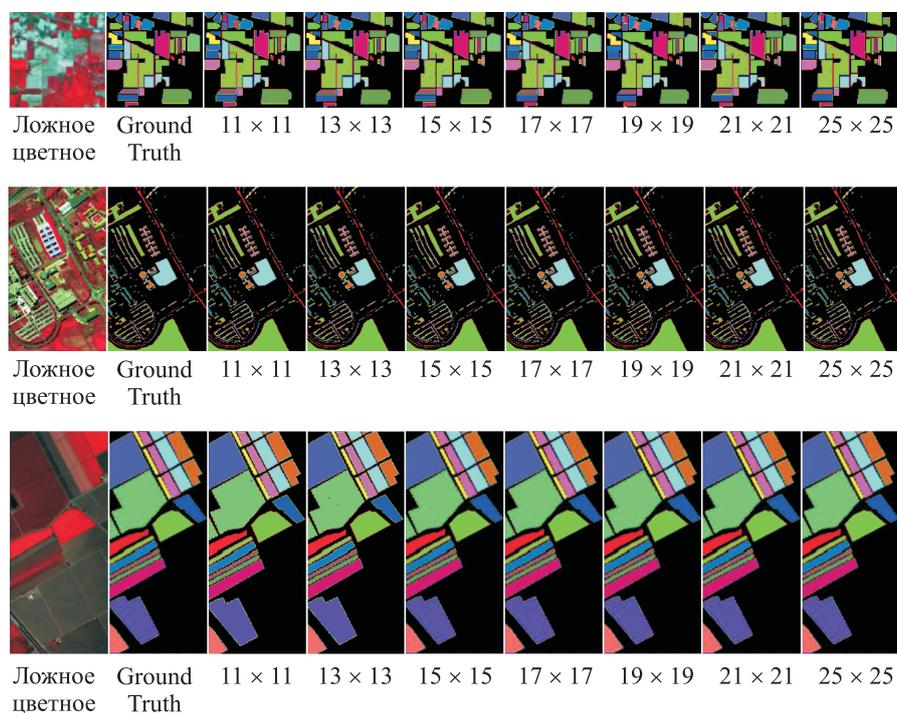


Рис. 3. Карты классификации по наборам данных IP, PU, SA с использованием окна размерами 11×11 ; 13×13 ; 15×15 ; 17×17 ; 19×19 ; 21×21 ; 25×25 предлагаемой модели 3D–2D СНС

деляют случайным образом, например, 30 % обучающих групп и 70 % тестовых (рис. 5). Точности обучения в зависимости от размеров окна для наборов данных SA, PU и IP приведены на рис. 6. Для трех наборов данных СТ более 99 %. Проведено сравнение результатов классификационных карт (см. рис. 3). Использованы методы 1D [8], 2D [9], 3D [10], Hybrid:SN [11] и предлагаемая модель 3D–2D СНС для наборов данных IP, UP и SA (табл. 2). Точности классификации и качество классификационной карты предлагаемой модели (3D–2D СНС) лучше, чем моделей 1D СНС, 2D СНС, 3D СНС, Hybrid:SN (рис. 7). Предлагаемая модель уменьшает погрешность обобщения (переоснащения и недостаточной подгонки) во время обучения и позволяет снизить вес модели. Модель разработана не только для уменьшения переоснащения, но и для быстрой оптимизации модели и лучшей общей производительности, чем модель Hybrid:SN. Модель значительно сократила число параметров (более 4 млн), чем Hybrid:SN. Все эксперименты проведены с использованием компьютера на базе шестиядерного процессора AMD Ryzen5 с тактовой частотой 3,6 ГГц и графическим процессором GTX 1070 (GPU) с 8 и 32 ГБ оперативной памяти.

Таблица 1

**Точность классификации наборов данных IP, PU, SA
в зависимости от размеров окна для предлагаемой модели 3D–2D СНС
и время обработки обучения t_o и тестирования t_t**

Набор данных	Точность классификации предложенной модели			t_o (за 100 эпох), с	t_t , с	Число параметров
	ОТ	КК	СТ			
<i>11 × 11</i>						
IP	91,38 ± 0,6	90,15 ± 0,3	88,56 ± 0,6	≈ 55,09	≈ 14,6	716 992
PU	99,93 ± 0,0	99,92 ± 0,0	99,88 ± 0,0	≈ 85,2	≈ 53,8	294 201
SA	99,97 ± 0,0	99,96 ± 0,0	99,96 ± 0,0	≈ 82,8	≈ 64,2	164 032
<i>13 × 13</i>						
IP	98,22 ± 0,2	97,22 ± 0,2	88,11 ± 0,4	≈ 74,4	≈ 14,2	716 992
PU	99,94 ± 0,0	99,93 ± 0,0	99,91 ± 0,0	≈ 85,8	≈ 52,5	163 129
SA	99,94 ± 0,0	99,93 ± 0,0	99,94 ± 0,0	≈ 123,6	≈ 66,2	164 032
<i>15 × 15</i>						
IP	99,38 ± 0,0	99,30 ± 0,0	97,06 ± 0,2	≈ 128	≈ 15,4	766 144
PU	99,96 ± 0,0	99,95 ± 0,0	99,92 ± 0,0	≈ 133	≈ 48,6	212 281
SA	99,99 ± 0,0	99,99 ± 0,0	99,99 ± 0,0	≈ 144,6	≈ 63,7	213 184
<i>17 × 17</i>						
IP	99,20 ± 0,0	99,09 ± 0,0	90,82 ± 0,0	≈ 210	≈ 12,8	848 064
PU	99,94 ± 0,0	99,93 ± 0,0	99,90 ± 0,0	≈ 180	≈ 50,4	294 201
SA	100,0 ± 0,0	100,0 ± 0,0	100,0 ± 0,0	≈ 206,4	≈ 64,2	295 104
<i>19 × 19</i>						
IP	99,67 ± 0,0	99,61 ± 0,0	99,38 ± 0,0	≈ 249,6	≈ 14,9	962 752
PU	99,94 ± 0,0	99,92 ± 0,0	99,90 ± 0,0	≈ 254,4	≈ 54,4	408 889
SA	100,0 ± 0,0	100,0 ± 0,0	100,0 ± 0,0	≈ 62,4	≈ 64,2	409 792
<i>21 × 21</i>						
IP	99,86 ± 0,0	99,84 ± 0,0	99,74 ± 0,0	≈ 325	≈ 15,3	1 110 208
PU	99,97 ± 0,0	99,96 ± 0,0	99,92 ± 0,0	≈ 315	≈ 52,5	556 345
SA	100,0 ± 0,0	100,0 ± 0,0	100,0 ± 0,0	≈ 437	≈ 64,2	557 248
<i>25 × 25</i>						
IP	99,80 ± 0,0	99,77 ± 0,0	99,37 ± 0,0	≈ 510	≈ 19,7	1 503 424
PU	99,99 ± 0,0	99,99 ± 0,0	99,99 ± 0,0	≈ 141,6	≈ 55,8	949 561
SA	100,0 ± 0,0	100,0 ± 0,0	100,0 ± 0,0	≈ 543,4	≈ 67,2	950 464

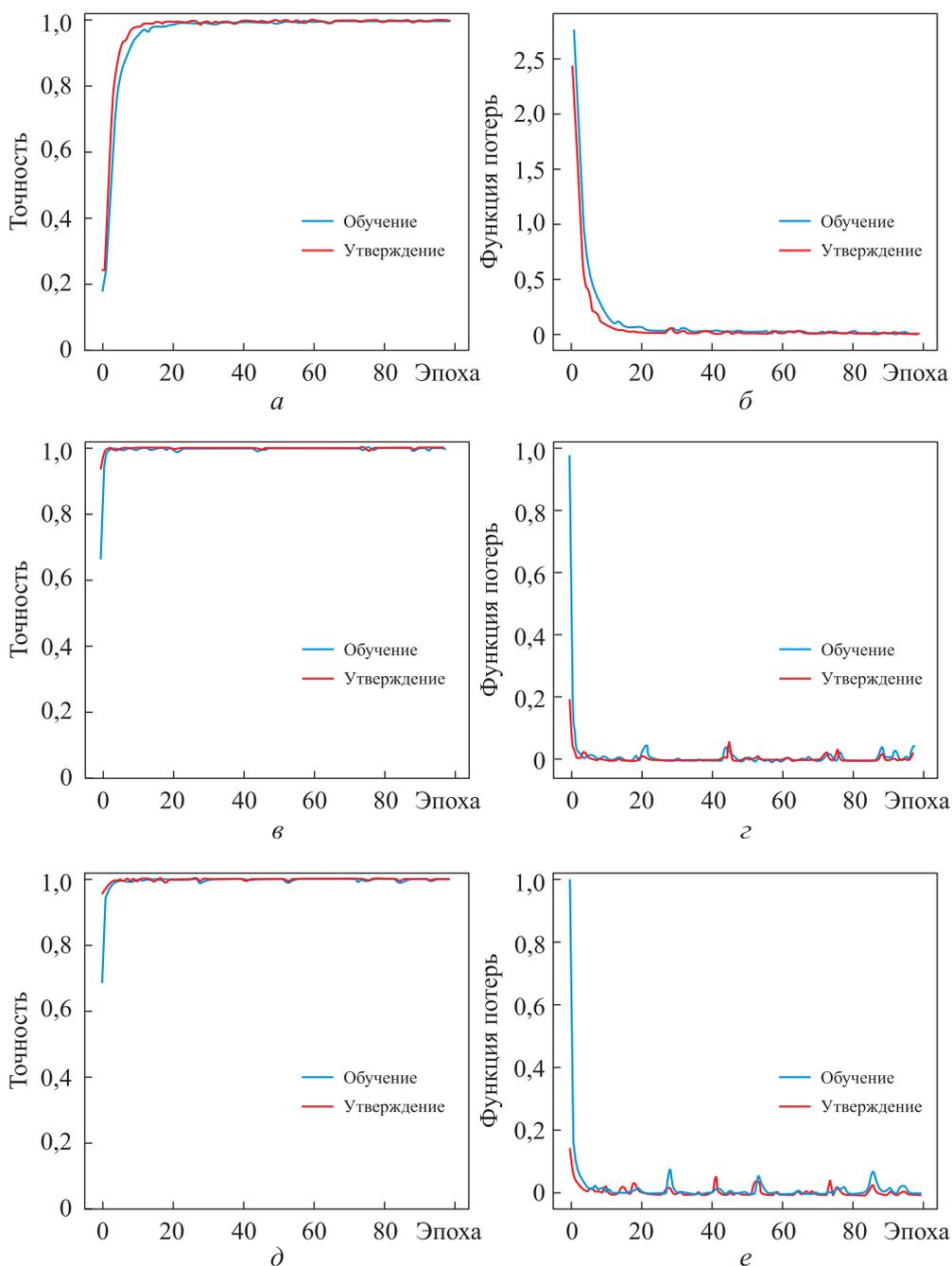


Рис. 4. Общая точность и функция потерь для трех наборов данных IP, SA и PU с использованием окна размером 25×25 для предлагаемой модели 3D-2D CNN

№	Цвет	Категория	Образец	Обуче- ние 70 %	Тестиро- вание 30 %	1D CHC	2D CHC	3D CHC	Hybrid: SN	3D-2D CHC
1		Люцерна	46	32	14	100%	97.83%	100%	100%	90%
2		Кукуруза по till	1428	1000	428	99.60%	94.82%	97.90%	99%	100%
3		Кукурузная мята	830	581	249	99.14%	97.23%	96.39%	100%	100%
4		Кукуруза	237	166	71	98.20%	99.58%	95.78%	98%	100%
5		Трава-пастбище	483	338	145	97.35%	99.59%	97.63%	100%	99%
6		Трава-деревья	730	511	219	99.80%	99.59%	99.02%	100%	100%
7		Скошенная трава-пастбище	28	20	8	100%	100%	85%	100%	100%
8		Сено валковское	478	335	143	100%	100%	98.81%	100%	100%
9		Овсянка	20	14	6	100%	100%	85.71%	83%	100%
10		Соя по till	972	680	292	99.12%	93.93%	96.62%	99%	100%
11		Соя min till	2455	1718	737	97.71%	97.23%	99.53%	100%	100%
12		Соя чистая	593	415	178	97.84%	98.99%	97.35%	100%	100%
13		Пшеница	205	143	62	100%	100%	96.60%	100%	100%
14		Леса	1265	885	380	99.77%	99.76%	99.21%	100%	100%
15		Здания-Трава-Деревья-Приводы	386	270	116	95.57%	97.93%	100%	98%	100%
16		Камень-Сталь-Башины	93	65	28	86.36%	98.92%	81.54%	96%	100%

а

№	Цвет	Категория	Образец	Обуче- ние 70 %	Тестиро- вание 30 %	1D CHC	2D CHC	3D CHC	Hybrid: SN	3D-2D CHC
1		Зеленые_сорняки_брокколи_1	46	32	14	100%	100%	100%	100%	100%
2		Зеленые_сорняки_брокколи_2	1428	1000	428	100%	99.92%	100%	98.41%	100%
3		Под паром	830	581	249	100%	99.65%	99.86%	100%	100%
4		Паровой_грубый_плуг	237	166	71	100%	99.78%	99.59%	100%	100%
5		Пар гладкий	483	338	145	99.89%	99.07%	100%	100%	100%
6		Стерня пшеницы	730	511	219	100%	99.97%	100%	98.65%	100%
7		Сельдерей	28	20	8	100%	99.75%	100%	100%	100%
8		Виноград_необученный	478	335	143	99.82%	94.28%	99.96%	100%	100%
9		Почва виноградника	20	14	6	100%	99.97%	100%	100%	100%
10		Кукуруза с зелеными сорняками	972	680	292	99.66%	99.63%	99.97%	100%	100%
11		Салат ромэн 4wk	2455	1718	737	100%	99.91%	98.66%	100%	100%
12		Салат ромэн 5wk	593	415	178	100%	100%	100%	99.90%	100%
13		Салат ромэн 6wk	205	143	62	100%	100%	100%	100%	100%
14		Салат ромэн 7wk	1265	885	380	99.33%	99.91%	100%	100%	100%
15		Виноградные лозы	386	270	116	99.57%	97.40%	99.27%	100%	100%
16		Вертикальная шпалера Винограда	93	65	28	99.92%	100%	100%	100%	100%

б

№	Цвет	Категория	Образец	Обуче- ние 70 %	Тестиро- вание 30 %	1D CHC	2D CHC	3D CHC	Hybrid: SN	3D-2D CHC
1		Асфальт	46	32	14	99.70%	99.27%	99.85%	100%	100%
2		Лука	1428	1000	428	99.95%	99.89%	99.92%	100%	100%
3		Гравий	830	581	249	98.37%	86.79%	97.82%	100%	100%
4		Деревья	237	166	71	99.44%	98.14%	98.55%	100%	100%
5		Окрашенные металлические листы	483	338	145	100%	100%	100%	100%	100%
6		Голая почва	730	511	219	100%	99.86%	98.21%	100%	100%
7		Битум	28	20	8	99.89%	94.52%	99.46%	100%	100%
8		Самоблокирующиеся кирпичи	478	335	143	98.33%	98.29%	99.26%	99%	100%
9		Тени	20	14	6	99.50%	99.85%	98.34%	100%	100%

в

Рис. 5. Сравнение точности методов 1D CHC, 2D CHC, 3D CHC и Hybrid:SN для наборов данных IP (а), SA (б) и PU (в)

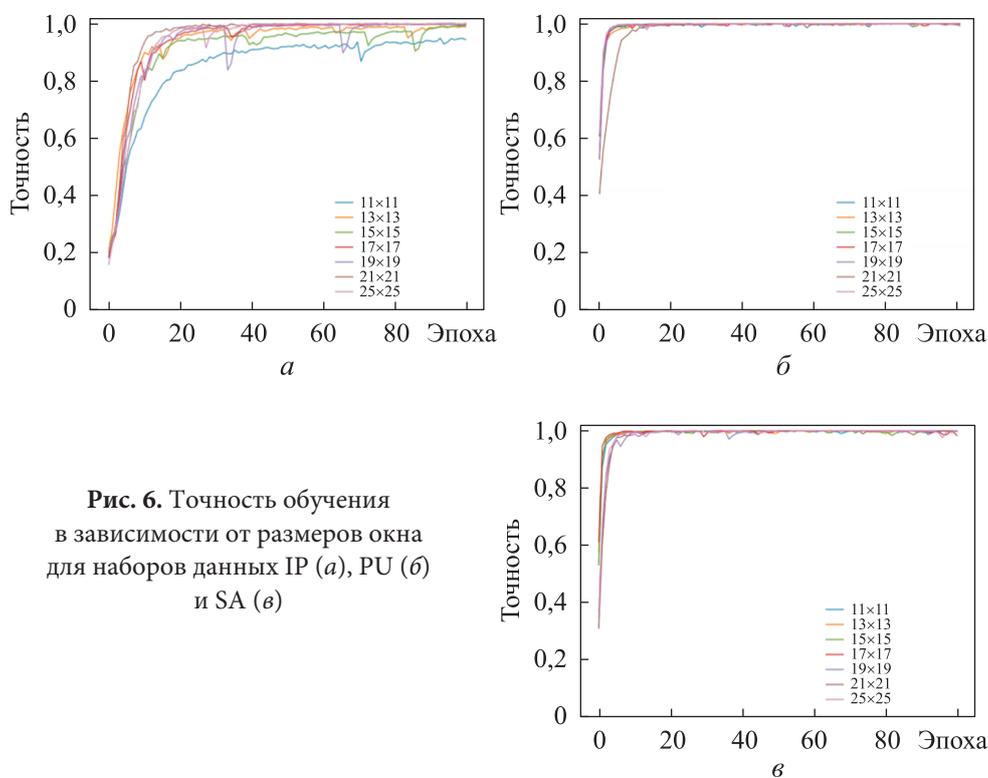


Рис. 6. Точность обучения в зависимости от размеров окна для наборов данных IP (а), PU (б) и SA (в)

Таблица 2

Точность классификации по эталонным трем наборам данных (IP, PU, SA) с использованием предложенной модели и современных методов глубокого обучения для объема обучающих данных 70 и 10 %

Метод	IP			PU			SA		
	ОТ	КК	СТ	ОТ	КК	СТ	ОТ	КК	СТ
<i>Результаты классификации по эталонным трем наборам данных для объема обучающих данных 70 %</i>									
1D CHC	99,18 ± ± 0,0	99,06 ± ± 0,0	98,85 ± ± 0,0	99,66 ± ± 0,0	99,55 ± ± 0,0	99,50 ± ± 0,0	99,88 ± ± 0,0	99,87 ± ± 0,0	99,91 ± ± 0,0
2D CHC	99,76 ± ± 0,2	99,73 ± ± 0,5	99,03 ± ± 0,8	99,49 ± ± 0,2	99,44 ± ± 0,5	99,78 ± ± 0,2	99,90 ± ± 0,2	99,86 ± ± 0,1	99,78 ± ± 0,6
3D CHC	98,40 ± ± 0,0	97,89 ± ± 0,0	97,89 ± ± 0,0	98,34 ± ± 0,0	98,90 ± ± 0,0	98,91 ± ± 0,0	98,01 ± ± 0,0	97,90 ± ± 0,0	97,68 ± ± 0,0
Hybrid:SN	99,75 ± ± 0,1	99,71 ± ± 0,1	99,63 ± ± 0,2	99,98 ± ± 0,0	99,98 ± ± 0,0	99,97 ± ± 0,0	100,0 ± ± 0,0	100,0 ± ± 0,0	100,0 ± ± 0,0
3D-2D CHC	99,86 ± ± 0,1	99,84 ± ± 0,1	99,44 ± ± 0,2	99,99 ± ± 0,0	99,99 ± ± 0,0	99,99 ± ± 0,0	100,0 ± ± 0,0	100,0 ± ± 0,0	100,0 ± ± 0,0

Метод	IP			PU			SA		
	ОТ	КК	СТ	ОТ	КК	СТ	ОТ	КК	СТ
<i>Результаты классификации по эталонным трем наборам данных для объема обучающих данных 10 %</i>									
1D CHC	97,18 ± ± 0,0	-	-	99,28 ± ± 0,0	-	-	99,56 ± ± 0,0	-	-
2D CHC	97,57 ± ± 0,0	97,23 ± ± 0,0	98,46 ± ± 0,0	99,64 ± ± 0,0	99,53 ± ± 0,0	99,74 ± ± 0,0	98,34 ± ± 0,0	98,15 ± ± 0,0	99,33 ± ± 0,0
3D CHC	82,62 ± ± 0,0	79,25 ± ± 0,0	76,51 ± ± 0,0	94,90 ± ± 0,0	94,90 ± ± 0,0	97,03 ± ± 0,0	85,00 ± ± 0,0	83,20 ± ± 0,0	89,63 ± ± 0,0
Hybrid:SN	98,39 ± ± 0,0	98,16 ± ± 0,0	98,01 ± ± 0,0	99,72 ± ± 0,0	99,64 ± ± 0,0	99,20 ± ± 0,0	99,98 ± ± 0,0	99,98 ± ± 0,0	99,98 ± ± 0,0
3D-2D CHC	98,26 ± ± 0,0	98,02 ± ± 0,0	98,03 ± ± 0,0	99,74 ± ± 0,0	99,66 ± ± 0,0	99,27 ± ± 0,0	100,0 ± ± 0,0	100,0 ± ± 0,0	100,0 ± ± 0,0

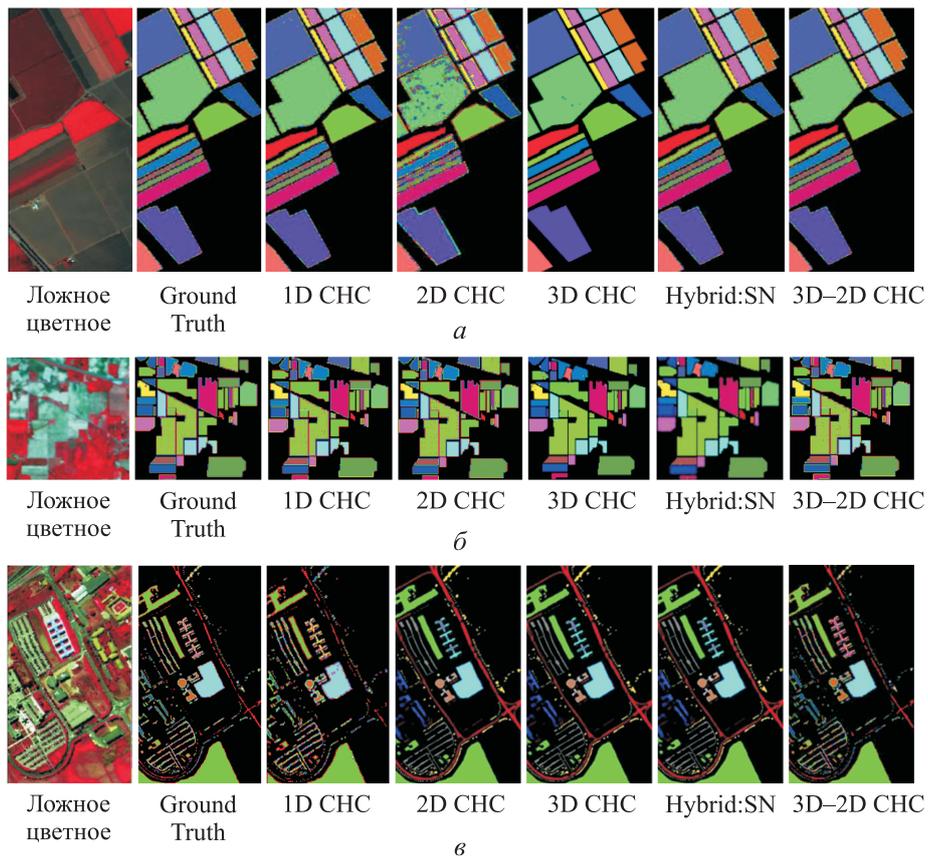


Рис. 7. Классификационные карты для набора данных IP (а), PU (б), SA (в) с использованием различных сверточных нейронных сетей (1D, 2D, 3D, Hybrid:SN, 3D-2D CHC)

В работе реализованы следующие глубокие методы обучения с различными типичными структурами сверточных нейронных сетей (1D, 2D и 3D СНС, Hybrid:SN, модель 3D–2D СНС) и проведено сравнение эффективности вычислительной точности (рис. 8). Установлено, что размер окна 21×21 наиболее удобен для классификации ГСИ по наборам данных.

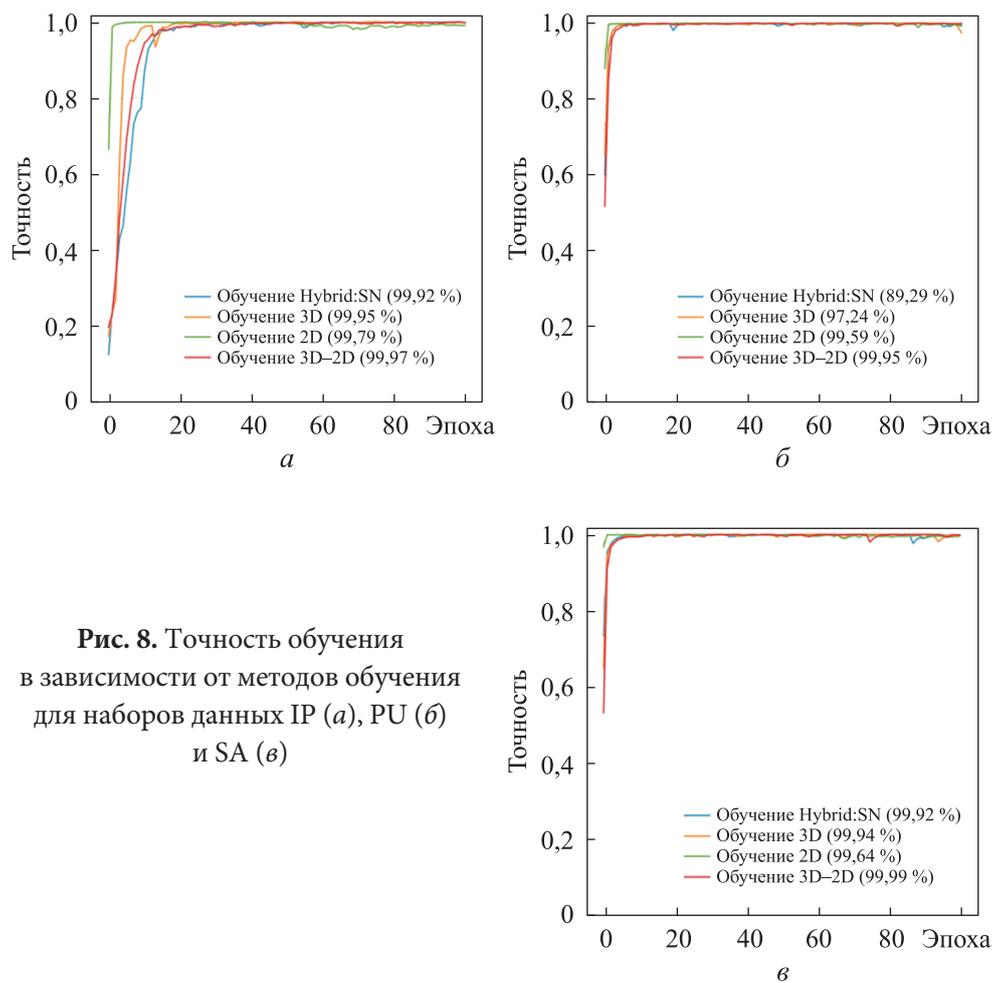


Рис. 8. Точность обучения в зависимости от методов обучения для наборов данных IP (а), PU (б) и SA (в)

Комбинированная модель 3D–2D СНС аналогична Hybrid:SN [11], за исключением глубокой сетевой архитектуры (более трех трехмерных слоев), применено четыре модуля регуляризации отсева для уменьшения переобучения и улучшения обобщения глубоких нейронных сетей. Во время обучения использован модуль максимального объединения для уменьшения числа параметров и размеров данных ГСИ.

Заключение. Основная проблема при использовании методов пространственно-спектрального выделения признаков — возможная высокая размерность векторов признаков, применяемых классификатором.

Введена комбинированная 3D–2D-модель для классификации ГСИ [18]. Предложенная комбинированная модель объединяет комплементарную информацию пространственно-спектрального и спектрального характера в виде 3D- и 2D-сверток. Эксперименты, проведенные над тремя эталонными наборами данных (см. табл. 2), подтверждают превосходство предложенного метода. Согласно изложенному, такая СНС имеет больше преимуществ по сравнению с другими традиционными методами глубокого обучения с позиции визуальных классификационных карт и количественных результатов (см. рис. 7). Предложенная модель 3D–2D СНС более эффективна по вычислительной точности, чем 1D СНС, 2D СНС, 3D СНС и Hybrid:SN.

ЛИТЕРАТУРА

- [1] Camps-Valls G., Tuia D., Bruzzone L., et al. Advances in hyperspectral image classification: Earth monitoring with statistical learning methods. *IEEE Signal Process. Mag.*, 2014, vol. 31, iss. 1, pp. 45–54. DOI: <https://doi.org/10.1109/MSP.2013.2279179>
- [2] Tun N.L., Gavrilov A., Tun N.M., et al. Hyperspectral remote sensing images classification using fully convolutional neural network. *IEEE ElConRus*, 2021. DOI: <https://doi.org/10.1109/ElConRus51938.2021.9396673>
- [3] Liang H., Li Q. Hyperspectral imagery classification using sparse representations of convolutional neural network features. *Remote Sens.*, 2016, vol. 8, no. 2, art. 99. DOI: <https://doi.org/10.3390/rs8020099>
- [4] Tun N.L., Gavrilov A., Tun N.M., et al. Remote sensing data classification using a hybrid pre-trained VGG16 CNN-SVM classifier. *IEEE ElConRus*, 2021. DOI: <https://doi.org/10.1109/ElConRus51938.2021.9396706>
- [5] Tun N.L., Gavrilov A., Tun N.M. Multi-classification of satellite imagery using fully convolutional neural network. *IEEE ICIEAM*, 2020. DOI: <https://doi.org/10.1109/ICIEAM48468.2020.9111928>
- [6] Liu S., Luo H., Tu Y., et al. Wide contextual residual network with active learning for remote sensing image classification. *IGARSS*, 2018. DOI: <https://doi.org/10.1109/IGARSS.2018.8517855>
- [7] Tun N.L., Gavrilov A., Tun N.M. Facial image denoising using convolutional autoencoder network. *IEEE ICIEAM*, 2020. DOI: <https://doi.org/10.1109/ICIEAM48468.2020.9112080>
- [8] Zhang H., Meng L., Wei X., et al. 1D-convolutional capsule network for hyperspectral image classification. *Computer Vision and Pattern Recognition*. URL: <https://arxiv.org/abs/1903.09834>
- [9] Gao Q., Lim S., Jia X. Hyperspectral image classification using convolutional neural networks and multiple feature learning. *Remote Sens.*, 2018, vol. 8, no. 2, art. 299. DOI: <https://doi.org/10.3390/rs10020299>

- [10] Ahmad M. A fast 3D CNN for hyperspectral image classification. *Image and Video Processing*. URL: <https://arxiv.org/abs/2004.14152>
- [11] Roy S.K., Krishna G., Dubey S.R., et al. HybridSN: exploring 3-D–2-D CNN feature hierarchy for hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.*, 2020, vol. 17, iss. 2, pp. 277–281. DOI: <https://doi.org/10.1109/LGRS.2019.2918719>
- [12] He M., Li B., Chen H. Multi-scale 3D deep convolutional neural network for hyperspectral image classification. *IEEE ICIP*, 2017, pp. 3904–3908. DOI: <https://doi.org/10.1109/ICIP.2017.8297014>
- [13] Mou L., Ghamisi P., Zhu X.X. Unsupervised spectral-spatial feature learning via deep residual conv–deconv network for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.*, 2018, vol. 56, iss. 1, pp. 391–406. DOI: <https://doi.org/10.1109/TGRS.2017.2748160>
- [14] Paoletti M.E., Haut J.M., Fernandez-Beltran R., et al. Capsule networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.*, 2019, vol. 57, iss. 4, pp. 2145–2160. DOI: <https://doi.org/10.1109/TGRS.2018.2871782>
- [15] Ji S., Xu W., Yang M., et al. 3D convolutional neural networks for human action recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2013, vol. 35, iss. 1, pp. 221–231. DOI: <https://doi.org/10.1109/TPAMI.2012.59>
- [16] Lv W., Wang X. Overview of hyperspectral image classification. *J. Sens.*, 2020, vol. 2020, art. ID 4817234. DOI: <https://doi.org/10.1155/2020/4817234>
- [17] Song W., Li S., Fang L., et al. Hyperspectral image classification with deep feature fusion network. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, vol. 56, iss. 6, pp. 3173–3184. DOI: <https://doi.org/10.1109/TGRS.2018.2794326>
- [18] Congalton R.G., Mead R.A. A quantitative method to test for consistency and correctness in photointerpretation. *Photogramm. Eng. Remote Sensing*, 1983, vol. 49, no. 1, pp. 69–74.

Ньян Линн Тун — аспирант кафедры «Системы автоматического управления» МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5, стр. 1).

Гаврилов Александр Игоревич — канд. техн. наук, доцент кафедры «Системы автоматического управления» МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5, стр. 1).

До Минь Чиеу — аспирант кафедры «Системы автоматического управления» МГТУ им. Н.Э. Баумана (Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5, стр. 1).

Просьба ссылаться на эту статью следующим образом:

Ньян Л.Т., Гаврилов А.И., До М.Ч. Классификация гиперспектральных данных дистанционного зондирования Земли с использованием комбинированных 3D–2D сверточных нейронных сетей. *Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение*, 2022, № 1 (138), с. 100–118.

DOI: <https://doi.org/10.18698/0236-3933-2022-1-100-118>

CLASSIFICATION OF HYPERSPECTRAL REMOTE EARTH SENSING DATA USING COMBINED 3D–2D CONVOLUTIONAL NEURAL NETWORKS

L.T. Nyan
A.I. Gavrilov
M.T. Do

nyanlin54@gmail.com
alexgavrilov@bmstu.ru
dominhtrieuvhp@gmail.com

Bauman Moscow State Technical University, Moscow, Russian Federation

Abstract

Hyperspectral image classification is used for analyzing remote Earth sensing data. Convolutional neural network is one of the most commonly used methods for processing visual data based on deep learning. The article considers the proposed hybrid 3D–2D spectral convolutional neural network for hyperspectral image classification. At the initial stage, a simple combined trained deep learning model was proposed, which was constructed by combining 2D and 3D convolutional neural networks to extract deeper spatial-spectral features with fewer 3D–2D convolutions. The 3D network facilitates the joint spatial-spectral representation of objects from a stack of spectral bands. Functions of 3D–2D convolutional neural networks were used for classifying hyperspectral images. The algorithm of the method of principal components is applied to reduce the dimension. Hyperspectral image classification experiments were performed on Indian Pines, University of Pavia and Salinas Scene remote sensing datasets. The first layer of the feature map is used as input for subsequent layers in predicting final labels for each hyperspectral pixel. The proposed method not only includes the benefits of advanced feature extraction from convolutional neural networks, but also makes full use of spectral and spatial information. The effectiveness of the proposed method was tested on three reference data sets. The results show that a multifunctional learning system based on such networks significantly improves classification accuracy (more than 99 %)

Keywords

Deep learning, convolutional neural networks, hyperspectral image classification

Received 19.04.2021

Accepted 04.08.2021

© Author(s), 2022

REFERENCES

[1] Camps-Valls G., Tuia D., Bruzzone L., et al. Advances in hyperspectral image classification: Earth monitoring with statistical learning methods. *IEEE Signal Process. Mag.*, 2014, vol. 31, iss. 1, pp. 45–54. DOI: <https://doi.org/10.1109/MSP.2013.2279179>

- [2] Tun N.L., Gavrilov A., Tun N.M., et al. Hyperspectral remote sensing images classification using fully convolutional neural network. *IEEE ElConRus*, 2021.
DOI: <https://doi.org/10.1109/ElConRus51938.2021.9396673>
- [3] Liang H., Li Q. Hyperspectral imagery classification using sparse representations of convolutional neural network features. *Remote Sens.*, 2016, vol. 8, no. 2, art. 99.
DOI: <https://doi.org/10.3390/rs8020099>
- [4] Tun N.L., Gavrilov A., Tun N.M., et al. Remote sensing data classification using a hybrid pre-trained VGG16 CNN-SVM classifier. *IEEE ElConRus*, 2021.
DOI: <https://doi.org/10.1109/ElConRus51938.2021.9396706>
- [5] Tun N.L., Gavrilov A., Tun N.M. Multi-classification of satellite imagery using fully convolutional neural network. *IEEE ICIEAM*, 2020.
DOI: <https://doi.org/10.1109/ICIEAM48468.2020.9111928>
- [6] Liu S., Luo H., Tu Y., et al. Wide contextual residual network with active learning for remote sensing image classification. *IGARSS*, 2018.
DOI: <https://doi.org/10.1109/IGARSS.2018.8517855>
- [7] Tun N.L., Gavrilov A., Tun N.M. Facial image denoising using convolutional autoencoder network. *IEEE ICIEAM*, 2020.
DOI: <https://doi.org/10.1109/ICIEAM48468.2020.9112080>
- [8] Zhang H., Meng L., Wei X., et al. 1D-convolutional capsule network for hyperspectral image classification. *Computer Vision and Pattern Recognition*.
Available at: <https://arxiv.org/abs/1903.09834>
- [9] Gao Q., Lim S., Jia X. Hyperspectral image classification using convolutional neural networks and multiple feature learning. *Remote Sens.*, 2018, vol. 8, no. 2, art. 299.
DOI: <https://doi.org/10.3390/rs10020299>
- [10] Ahmad M. A fast 3D CNN for hyperspectral image classification. *Image and Video Processing*. Available at: <https://arxiv.org/abs/2004.14152>
- [11] Roy S.K., Krishna G., Dubey S.R., et al. HybridSN: exploring 3-D-2-D CNN feature hierarchy for hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.*, 2020, vol. 17, iss. 2, pp. 277–281. DOI: <https://doi.org/10.1109/LGRS.2019.2918719>
- [12] He M., Li B., Chen H. Multi-scale 3D deep convolutional neural network for hyperspectral image classification. *IEEE ICIP*, 2017, pp. 3904–3908.
DOI: <https://doi.org/10.1109/ICIP.2017.8297014>
- [13] Mou L., Ghamisi P., Zhu X.X. Unsupervised spectral-spatial feature learning via deep residual conv-deconv network for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.*, 2018, vol. 56, iss. 1, pp. 391–406.
DOI: <https://doi.org/10.1109/TGRS.2017.2748160>
- [14] Paoletti M.E., Haut J.M., Fernandez-Beltran R., et al. Capsule networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.*, 2019, vol. 57, iss. 4, pp. 2145–2160. DOI: <https://doi.org/10.1109/TGRS.2018.2871782>
- [15] Ji S., Xu W., Yang M., et al. 3D convolutional neural networks for human action recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2013, vol. 35, iss. 1, pp. 221–231.
DOI: <https://doi.org/10.1109/TPAMI.2012.59>

- [16] Lv W., Wang X. Overview of hyperspectral image classification. *J. Sens.*, 2020, vol. 2020, art. ID 4817234. DOI: <https://doi.org/10.1155/2020/4817234>
- [17] Song W., Li S., Fang L., et al. Hyperspectral image classification with deep feature fusion network. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, vol. 56, iss. 6, pp. 3173–3184. DOI: <https://doi.org/10.1109/TGRS.2018.2794326>
- [18] Congalton R.G., Mead R.A. A quantitative method to test for consistency and correctness in photointerpretation. *Photogramm. Eng. Remote Sensing*, 1983, vol. 49, no. 1, pp. 69–74.

Nyan L.T. — Post-Graduate Student, Department of Automatic Control Systems, Bauman Moscow State Technical University (2-ya Baumanskaya ul. 5, str. 1, Moscow, 105005 Russian Federation).

Gavrilov A.I. — Cand. Sc. (Eng.), Assoc. Professor, Department of Automatic Control Systems, Bauman Moscow State Technical University (2-ya Baumanskaya ul. 5, str. 1, Moscow, 105005 Russian Federation).

Do M.T. — Post-Graduate Student, Department of Automatic Control Systems, Bauman Moscow State Technical University (2-ya Baumanskaya ul. 5, str. 1, Moscow, 105005 Russian Federation).

Please cite this article in English as:

Nyan L.T., Gavrilov A.I., Do M.T. Classification of hyperspectral remote Earth sensing data using combined 3D–2D convolutional neural networks. *Herald of the Bauman Moscow State Technical University, Series Instrument Engineering*, 2022, no. 1 (138), pp. 100–118 (in Russ.). DOI: <https://doi.org/10.18698/0236-3933-2022-1-100-118>