Article

Lithology Classification Using TASI Thermal Infrared Hyperspectral Data with Convolutional Neural Networks

Huize Liu [1], Ke Wu [1,*], Honggen Xu [2] and Ying Xu [3,4]

1   Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan 430074, China; huizeliu@cug.edu.cn
2   Changsha Center of Natural Resources Comprehensive Survey, China Geological Survey, Changsha 410699, China; xuhonggen@mail.cgs.gov.cn
3   National Satellite Ocean Application Service, Ministry of Natural Resources, Beijing 100081, China; xuying@mail.nsoas.org.cn
4   Key Laboratory of Space Ocean Remote Sensing and Application, Ministry of Natural Resources, Beijing 100081, China
*   Correspondence: kewu@cug.edu.cn; Tel.: +86-136-3862-2685

Abstract: In recent decades, lithological mapping techniques using hyperspectral remotely sensed imagery have developed rapidly. The processing chains using visible-near infrared (VNIR) and shortwave infrared (SWIR) hyperspectral data are proven to be available in practice. The thermal infrared (TIR) portion of the electromagnetic spectrum has considerable potential for mineral and lithology mapping. In particular, the abovementioned rocks at wavelengths of 8–12 μm were found to be discriminative, which can be seen as a characteristic to apply to lithology classification. Moreover, it was found that most of the lithology mapping and classification for hyperspectral thermal infrared data are still carried out by traditional spectral matching methods, which are not very reliable due to the complex diversity of geological lithology. In recent years, deep learning has made great achievements in hyperspectral imagery classification feature extraction. It usually captures abstract features through a multilayer network, especially convolutional neural networks (CNNs), which have received more attention due to their unique advantages. Hence, in this paper, lithology classification with CNNs was tested on thermal infrared hyperspectral data using a Thermal Airborne Spectrographic Imager (TASI) at three small sites in Liuyuan, Gansu Province, China. Three different CNN algorithms, including one-dimensional CNN (1-D CNN), two-dimensional CNN (2-D CNN) and three-dimensional CNN (3-D CNN), were implemented and compared to the six relevant state-of-the-art methods. At the three sites, the maximum overall accuracy (OA) based on CNNs was 94.70%, 96.47% and 98.56%, representing improvements of 22.58%, 25.93% and 16.88% over the worst OA. Meanwhile, the average accuracy of all classes (AA) and kappa coefficient (kappa) value were consistent with the OA, which confirmed that the focal method effectively improved accuracy and outperformed other methods.

Keywords: hyperspectral; thermal infrared remote sensing; convolutional neural networks; TASI

1. Introduction

Hyperspectral remote sensing imagery has an extremely high spectral resolution and has been widely applied in lithological geological exploration, classification and mapping [1,2]. Therefore, it is very beneficial to mine physicochemical characteristics or subtle recognition of minerals and rocks by using hyperspectral remote sensing technology. In general, the spectra at visible and near-infrared (VNIR; 0.4–1.1 μm) and shortwave infrared (SWIR; 1.1–2.5 μm) [3] wavelengths are successfully utilized for mineral mapping and rock classification [4]. Different types of minerals and rocks have different chemical compositions and crystalline structures that control the shape of the spectral curve and the presence and positions of specific absorption bands [5]. For example, in the ferrous and

ferric minerals spectrum, the strong absorption in the VNIR range is caused by ferric iron; hydroxyl minerals show significant absorption features in the SWIR spectral properties [6,7]. In contrast to these examples, orthoclase, a dominant mineral in granite, shows almost no significant absorption features in the VNIR to SWIR spectral range. Simultaneously, studies show that no more than 50% of the minerals can be identified within the VNIR/SWIR range, which leads to inaccuracy in lithology mapping and classification [8]. Attention has shifted to the thermal infrared (TIR; 8–12 μm) spectrum. TIR can show some advantages over other spectral bands in the classification of rock-forming minerals due to its distinctive characteristics. In this spectrum, features are dominated by their radiant energy for some minerals, reflecting their characteristic properties. Some typical minerals have their spectral characteristics in the thermal infrared spectrum. For example, olivines, pyroxenes and quartz [1,9,10] have typical absorption features at TIR wavelengths, which can be seen as exclusive characteristics to be applied [3,7]; silicate minerals are responsible for the absorption bands near 10 μm due to fundamental Si-O stretching vibrations; in the spectrum of carbonates, the strong features at 11.4 and 14.3 μm are caused by C-O bending [11]. Therefore, characterizing, modeling and analyzing the lithology and minerals from the TIR remotely sensed images is very meaningful.

In addition, remote sensing sensors can play a big role in lithology classification. The majority of geological studies based on TIR remote sensing data have focused on multi-spectral data. Some typical sensors, such as the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) [12–14] and Thermal Infrared Multispectral Scanner (TIMS) [9,11], have been successfully used to identify lithologic units based on spectral indices of quartz, carbonate and so on [12,14]. With the development of technology, thermal infrared hyperspectral data can be obtained from airborne platforms, especially Thermal Airborne Spectrographic Imagery (TASI). Cui et al., using airborne hyperspectral TASI data, illustrated the exceptional potential for mapping quartz, calcite, diopside, hornblende and microcline [9]. Black et al. applied the traditional hyperspectral remote sensing lithology mapping process to TASI remote sensing images in the Antarctic region and successfully extracted three lithologies of granite, granodiorite and diabase [11]. Great progress has been made in the past several years, and the superiority of TASI has been proven.

The methods of classifying rocks and minerals with TASI images are currently focused on traditional methods, which can have some drawbacks. Traditional lithology classification algorithms are mainly aimed at the spectra, and they can be grouped into two aspects: based on spectral similarity and spectral characteristics [15–17]. To the former, the main idea of the spectral comparison is to construct a spectral similarity measure to accomplish the lithology classification. For example, spectral angle mapping (SAM) [18,19] and spectral information divergence (SID) [20] are both effective in similarity measurement when the spectral vector is used for direct retrieval. However, this type of method tends to focus on the overall waveform features of the spectrum and ignores some of the detailed features, which makes it difficult to classify similar rock types. For the latter, several common absorption characteristics, including absorption depth, absorption width, absorption area, absorption position, absorption symmetry and spectral absorption index, can be defined to identify rocks and minerals [21,22]. However, the spectral absorption position features are easily affected due to the complex chemical composition of different minerals. In addition, a spectrum may be a collection of multiple substances due to the mixed pixel problem. Fully constrained linear spectral unmixing (FCLSU) is proposed to resolve this problem, working on the inversion of mineral abundances [23]. However, it is usually challenging to obtain the accuracy of the mineral type of rocks, which makes adequate capture of detail in the resulting lithology classification map challenging. In general, these algorithms perform classification or recognition by extracting the shallow features of each pixel, i.e., spectral features, without considering the deeper features.

With improved technology, machine learning algorithms (MLAs) have also been applied to mineral classification in hyperspectral remote sensing to achieve good results [23]. Similar to human brains, classifiers such as support vector machine (SVM) [24,25], random

forest (RF) [23] and neural network (NN) [26] all have multiple layers to perform well in tasks. Deep learning has recently emerged as the state-of-the-art machine learning technique with great potential for hyperspectral image classification [27]. Deep learning techniques can automatically learn hierarchical features (from low to high levels) from input data rather than relying on shallow and artificially engineered features [28]. There are a few deep learning models in the literature, including deep belief network (DBN) [29], stacked autoencoder (SAE) [30] and convolutional neural networks (CNNs) [31,32]. Compared to other deep learning models, CNNs were found to be a good alternative to other deep learning models in classification and detection [33]. The network architecture of one-dimensional CNN (1-D CNN) is designed to use the pixel vector along the radiometric dimension as a training sample to extract deep features [34]. Furthermore, two-dimensional CNN (2-D CNN) and three-dimensional CNN (3-D CNN) are both proposed based on the original 1-D CNN to further improve the performance. The 2-D CNN extracts the local spatial features of each pixel [33]. The 3-D CNN can extract spatial features and spectral features simultaneously. The use of explicit spatial and spectral information allows accurate predictions to be made [35].

Based on the above considerations, CNNs are introduced into the lithology classification with TIR hyperspectral data in this paper. Our work focused on applying typical CNN models to finish the lithology classification map with TASI collected at three small sites in Liuyuan, Gansu Province, China. Liuyuan is a mining town with rich mineral resources. The bedrock on the land surface is bare without vegetation cover, which is suitable for lithology classification work. Six state-of-the-art methods, including SAM, SID, FCLSU, SVM, RF and NN, were implemented to make comparisons with the performance of the CNNs in three images. Experiments have proven that CNNs significantly improve the overall classification accuracy compared with other methods, which shows that CNNs have great advantages.

The sections in this article are organized as follows. The data are described in Section 2. The classification methodology is described in Section 3. The classification analysis for these data are provided in Section 4. The discussion is described in Section 5. Finally, the conclusions are given in Section 6.

2. Data Description

2.1. Study Area

The study area is located near Huaniu Mountains, Liuyuan Town, Guazhou County, Jiuquan City, Gansu Province in Figure 1. Liuyuan is a mining town with rich mineral resources and a typical continental climate. The land surface of the study area is an arid bare stone desert [9]. The study area is in the Yujingzi and Liuyuan intracontinental rift zones on the southern margin of the Beishan active epicontinental belt, which is between the Tarim and SinoKorean plates [36]. The geological map of the study area is shown in Figure 2. The strata developed in the area include the Ordovician Huaniu Mountains Group and the Sinian Xichangjing Group. The lithology of the middle formation of the Ordovician Huaniu Mountains group is dominated by basalt, with slate and marble. The lithology of the second rock group of the Sinian Xichangjing Group is sericite phyllite. The lithology of the fourth rock group of the Sinian Xichangjing Group is slate and hornfels. In addition, due to the frequent magmatic activities and strong intrusive activities in the area, the intrusive rocks of Variscan that are granodiorite and Indosinian that are syenogranite are exposed.
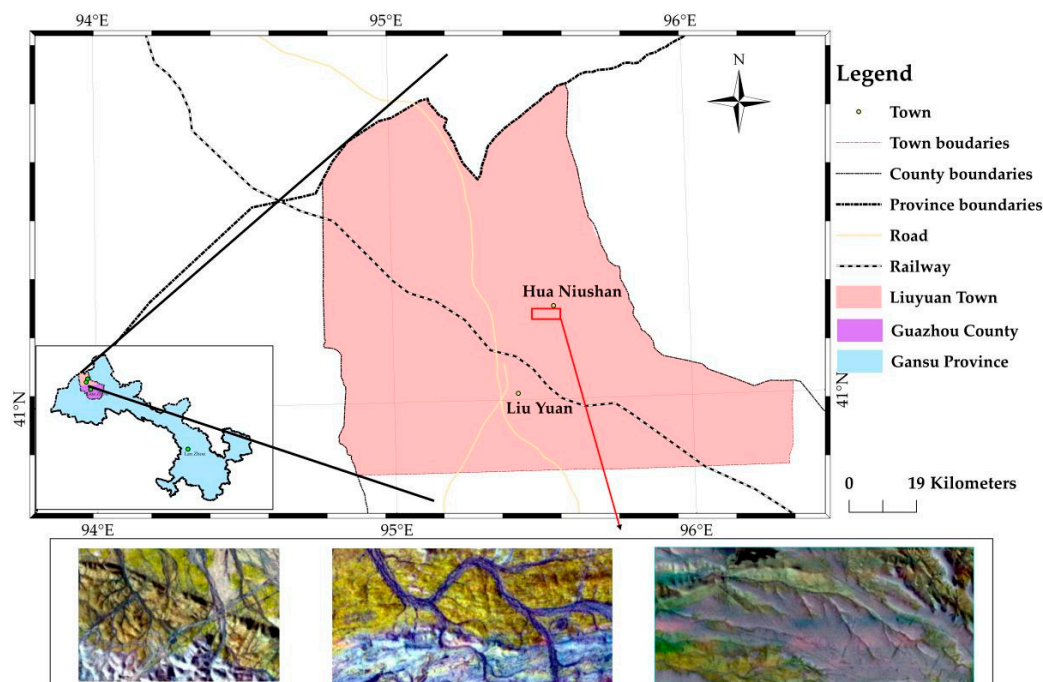
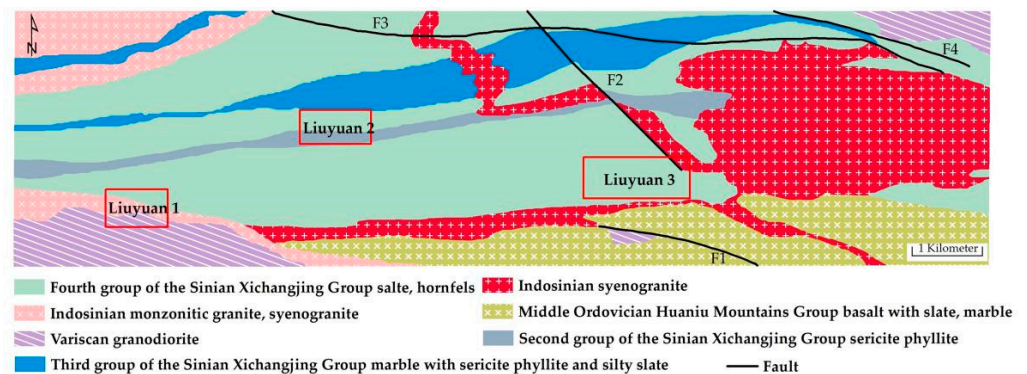**Figure 1.** Location of the study area and experimental data.

Figure 2. Geological map of the study area.

Airborne hyperspectral TIR imagery was acquired in September 2010 between 17:00–18:00 by the TASI sensor with a height of 2 km. TASI was developed by ITRES Canada, which sets 32 channels in the 8–11.5 μm band range, with a band interval of 0.1095 μm and a half-height width of 0.0548 μm. The sensor's total field of view angle is 40°, and the spatial resolution is 2.25 m. The data have the advantages of mobility and flexibility and high spatial and spectral resolution. TASI ensures that each pixel of the image is a clear and independent spatial sample, which has a broader application in mineral mapping. Three small sites of Liuyuan are chosen, and the false-color composite (R: 11.449 μm, G: 10.354 μm, B: 9.914 μm) TIR images are shown in Figure 1. The imagery of the three study areas is named Liuyuan 1, Liuyuan 2 and Liuyuan 3. The sizes of the three images are 270×407, 240×495, 300×700, respectively. The lithological boundaries in this study area are obvious, and the lithological types are easily identified. The main types are slate, syenogranite and diorite, according to the historical geological map.

2.2. Data Preprocessing

The pretreatment of TASI data can be divided into three parts: 1. radiometric calibration; 2. atmospheric correction; and 3. separation of temperature and emissivity.

Radiometric calibration was performed automatically by TASI data processing software. Atmospheric correction and separation of temperature and emissivity are described in the following sections.

2.2.1. Atmospheric Correction

For TASI airborne thermal infrared hyperspectral data, atmospheric correction, i.e., elimination of atmospheric absorption and atmospheric upwelling radiation effects, is the first requirement. The simultaneous atmospheric temperature and humidity profile acquisition experiment was not planned on the day of the aerial photograph; therefore, atmospheric reanalysis data were used. Atmospheric reanalysis data assimilate a large amount of satellite information and conventional observations such as ground and upper air with the advantages of long time series and high spatial resolution, which can be used not only for diagnostic analysis of weather and climate but also as a driving field for weather and climate models. This study used data from the fifth generation of the reanalysis product (ERA5) in November 2016, which can be downloaded free of charge from the European Centre for Medium-Range Weather Forecasts (ECMWF) official website (https://cds.climate.copernicus.eu, accessed on 17 July 2021). The spatial resolution of ERA5 is 25 km, and the temporal resolution is 1 h. The ERA5 products were resampled to the same spatial scale as TASI. The atmospheric reanalysis data were acquired only every hour on the hour, so the simultaneous atmospheric parameters were acquired according to the specific temporal and spatial coordinates of the TASI images taken. The raw ERA5 atmospheric profile data obtained were different in time and space, and the interpolated profiles were entered as parameters. Finally, the atmospheric parameters were simulated by a moderate spectral resolution atmospheric radiative transfer model (MODTRAN) to extract the atmospheric transmittance, atmospheric upward radiation L_{up} and atmospheric downward radiation L_{down}.

2.2.2. Temperature Emissivity Separation

The TASI thermal infrared hyperspectral sensor has a total of 32 thermal infrared channels. Considering that some channels are located in the non-atmospheric window and are greatly affected by the atmosphere, some channels have a large signal-to-noise ratio. To ensure the inversion accuracy, the wavelength range was used to select the channels, and channels with wavelengths less than 8.5 μm and wavelengths greater than 11 μm were eliminated. Finally, 22 channels from channel 6 to 27 were selected. The separation of land surface temperature and emissivity was mainly divided into three steps, which are realized by three different modules: normalized emissivity module (NEM), emissivity ratio module (RATIO) and average/maximum–minimum difference module (MMD) [37]. The NEM module is a preliminary estimate of the target surface temperature and removes the reflected atmospheric radiation from the surface radiance observations. The equations used in this model are as follows:

$$T_b = B_b^{-1}(R_b/\varepsilon_{max}) = \frac{c_2}{\lambda_b}(ln(\frac{c_1\varepsilon_{max}}{\pi R_b \lambda_b^5} + 1))^{-1} \tag{1}$$

$$T_{NEM} = max(T_b) \tag{2}$$

$$\varepsilon_b = \frac{R_b}{B_b(T_{NEM})} \tag{3}$$

where T_b is the estimated surface temperature of channel b (b = 10, 14), B_b is the Planck function, c_1 and c_2 are constants in the Planck function, λ_b is the central wavelength of channel b, and R_b is the ground object emission radiation after the second channel removes the reflected atmospheric down radiation. T_{NEM} is the maximum surface temperature estimated by the NEM module. The RATIO module calculates the emissivity ratio β_b using

the emissivity estimated by the NEM module, which was first proposed by Watson [38]. β_b is calculated using Equation (4).

$$\beta_b = \frac{\varepsilon_b}{\frac{1}{N}\sum \varepsilon_b} \tag{4}$$

The MMD module further estimates surface emissivity and temperature and reveals that there is a certain statistical relationship between the maximum and minimum emissivity. The MMD index can be estimated from the emissivity ratio β_b:

$$MMD = max(\beta_b) - min(\beta_b) \quad b = 6,\ldots\ldots, 27 \tag{5}$$

After the temperature emissivity separation, three emissivity images of the study area are shown in Figure 3a–c.
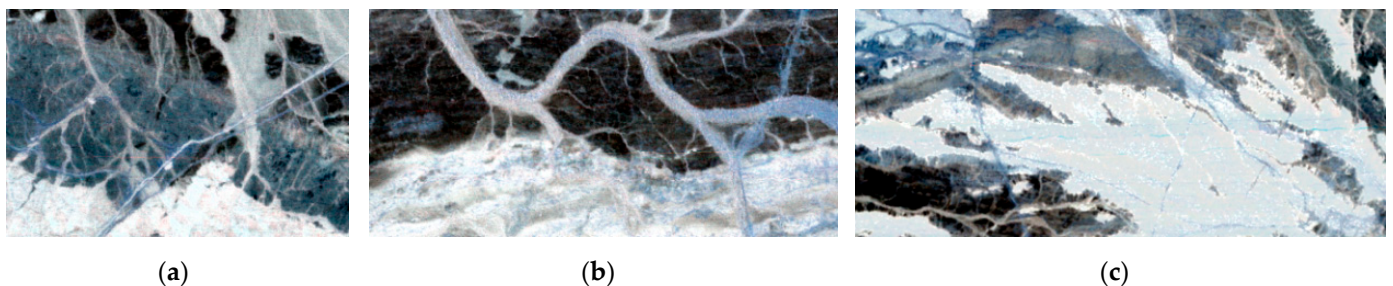

Figure 3. Emissivity image of the study area (R: 8.602 μm, G: 8.711 μm, B: 8.821 μm). (**a**) Liuyuan 1; (**b**) Liuyuan 2; (**c**) Liuyuan 3.

2.3. Reference Map Generation

The study utilized the TASI hyperspectral datasets for lithological identification using CNNs. The proposed methods were designed in a supervised way and, thus, are dependent on the availability of prior knowledge and manual assistance in real applications. During the process, a high spatial resolution reference map was critical for better training data sampling for CNN modeling. In general, conventional geological maps can be regarded as references [23,39,40]. However, they are usually prepared at a large scale and may not be suitable in this study due to the high spatial resolution of TASI. Therefore, the field surveying data, historical geological map and spectral comparison were all used to prepare a reference map. In addition, the minimum noise fraction (MNF) transformation is applied to the emissivity data to demarcate the lithological boundary and help to confirm the type of lithological unit. The process of reference map generation is discussed in the following subsections.

#### 2.3.1. MNF Transformation

MNF can be seen as a character transformation method used in remote sensing images. It is used to determine the intrinsic dimensionality (i.e., number of bands) of the image data and to separate the noise from the data. The MNF transform can arrange the principal components after the transformation according to the image quality. The band combinations of MNF (R: MNF1, G: MNF2, B: MNF3) are shown in Figure 4. It is worth noting that the MNF image shows higher spectral contrast, which effectively discriminates various lithological units of the area. The image pixels of the same color can be grouped together according to the results of MNF transformation, and then the corresponding boundaries can be effectively drawn. Therefore, the MNF transformation confirms the boundaries of the lithology.
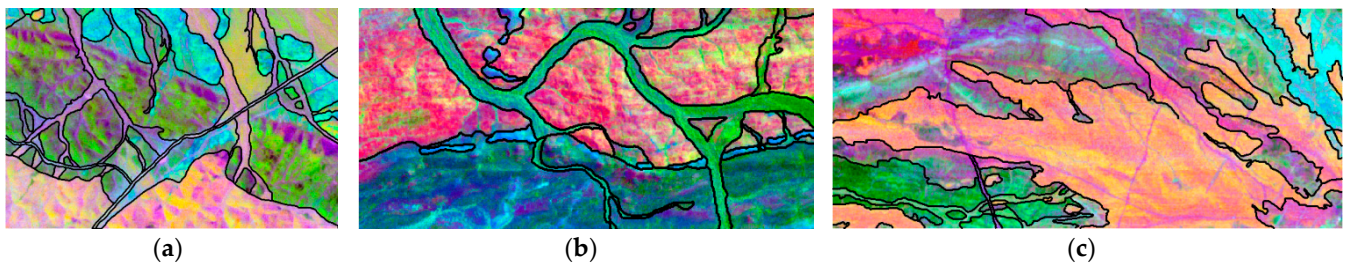
Figure 4. Minimum noise fraction (MNF) transformation image of the study area (R: MNF1, G: MNF2, B: MNF3). (**a**) Liuyuan 1; (**b**) Liuyuan 2; (**c**) Liuyuan 3.

2.3.2. Field Surveying Data

The lithological unit types of different colors can be determined by the field data and geological map. The field surveying locations are displayed in the three study areas. There are 15 positions, which are labeled 1 to 8, 1 to 4 and 1 to 3 for Liuyuan 1, Liuyuan 2 and Liuyuan 3 in Figure 5a–c, respectively. Ground TIR emissivity spectra were acquired using a 102F portable Fourier transform infrared spectrometer (FTIR) under cloud-free conditions at the end of August 2020. The spectrum range is 8–12 μm with a sampling interval of 1 nm. Moreover, the samples were collected and brought to the laboratory for indoor spectroscopic measurements. After the measurement, the lithological types of the positions are determined. In Figure 5, red is the slate, green is the syenogranite and blue is the diorite. The primary contents of the 15 measured positions are described in Table 1. Some of the measured locations were not subjected to spectrum field measurements, and only indoor spectra of the corresponding samples were analyzed. In contrast, some samples were subjected to field spectroscopic measurements, and indoor spectroscopic measurements were not carried out, as the measured lithological samples were more difficult to obtain. For Liuyuan 1, the field measurement spectra were encoded according to different positions, such as LY1_1~LY1_7. Due to the absence of position 6, there are 7 encoded spectra. Moreover, the indoor measuring spectra were also encoded in the same way, such as LY1_YP_1~LY1_YP_6. Due to the absence of positions 2 and 5, there are 6 encoded spectra. For Liuyuan 2 and 3, there is only 1 encoded spectrum of the field measurement at positions 1 and 3, respectively. All of the spectra can be collected and encoded by indoor measurements. Some mineral compositions were determined by X-ray diffraction (XRD) analyses.
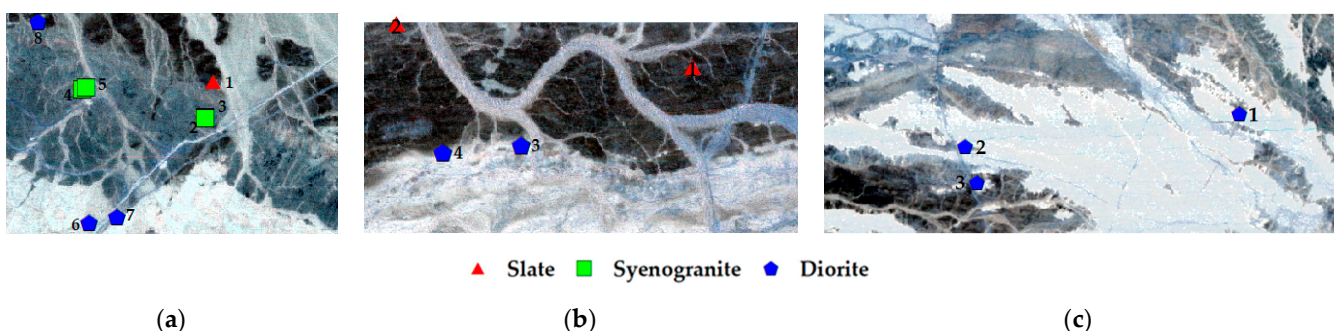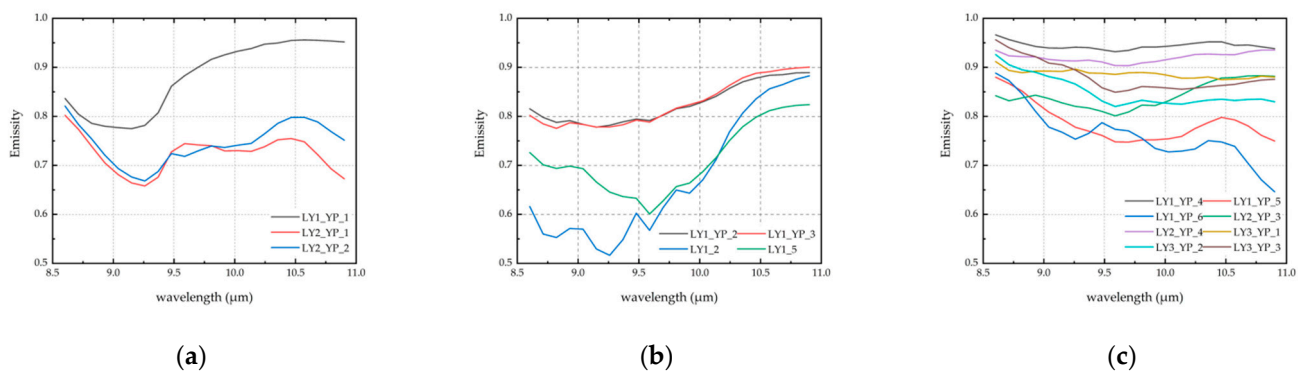

Figure 5. Measured positions. (**a**) Liuyuan 1; (**b**)Liuyuan 2; (**c**) Liuyuan 3.

Table 1. Information of measured positions.

Emissivity Image	Measured Positions	Lithologic Types	Compositions	Field Measured Spectrum	Indoor Measured Spectrum
Liuyuan 1	1	Slate detritus	—	LY1_1	LY1_YP_1
	2	Syenogranite	—	LY1_2	—
	3	Syenogranite detritus	Orthoclase (38.7%) + quartz (27.1%) + biotite (18.9%) + monoclinic pyroxene (15.3%)	LY1_3	LY1_YP_2
	4	Syenogranite detritus	—	LY1_4	LY1_YP_3
	5	Syenogranite	—	LY1_5	—
	6	Diorite	—	—	LY1_YP_4
	7	Diorite	Albite (32.1%) + hornblende (17.0%) + diopside (17%) + quartz (15.3%) + montmorillonite (10.5%) + biotite (8.1%)	LY1_6	LY1_YP_5
	8	Diorite	Chlorite (41.6%) + albite (30.1%) + quartz (13.2%) + biotite (7.6%)+ diopside (7.4%)	LY1_7	LY1_YP_6
Liuyuan 2	1	Slate	Quartz (49.7%) + anorthite (36.9%) + seraphinite (14.7%)	LY2_1	LY2_YP_1
	2	Slate	—	—	LY2_YP_2
	3	Diorite	—	—	LY2_YP_3
	4	Diorite	Albite (51.3%) + quartz (24.6%) + biotite (24.1%)	—	LY2_YP_4
Liuyuan 3	1	Diorite	Orthoclase (51.9%) + magnesia hornblende (39.3%) + albite (14.7%) + biotite (13.9%) + pyroxene (13.1%) + quartz (9.1%)	—	LY3_YP_1
	2	Diorite	—	—	LY3_YP_2
	3	Diorite	Magnesia hornblende (48%) + albite (40.2%) + bixbyite (6.0%) + quartz (5.8%)	LY3_1	LY3_YP_3

The spectra measured by field and indoor measurements according to different lithology types are shown in Figure 6. The spectra for slate, syenogranite and diorite all have different spectral features. In Figure 6a, the spectra of slate have a distinct emission valley between 8.5 and 9.5 μm. Figure 6b has a continuous broad emission valley between 8.5 and 10 μm. As shown in Figure 6c, the spectra form a weak emission valley near 9.5 μm. There is likely a difference in the chemical composition of these samples to lead to the different spectra. For example, LY1_YP_6 and LY3_YP_3 are the spectra of diorite, but their spectral curves are very different.

Figure 6. Spectrum of measured points at three small sites in Liuyuan. (**a**) Slate; (**b**) syenogranite; (**c**) diorite.

2.3.3. Confirmation of Lithology Type

Combining the information of the field positions and the geological map, the main lithology of the study areas can be determined. There is slate, granite and diorite in Liuyuan 1 and slate, diorite and sericite phyllite in Liuyuan 2. For Liuyuan 3, there is much slate according to the geologic map. Even so, there are multiple colors on the MNF transformed image of Liuyuan 3, indicating that the lithology is not unique at the site. This means that there are other lithological unit types in addition to slate. To discriminate the unknown lithologic types in Liuyuan 3, a spectrum analysis model, such as the SAM, was implemented to compare and analyze the known spectrum and unidentified spectrum. The unidentified regions of interest (ROIs) spectra were selected for comparison with the known field-measured spectra. The lithology types were confirmed using the highest score with the measured spectrum. ROIs chosen in red, green and blue are shown as ROI #1, ROI #2 and ROI #3, respectively, in Figure 7. The comparison of lithologic spectral curves between the measured known spectrum and unidentified ROI spectrum is shown in Figure 8. Through spectral comparison analysis, it is determined that ROI #1 is slate, ROI #2 is syenogranite, and ROI #3 is diorite. In addition, there were many alluvial sediments in Figure 4a–c, representing Quaternary sediments, which can be determined by the Google Earth map.
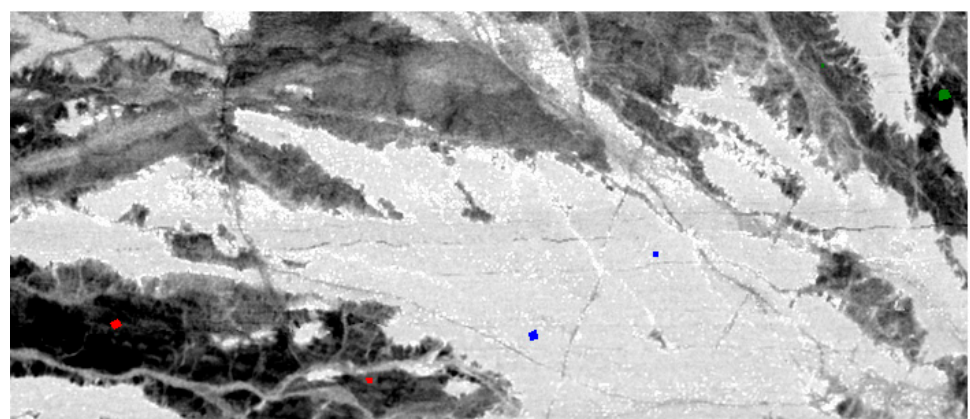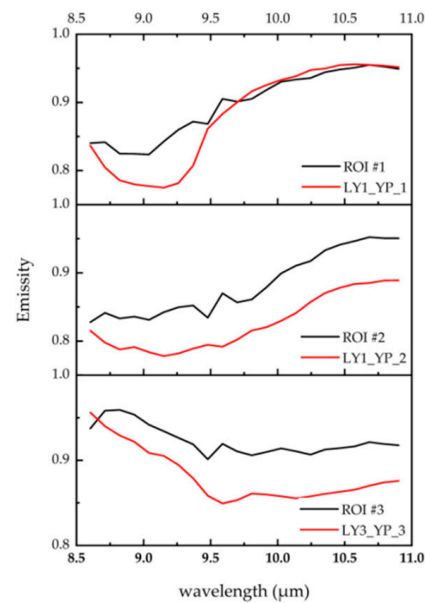

Figure 7. The ROI image of Liuyuan 3.

Figure 8. Comparison of lithologic spectral curves of Liuyuan 3.

Based on the above information, the classification reference maps for Liuyuan 1, 2 and 3 are drawn in Table 2. Slate is red. Syenogranite is green. Diorite is blue. Quaternary sediment is yellow, and sericite phyllite is cyan. Meanwhile, there is some uncertainty labelled black in the reference map.

Table 2. Composition of lithologic categories.

Emissivity Image	Lithologic Types	Classification Reference Map	Category Label
Liuyuan 1	Slate Syenogranite Diorite Quaternary sediment Unclassified		
Liuyuan 2	Slate Diorite Quaternary sediment Sericite phyllite Unclassified		
Liuyuan 3	Slate Syenogranite Diorite Quaternary sediment Unclassified		

3. Convolutional Neural Networks

In recent times, CNNs have achieved outstanding achievements in the analysis of hyperspectral remote sensing images [41]. As shown in Figure 9, the fundamental structure of a CNN is constructed as a set of layers, consisting of a convolutional layer, pooling layer and fully connected layer. The main task of hyperspectral image processing of CNN architecture is handling spatial and spectral information via adjacent layers [34]. The convolutional layer is used to process data by a convolutional operation, aiming to extract features of different dimensions of the input data. A nonlinear mapping layer such as ReLU uses a nonlinear function to map the results of the convolution layer to

increase the nonlinear characteristics of the system and usually follows a convolution layer. The pooling operation follows the convolution layer to streamline the convolution information. The pooling operation reduces the probability of overfitting by reducing the dimensionality of the target features by calculating the mean or extreme value of some regions [42]. A commonly used pooling operation is the max-pooling operation that calculates the maximum of a local patch of units into a single feature map. The last few layers of a convolutional neural network are usually fully connected layers, which helps to better aggregate the information conveyed at lower levels and make final decisions [43].
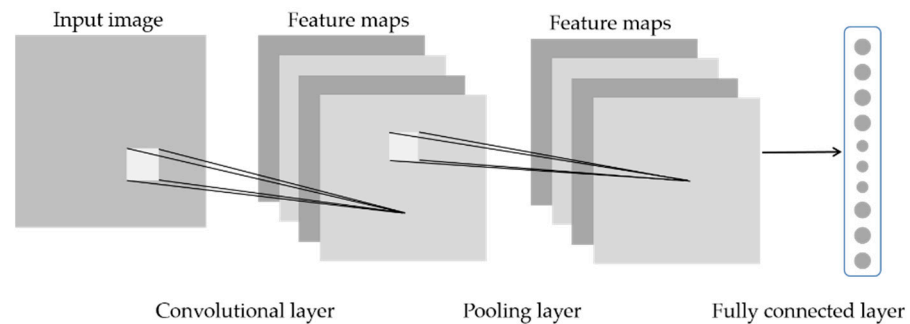
Figure 9. Fundamental structure of the convolutional neural network (CNN).

3.1. One-Dimensional CNN

Let the hyperspectral data cube be denoted by $\mathbf{I} \in R^{M \times N \times D}$, where $\mathbf{I}$ is the original input, M is the height, N is the width, and D is the number of spectral bands. Every HSI pixel in $\mathbf{I}$ contains D spectral measures and forms a one-hot label vector $Y = (y_1, y_2 \ldots, y_C)$ $\in R^{1 \times 1 \times C}$, where C represents the categories [44]. However, the convolution of hundreds of spectral bands of hyperspectral images leads to overfitting due to the increase in operation cost. To address this problem, the input data were processed by principal component analysis (PCA) in the spectral dimension, which greatly reduced the redundancy within hyperspectral data [45]. PCA reduces the number of spectral bands from D to B while maintaining the same spatial dimensions. The PCA reduced data cube is denoted by $\mathbf{X} \in R^{M \times N \times B}$, where $\mathbf{X}$ is the modified input after PCA. The labeled pixels are selected from $\mathbf{X}$ and are fed into input layers as training samples and then convoluted by kernels in convolutional layers [44].

The convolutional layer is introduced first. The value of a neuron v_{ij}^x at position x of the jth feature map in the ith layer is denoted as follows [33]:

$$v_{ij}^x = g\left(b_{ij} + \sum_m \sum_{p=0}^{P_i-1} w_{ijm}^p v_{(i-1)m}^{x+p} \right) \tag{6}$$

where m indexes the feature map in the previous layer $((i-1)$ th layer) connected to the current feature map, w_{ijm}^p is the weight of position p connected to the mth feature map, Pi is the width of the kernel towards the spectral dimension, and b_{ij} is the bias of jth feature map in the ith layer.

Each pooling layer corresponds to the previous convolutional layer. The neuron in the pooling layer combines a small $N \times 1$ patch of the convolution layer. The most common pooling operation is max pooling, which is used throughout this paper. The max-pooling is as follows:

$$a_j = \max_{N \times 1}\left(a_i^{n \times 1} u(n, 1) \right) \tag{7}$$

where $u (n, 1)$ is a window function to the patch of the convolution layer, and a_j is the maximum in the neighborhood.

After a series of convolutions, deep features were extracted and flattened into a column of neurons, which served as the input layer for subsequent FC layers [43]. The output of

the fully connected layers is fed to a classifier to generate the required classification result. The input of the 1-D CNN is a pixel vector, and the output of the system is the label of the pixel vector. It consists of several convolutional and pooling layers, as shown in Hu's 1-D CNN (Figure 10) [46].
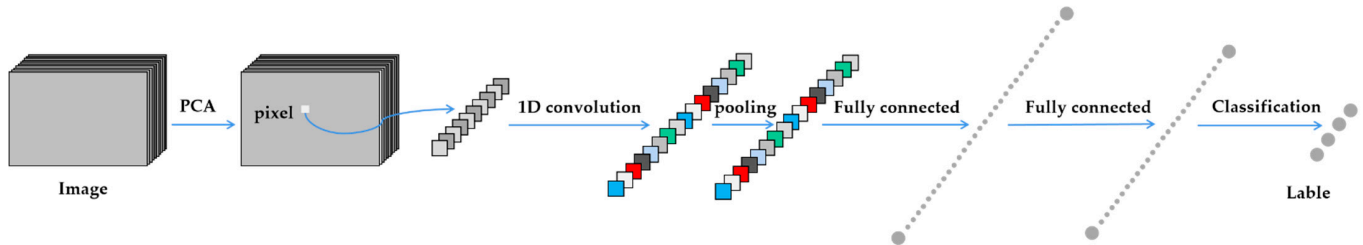
Figure 10. Illustration of the one-dimensional convolutional neural network (1-D CNN) framework.
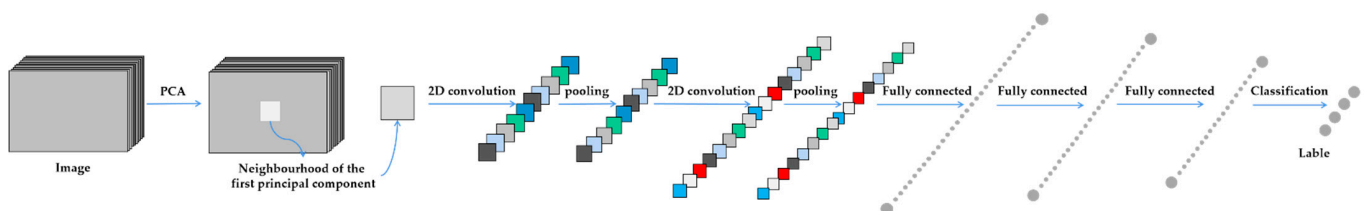
3.2. Two-Dimensional CNN

The 2-D convolutional layer is obtained by the extension of (6). The value of a neuron v_{xy}^{ij} at position (x, y) of the jth feature map in the ith layer is denoted as follows [33]:

$$v_{ij}^{xy} = g\left(b_{ij} + \sum_m \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} w_{ijm}^{pq} v_{(i-1)m}^{(x+p)(y+q)} \right) \tag{8}$$

where m indexes the feature map in the $(i-1)$th layer connected to the current (jth) feature map, w_{ijm}^{pq} is the weight of position (p, q) connected to the mth feature map, P_i and Q_i are the height and width of the spatial convolution kernel, and b_{ij} is the bias of the jth feature map in the ith layer. Pooling is carried out in a similar way to the 1-D CNN.

Based on the theory described previously, a variety of CNN architectures can be developed. The architecture is shown in Figure 11. The 11×11 neighborhoods of a current pixel are selected as the input to the 2-D CNN model. The 2-D CNN model includes two convolution layers and two pooling layers. Each layer of CNN contains 2D convolution and pooling. A 3×3 kernel can be selected to run convolution, and a 2×2 kernel can be selected for pooling. After several layers of convolution and pooling, the input image can be represented by some feature vectors, which capture the spatial information contained in the 11×11 neighborhood region of the input pixel. Then, the learned features are fed to the fully connected layer for classification.

Figure 11. Illustration of the two-dimensional convolutional neural network (2-D CNN) framework.
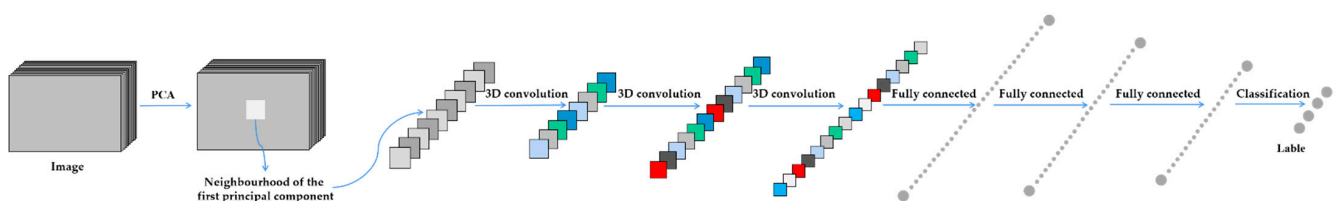
3.3. Three-Dimensional CNN

As can be seen in Sections 3.1 and 3.2, the 1-D CNN extracts spectral features and the 2-D CNN extracts the local spatial features of each pixel. Furthermore, the 3-D CNN can learn both spatial and spectral features. The value of a neuron v_{ij}^{xyz} at position (x, y, z) of the jth feature map in the ith layer is given by [33]:

$$v_{ij}^{xyz} = g\left(\sum_m \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} \sum_{r=0}^{R_i-1} w_{ijm}^{pqr} v_{(i-1)m}^{(x+p)(y+q)(z+r)} + b_{ij} \right) \tag{9}$$

where m indexes the feature map in the $(i - 1)$th layer connected to the current (jth) feature map, and P_i and Q_i are the height and width of the spatial convolution kernel, respectively. R_i is the size of the kernel along the spectral dimension, w_{ijm}^{pqr} is the value of position (p, q, r) connected to the mth feature map, and b_{ij} is the bias of the jth feature map in the ith layer.

The structure diagram of the 3-D CNN classification framework is shown in Figure 12. The 11×11 neighborhoods of a current pixel are selected as the input to the 3-D CNN model. The 3-D CNN model includes three convolution layers. The spatial size of the first 3D convolution kernel is set to $5 \times 3 \times 3$. The spatial sizes of the second and third convolution kernels are set to $3 \times 3 \times 3$. Here, the pooling operations are not applied, which are known for reducing the spatial resolution [27]. After a series of convolutions, deep features were extracted and flattened into a column of neurons, which served as the input layer for subsequent FC layers. The output of the fully connected layers is fed to a classifier to generate the required classification result.


Figure 12. Illustration of the three-dimensional convolutional neural network (3-D CNN) framework.

## 4. Experimental Results and Analysis

4.1. Classification Results

To prove the feasibility of the methods, the CNNs were tested and compared with six different typical methods, including SAM, SID, FCLSU, SVM, RF and NN. The hardware and software parameters were as follows: Lenovo Desktops, Intel(R) Core (TM) i7-10700K CPU @3.80 GHz 3.79 GHz, 16 GB RAM, NVIDIA GeForce RTX 2080 SUPER 8 GB GPU, Windows 10, python3.8. For SAM, SID and FCLSU, the final classification results were obtained through comparison with the reference spectral curves. For SVM, RF, NN and CNNs, 10% of the number of pixels from each class were extracted as training samples, and the rest were test samples. Then, the trained model was used to predict the images and obtain the final classification results. The CNN models were implemented based on the PyTorch deep learning framework for the training and testing models. All the experiments were repeated several times. Under the same parameters, there are very small variations according to their results. The OA, AA and kappa denote the overall accuracy (OA), the average accuracy of all classes (AA) and the kappa coefficient (kappa), which are calculated to evaluate and test the precision. OA represents the number of correctly classified samples in all categories divided by the total sample size by the model. AA stands for the average precision of all classes. Kappa is an accuracy measure based on the confusion matrix, which represents the percentage of errors reduced by classification versus a completely random classification. The classification maps of the three images are shown in Figures 13–15. The classification accuracies over different classification methods are reported in Tables 3–5.
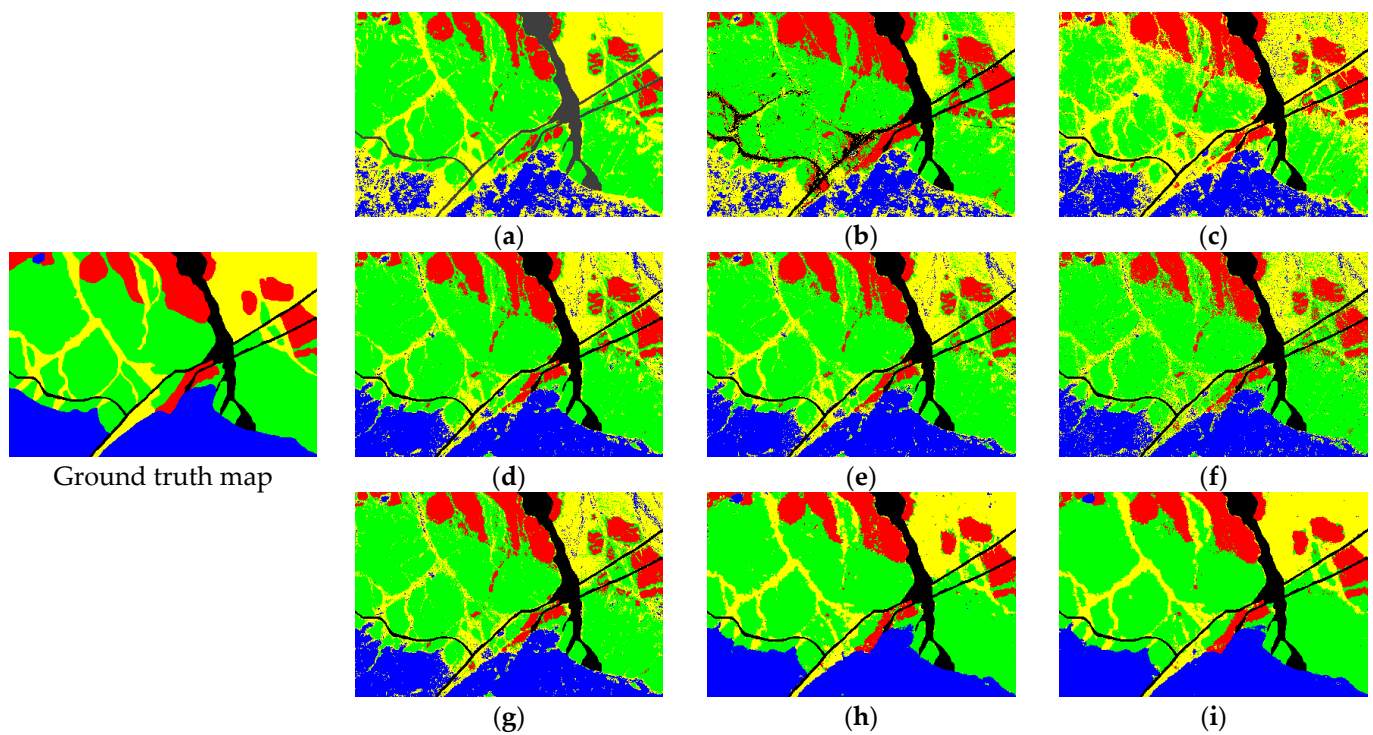
Figure 13. Classification results of all the methods for Liuyuan 1. (**a**) Spectral angle mapping (SAM); (**b**) spectral information divergence (SID); (**c**) fully constrained linear spectral unmixing (FCLSU); (**d**) support vector machine (SVM); (**e**) random forest (RF); (**f**) neural network (NN); (**g**) 1-D CNN; (**h**) 2-D CNN; (**i**) 3-D CNN.
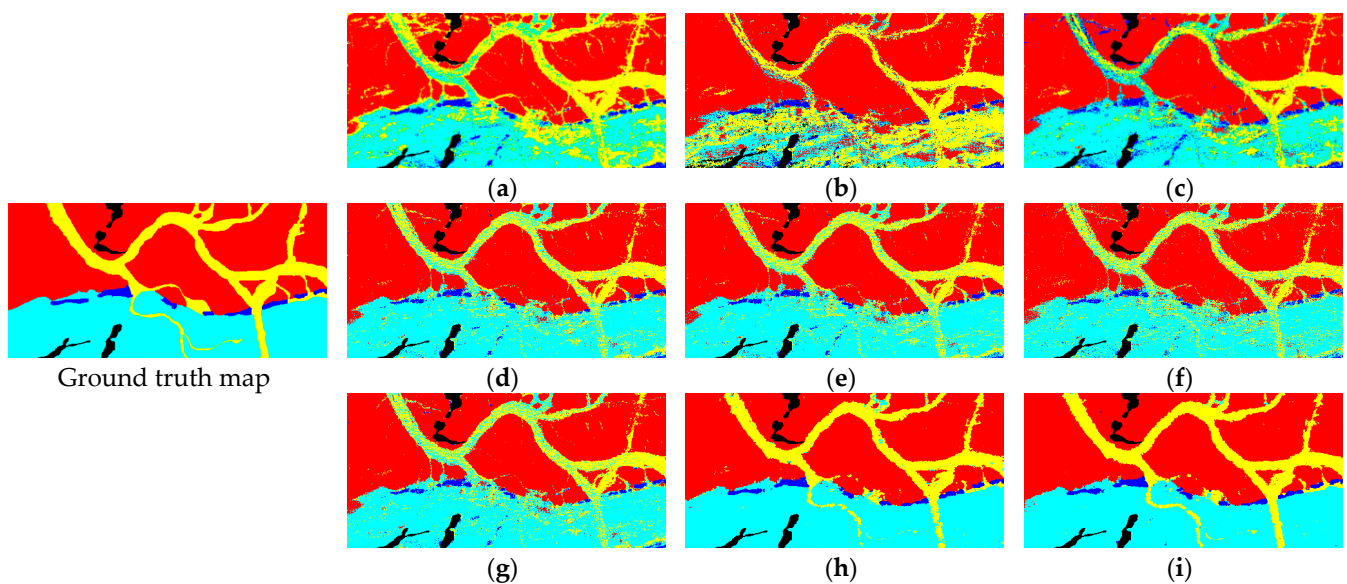
Figure 14. Classification results of all the methods for Liuyuan 2. (**a**) SAM; (**b**) SID; (**c**) FCLSU; (**d**) SVM; (**e**) RF; (**f**) NN; (**g**) 1-D CNN; (**h**) 2-D CNN; (**i**) 3-D CNN.
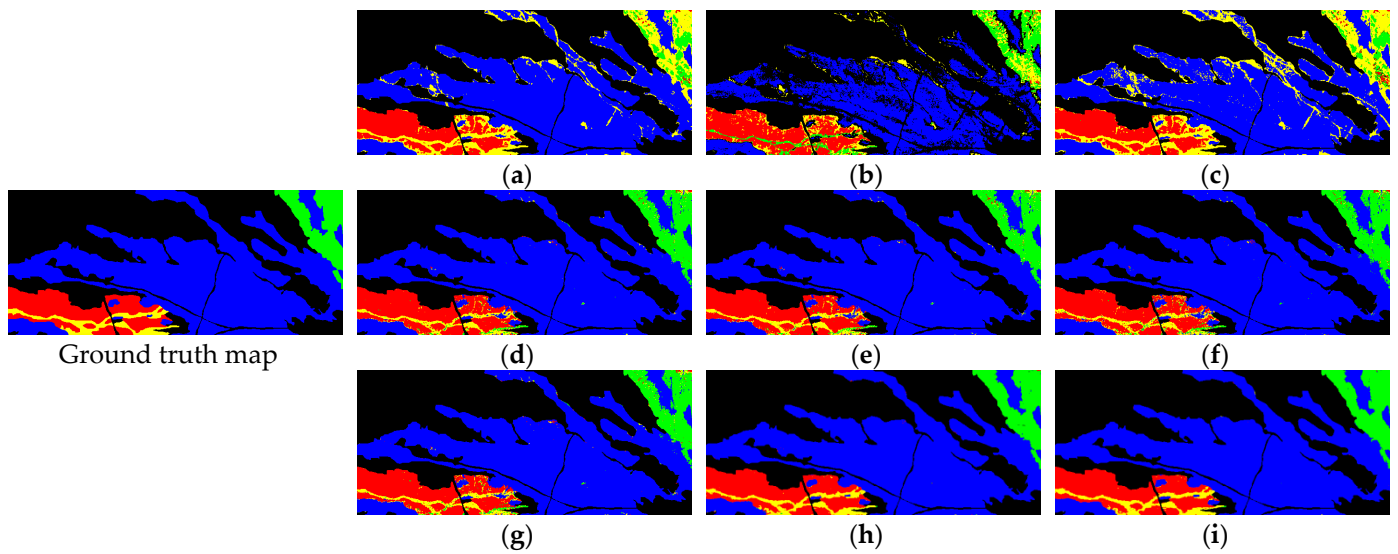
Figure 15. Classification results of all the methods for Liuyuan 3. (**a**) SAM; (**b**) SID; (**c**) FCLSU; (**d**) SVM; (**e**) RF; (**f**) NN; (**g**) 1-D CNN; (**h**) 2-D CNN; (**i**) 3-D CNN.

Table 3. Accuracy of classification results of Liuyuan 1.

Lithology	SAM	SID	FCLSU	SVM	RF	NN	1-D CNN	2-D CNN	3-D CNN
Slate	69.43	93.59	86.26	82.91	84.02	83.90	85.27	93.00	93.97
Syenogranite	87.83	84.91	64.61	90.57	90.67	87.59	89.36	94.72	95.55
Diorite	52.48	50.82	75.53	89.96	90.74	88.91	90.10	98.40	98.32
Quaternary sediment	75.60	48.16	84.89	65.33	70.48	64.67	65.11	89.02	89.12
OA (%)	75.87	72.12	73.42	84.68	86.01	81.27	84.38	94.18	94.70
AA (%)	71.34	69.37	77.82	82.19	83.98	80.48	82.46	93.78	94.24
Kappa	0.6393	0.5869	0.6314	0.7707	0.7915	0.7757	0.7673	0.9142	0.9217

Table 4. Accuracy of classification results of Liuyuan 2.

Lithology	SAM	SID	FCLSU	SVM	RF	NN	1-D CNN	2-D CNN	3-D CNN
Slate	87.50	97.60	94.03	96.59	96.77	95.52	96.59	98.22	98.29
Diorite	69.66	60.09	70.79	53.07	59.37	52.48	56.24	82.97	83.97
Quaternary sediment	75.47	57.42	60.00	59.70	64.58	59.32	57.24	87.32	88.94
Sericite phyllite	68.92	36.52	76.52	87.41	87.80	84.71	87.44	97.73	98.02
OA (%)	79.06	70.54	82.45	86.93	88.03	85.39	86.61	96.07	96.47
AA (%)	75.39	62.91	75.34	74.19	77.13	73.01	74.38	91.56	93.11
Kappa	0.6794	0.5237	0.7228	0.7879	0.8063	0.7641	0.7825	0.9367	0.9432

Table 5. Accuracy of classification results of Liuyuan 3.

Lithology	SAM	SID	FCLSU	SVM	RF	NN	1-D CNN	2-D CNN	3-D CNN
Slate	85.46	84.32	84.76	94.39	94.89	91.76	93.72	97.04	96.18
Syenogranite	31.63	63.27	25.88	91.45	91.03	91.32	91.19	97.46	97.49
Diorite	93.71	96.55	87.77	99.13	99.32	99.26	98.93	99.53	99.61
Quaternary sediment	85.60	22.28	85.06	59.57	61.91	60.60	64.36	87.55	89.36
OA (%)	86.60	86.89	81.68	96.22	96.49	96.01	96.16	98.53	98.56
AA (%)	74.1	66.61	70.86	86.13	86.79	85.73	87.05	95.39	95.66
Kappa	0.7107	0.7549	0.6285	0.9117	0.9177	0.9062	0.9105	0.9659	0.9664

4.1.1. Liuyuan 1

The thematic maps of Liuyuan 1 by different methods are visually shown in Figure 13a–i. It shows that the 3-D CNN algorithm achieved the best classification results for most land cover classes. As shown in Figure 13a–c, the classification results based on the traditional methods SAM, SID and FCLSU appear to be wrongly misclassified, and a part of the diorite is misclassified as quaternary sediment. In Figure 13d–i, almost nothing diorite is misclassified as Quaternary sediment. In particular, the 3-D CNN method compared with the other methods is largely free of misclassification for each lithology class due to its powerful learning capabilities. The 3-D CNN offers better classification performance, greater noise immunity and more accurate boundary classification. Additionally, as shown in Table 3, the evaluation results of different algorithms for Liuyuan 1 are shown. The accuracy of diorite is only 52.48%, 50.82% and 75.53% based on SAM, SID and FCLSU, respectively. The classification effects of traditional methods, including the SAM, SID and FCLSU methods, are the worst. MALs, such as SVM, RF and NN, show a good advantage, and the RF model obtains the best result. The 2-D CNN and 3-D CNN models obtained the best result, with overall accuracies of 94.18% and 94.70%, respectively. The 3D-CNN model is 0.52% higher than 2D-CNN, indicating that the inclusion of spectral features in 3D-CNN improves the classification accuracy. Particularly for slate, 3D-CNN effectively improves the classification accuracy, 0.97% higher than 2D-CNN. However, the overall accuracy of the 1-D CNN is only 84.38%, which is lower than that of SVM and RF. The 1-D CNN input is a single pixel, and the 1-D network used in this case has only one convolutional layer and one pooling layer, which may lead to poorer classification results. Therefore, CNN algorithms are conducive to improving the classification of high-resolution and hyperspectral TIR remote sensing images.

4.1.2. Liuyuan 2

The thematic maps of Liuyuan 2 are visually shown in Figure 14a–i. The image of Liuyuan 2 has a complex lithology distribution, and it is very difficult to distinguish between sericite phyllite and quaternary sediments in the spectral domain because sericite phyllite is produced by metamorphism of sedimentary rocks, and Quaternary sediments are also clastic sedimentary rocks. SAM, SID and FCLSU have a serious misclassification of quaternary sediments and sericite phyllite in Figure 14a–c. The sericite phyllite is wrongly misclassified as quaternary sediments. The classification area is smaller than the real area. The boundary of classification is fuzzy, and the noise level is high. The sericite phyllite classification results of SVM, RF, NN and 1-D CNN are slightly better than SAM, SID and FCLSU. Compared with other algorithms, the 2-D CNN and 3-D CNN algorithms have a stronger classification capability at the corresponding location, such as the classification of sericite phyllite and quaternary sediments in Figure 14h,i. In addition, as shown in Table 4, the 3D-CNN achieved the best result with an overall accuracy of 96.47%, which is 0.4% higher than obtained by 2D-CNN. Particularly for diorite, 3D-CNN effectively improves the classification accuracy, 1% higher than 2D-CNN. Meanwhile, the average precision obtained by the 3-D CNN and 2-D CNN is significantly higher than that of the other six methods by approximately 20%, which significantly improves the classification precision of per class. The 3-D CNN achieved the best result with an average accuracy, which is 1.55% higher than 2-D CNN. The 3-D CNN is significantly better than 2-D CNN. Other evaluation indexes also have some potential advantages. Therefore, the 3-D CNN algorithm can effectively improve the classification of complex remote sensing images.

4.1.3. Liuyuan 3

Figure 15a–f shows the thematic maps of Liuyuan 3. For SAM and FCLSU, the syenogranite classification ability is very poor in Figure 15a–c because a part of the syenogranite is misclassified as quaternary sediment. For SID, the classification result of syenogranite in Figure 15b is slightly better than SAM and FCLSU. A portion of the syenogranite is also misclassified as quaternary deposits. Among these three methods,

slate and diorite classification results are significantly better than those of syenogranite and quaternary sediments because the spectral diagnostic features of slate and diorite are more pronounced. SVM, RF, NN and 1-D CNN obtain higher partition precision and more accurate partition boundaries than SAM, SID and FCLSU in Figure 15d–g. The classification results of slate, syenogranite and diorite are correct, but there is still some noise. However, the classification accuracy of Quaternary sediments is lower than that of SAM, SID and FCLSU. Because of the spectral similarity between syenogranite and quaternary sediments, it led to a better classification of quaternary sediments than syenogranite by SAM, SID and FCLSU. For MALs and 1-D CNN, the number of training samples depends on the classification results. There are fewer training samples of quaternary sediments, and the calculations of these algorithms are relatively simple, leading to the worst classification results for this class. For 2-D CNN and 3-D CNN, their results have a noticeable improvement in Figure 15h,i and the boundary partition is very satisfactory. Additionally, as shown in Table 5, the evaluation results of different algorithms for Liuyuan 1 are shown. The 3-D CNN model obtained the best result, with a 98.56% overall accuracy, which is 0.03% higher than 2D-CNN. MALs such as SVM, RF and NN also show a good advantage with an overall classification accuracy that is even closer to 3-D CNN and 2-D CNN. The classification effects of the SAM, SID and FCLSU methods are the worst. Their overall classification accuracy is only approximately 85%. Compared with Liuyuan 1 and Liuyuan 2, this scene has more explicit lithological units. In general, all classification methods have high accuracy. The 2-D CNN has already achieved such a high overall accuracy of 98.53%, so 3-D CNN has a smaller performance improvement of 0.03% on top of that.

5. Discussion

According to the experimental results presented above, CNNs achieve better overall classification results than SAM, SID, FCLSU, SVM, RF and NN. Therefore, CNN methods are conducive to promoting rock classification using hyperspectral TIR imagery. In addition, further discussion and analysis of the effect of MLA and CNN parameter settings, i.e., number of training samples, on the recognition results are presented below.

The classification accuracies of Liuyuan 1, 2 and 3, including the OA, AA and kappa, of all comparison methodologies for three datasets with different numbers of training samples are shown in Figures 16–18. In the figure, the two horizontal axes represent the number of training samples and the different supervised methods. The vertical axis represents the classification accuracy of the mentioned algorithm. For each of the six methods, the figure indicates the algorithm's performance within given training samples. The numbers of training samples per class of all the comparison methods were set up to 5%, 10% and 15% of the total sample number per class. When the number of samples increases proportionally, the overall classification accuracy of 3-D CNN and 2-D CNN is improved by approximately 0.4–3.7%. The overall classification accuracy of SVM, RF, NN and 1-D CNN under different sample sizes has small differences, and the overall accuracy fluctuates within 1%. For instance, in terms of the difference in classification overall accuracy, NN was greater when the number of samples was 10% per class than when the number of samples was 15% per class, as shown in Figure 16a. For the average accuracy of each category in the three datasets, compared to 5% of the total sample number per class, 2-D CNN and 3-D CNN have the highest average accuracy when the sample size is 15%. The average accuracy of 10% of the total sample number per class is increased by 5.17% and 5.95%, especially using the 3-D CNN and 2-D CNN, compared with the average accuracy of 5% of the total sample number per class in Figure 17b. The average accuracy of SVM, RF, NN and 1-D CNN under different sample sizes has small differences, and the accuracy fluctuates within 1%. When the number of samples increases proportionally, the variation in kappa coefficient values is consistent with the variation in the average accuracy and overall accuracy values. It can also be found in Figures 16–18 that the difference is approximately 1% in terms of the overall accuracy, the average accuracy and the kappa coefficient when the sample size is

10% and 15%. Therefore, when the sample size of each class is 10%, good classification results can be achieved.
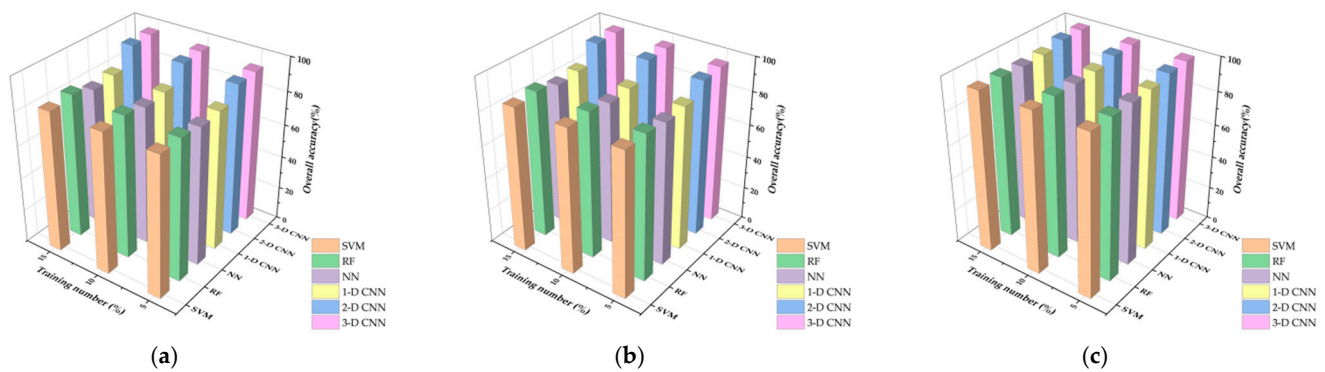
Figure 16. Classification performances using different numbers of training samples for the Liuyuan 1 dataset. (**a**) Overall accuracy (OA); (**b**) average accuracy of all classes (AA); (**c**) kappa.
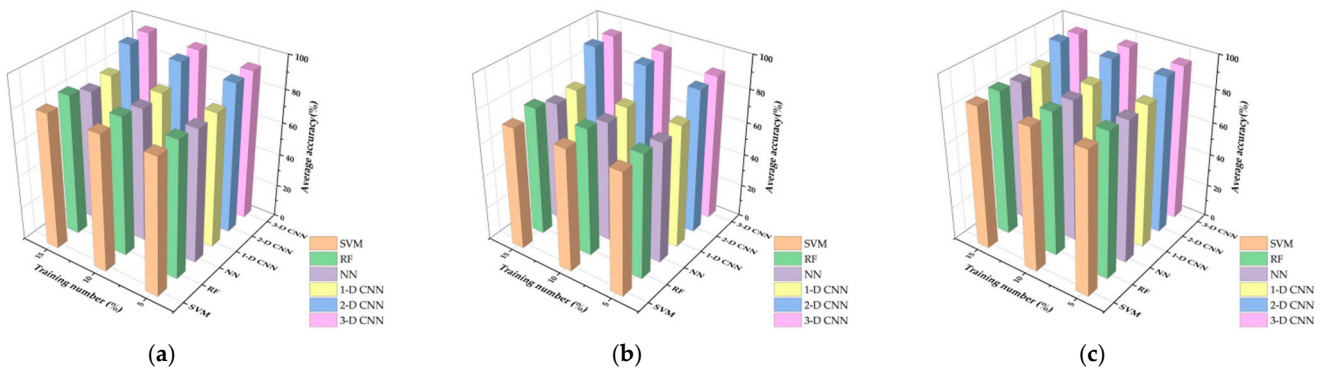
Figure 17. Classification performances using different numbers of training samples for the Liuyuan 2 dataset. (**a**) OA; (**b**) AA; (**c**) kappa.
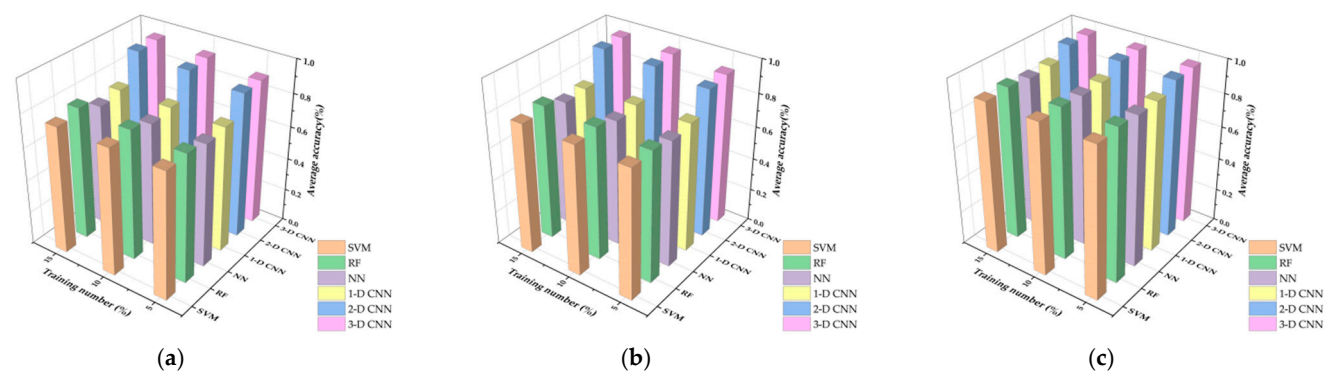
Figure 18. Classification performances using different numbers of training samples for the Liuyuan 3 dataset. (**a**) OA; (**b**) AA; (**c**) kappa.

6. Conclusions

The application of CNNs to complete lithology classification using thermal infrared hyperspectral TASI data is the main contribution of this paper. Because the methods are designed in a supervised way, it is also important to provide prior knowledge and manual assistance in determining the label of the image elements in practical applications. The field measurement data coupled with MNF transformation using TASI emissivity images were able to obtain better pixel spectra and training samples. Six state-of-the-art

algorithms, including SAM, SID, FCLSU, SVM, RF and NN, were implemented to make comparisons with the performance of the CNNs. Compared with other algorithms, the classification accuracy of CNNs, especially 2-D CNN and 3-D CNN, are approximately 2.5–12% higher SVM, RF and NN and are approximately 12–25% higher SAM, SID and FCLSU. The classification result of the 1-D CNN is inferior to the other two CNNs, and its classification accuracy is roughly comparable to that of MALs. The numerical results also verify the effective classification of CNNs, which shows that the algorithm can obtain satisfactory classification results. It is demonstrated that the 2-D CNN and 3-D CNN algorithms improve the classification of hyperspectral TIR remote sensing images.

References

1. Van der Meer, F.D.; van der Werff, H.M.A.; van Ruitenbeek, F.J.A.; Hecker, C.A.; Bakker, W.H.; Noomen, M.F.; van der Meijde, M.; Carranza, E.J.M.; Smeth, J.B.D.; Woldai, T. Multi- and Hyperspectral Geologic Remote Sensing: A review. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *14*, 112–128. [CrossRef]
2. Cloutis, E.A. Review Article Hyperspectral Geological Remote Sensing: Evaluation of Analytical Techniques. *Int. J. Remote Sens.* **1996**, *17*, 2215–2242. [CrossRef]
3. Chen, X.; Warner, T.A.; Campagna, D.J. Integrating Visible, Near-Infrared and Short-Wave Infrared Hyperspectral and Multispectral Thermal Imagery for Geological Mapping at Cuprite, Nevada. *Remote Sens. Environ.* **2007**, *110*, 344–356. [CrossRef]
4. Salles, R.D.R.; de Souza Filho, C.R.; Cudahy, T.; Vicente, L.E.; Monteiro, L.V.S. Hyperspectral Remote Sensing Applied to Uranium Exploration: A Case Study at the Mary Kathleen Metamorphic-Hydrothermal U-REE Deposit, NW, Queensland, Australia. *J. Geochem. Explor.* **2017**, *179*, 36–50. [CrossRef]
5. Zhang, T.-T.; Liu, F. Application of Hyperspectral Remote Sensing in Mineral Identification and Mapping. In Proceedings of the 2012 2nd International Conference on Computer Science and Network Technology (ICCSNT 2012), Changchun, China, 29–31 December 2012; pp. 103–106.
6. Ni, L.; Xu, H.; Zhou, X. Mineral Identification and Mapping by Synthesis of Hyperspectral VNIR/SWIR and Multispectral TIR Remotely Sensed Data with Different Classifiers. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 3155–3163. [CrossRef]
7. Ninomiya, Y.; Fu, B. Thermal Infrared Multispectral Remote Sensing of Lithology and Mineralogy Based on Spectral Properties of Materials. *Ore Geol. Rev.* **2019**, *108*, 54–72. [CrossRef]
8. Asadzadeh, S.; de Souza Filho, C.R. A Review on Spectral Processing Methods for Geological Remote Sensing. *Int. J. Appl. Earth Obs.* **2016**, *47*, 69–90. [CrossRef]
9. Cui, J.; Yan, B.; Dong, X.; Zhang, S.; Zhang, J.; Tian, F.; Wang, R. Temperature and Emissivity Separation and Mineral Mapping Based On Airborne TASI Hyperspectral Thermal Infrared Data. *Int. J. Appl. Earth Obs.* **2015**, *40*, 19–28. [CrossRef]
10. Boubanga-Tombet, S.; Huot, A.; Vitins, I.; Heuberger, S.; Veuve, C.; Eisele, A.; Hewson, R.; Guyot, E.; Marcotte, F.; Chamberland, M. Thermal Infrared Hyperspectral Imaging for Mineralogy Mapping of a Mine Face. *Remote Sens.* **2018**, *10*, 1518. [CrossRef]
11. Black, M.; Riley, T.R.; Ferrier, G.; Fleming, A.H.; Fretwell, P.T. Automated Lithological Mapping Using Airborne Hyperspectral Thermal Infrared Data: A Case Study from Anchorage Island, Antarctica. *Remote Sens. Environ.* **2016**, *176*, 225–241. [CrossRef]

12. Ninomiya, Y.; Fu, B.; Cudahy, T.J. Detecting Lithology with Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Multispectral Thermal Infrared "Radiance-at-Sensor" Data. *Remote Sens. Environ.* **2005**, *99*, 127–139. [CrossRef]

13. Reath, K.A.; Ramsey, M.S. Exploration of Geothermal Systems Using Hyperspectral Thermal Infrared Remote Sensing. *J. Volcanol. Geothem. Res.* **2013**, *265*, 27–38. [CrossRef]

14. Kurata, K.; Yamaguchi, Y. Integration and Visualization of Mineralogical and Topographical Information Derived from ASTER and DEM Data. *Remote Sens.* **2019**, *11*, 162. [CrossRef]

15. Zhao, H.; Zhang, L.; Zhao, X.; Yang, H.; Yang, K.; Zhang, X.; Wang, S.; Sun, H. A New Method of Mineral Absorption Feature Extraction from Vegetation Covered Area. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 5437–5440.

16. Vignesh, K.M.; Kiran, Y. Comparative Analysis of Mineral Mapping for Hyperspectral and Multispectral Imagery. *Arab. J. Geosci.* **2020**, *13*, 160. [CrossRef]

17. Ni, L.; Wub, H. Mineral Identification and Classification by Combining Use of Hyperspectral VNIR/SWIR and Multispectral TIR Remotely Sensed Data. In Proceedings of the IGARSS 2019—2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 3317–3320.

18. Krupnik, D.; Khan, S.D. High-Resolution Hyperspectral Mineral Mapping: Case Studies in the Edwards Limestone, Texas, USA and Sulfide-Rich Quartz Veins from the Ladakh Batholith, Northern Pakistan. *Minerals.* **2020**, *10*, 967. [CrossRef]

19. Chen, X.; Chen, J.; Pan, J. Using Geochemical Imaging Data to Map Nickel Sulfide Deposits in Daxinganling, China. *SN Appl. Sci.* **2021**, *3*, 324. [CrossRef]

20. Villa, P.; Pepe, M.; Boschetti, M.; de Paulis, R. Spectral Mapping Capabilities of Sedimentary Rocks Using Hyperspectral Data in Sicily, Italy. In Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium, Vancouver, BC, Canada, 24–29 July 2011; pp. 2625–2628.

21. Hecker, C.; van Ruitenbeek, F.J.A.; van der Werff, H.M.A.; Bakker, W.H.; Hewson, R.D.; van der Meer, F.D. Spectral Absorption Feature Analysis for Finding Ore: A Tutorial on Using the Method in Geological Remote Sensing. *IEEE Geosci. Remote Sens. Mag.* **2019**, *7*, 51–71. [CrossRef]

22. Kopačková, V.; Koucká, L. Integration of Absorption Feature Information from Visible to Longwave Infrared Spectral Ranges for Mineral Mapping. *Remote Sens.* **2017**, *9*, 1006. [CrossRef]

23. Kumar, C.; Chatterjee, S.; Oommen, T.; Guha, A. Automated Lithological Mapping by Integrating Spectral Enhancement Techniques and Machine Learning Algorithms Using AVIRIS-NG Hyperspectral Data in Gold-Bearing Granite-Greenstone Rocks in Hutti, India. *Int. J. Appl. Earth Obs.* **2020**, *86*, 102006. [CrossRef]

24. Wan, S.; Lei, T.C.; Ma, H.L.; Cheng, R.W. The Analysis on Similarity of Spectrum Analysis of Landslide and Bareland through Hyper-Spectrum Image Bands. *Water* **2019**, *11*, 2414. [CrossRef]

25. Othman, A.; Gloaguen, R. Improving Lithological Mapping by SVM Classification of Spectral and Morphological Features: The Discovery of a New Chromite Body in the Mawat Ophiolite Complex (Kurdistan, NE Iraq). *Remote Sens.* **2014**, *6*, 6867–6896. [CrossRef]

26. Okada, N.; Maekawa, Y.; Owada, N.; Haga, K.; Shibayama, A.; Kawamura, Y. Automated Identification of Mineral Types and Grain Size Using Hyperspectral Imaging and Deep Learning for Mineral Processing. *Minerals* **2020**, *10*, 809. [CrossRef]

27. Li, Y.; Zhang, H.; Shen, Q. Spectral–Spatial Classification of Hyperspectral Imagery with 3D Convolutional Neural Network. *Remote Sens.* **2017**, *9*, 67. [CrossRef]

28. Zhang, J.; Wei, F.; Feng, F.; Wang, C. Spatial–Spectral Feature Refinement for Hyperspectral Image Classification Based on Attention-Dense 3D-2D-CNN. *Sensors* **2020**, *20*, 5191. [CrossRef]

29. Hinton, G.E.; Osindero, S.; Teh, Y. A Fast Learning Algorithm for Deep Belief Nets. *Neural Comput.* **2006**, *18*, 1527–1554. [CrossRef]

30. Ding, H.; Xu, L.; Wu, Y.; Shi, W. Classification of Hyperspectral Images by Deep Learning of Spectral-Spatial Features. *Arab. J. Geosci.* **2020**, *13*, 464. [CrossRef]

31. Lei, T.; Zhang, Y.; Lv, Z.; Li, S.; Liu, S.; Nandi, A.K. Landslide Inventory Mapping from Bitemporal Images Using Deep Convolutional Neural Networks. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 982–986. [CrossRef]

32. Zuo, R.; Xiong, Y.; Wang, J.; Carranza, E.J.M. Deep Learning and Its Application in Geochemical Mapping. *Earth Sci. Rev.* **2019**, *192*, 1–14. [CrossRef]

33. Chen, Y.; Jiang, H.; Li, C.; Jia, X.; Ghamisi, P. Deep Feature Extraction and Classification of Hyperspectral Images Based On Convolutional Neural Networks. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6232–6251. [CrossRef]

34. Yu, C.; Han, R.; Song, M.; Liu, C.; Chang, C. A Simplified 2D-3D CNN Architecture for Hyperspectral Image Classification Based on Spatial–Spectral Fusion. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 2485–2501. [CrossRef]

35. Cracknell, M.J.; Reading, A.M. Geological Mapping Using Remote Sensing Data: A Comparison of Five Machine Learning Algorithms, Their Response to Variations in the Spatial Distribution of Training Data and the Use of Explicit Spatial Information. *Comput. Geosci.* **2014**, *63*, 22–33. [CrossRef]

36. Jing, C.; Bokun, Y.; Runsheng, W.; Feng, T.; Yingjun, Z.; Dechang, L.; Suming, Y.; Wei, S. Regional-scale Mineral Mapping Using ASTER VNIR/SWIR Data and Validation of Reflectance and Mineral Map Products Using Airborne Hyperspectral CASI/SASI Data. *Int. J. Appl. Earth Obs.* **2014**, *33*, 127–141. [CrossRef]

37. Gillespie, A.; Rokugawa, S.; Matsunaga, T.; Cothern, J.S.; Hook, S.; Kahle, A.B. A Temperature and Emissivity Separation Algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Images. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1113–1126. [CrossRef]

38. Watson, K. Spectral Ratio Method for Measuring Emissivity. *Remote Sens. Environ.* **1992**, *42*, 113–116. [CrossRef]

39. Zhao, H.; Deng, K.; Li, N.; Wang, Z.; Wei, W. Hierarchical Spatial-Spectral Feature Extraction with Long Short Term Memory (LSTM) for Mineral Identification Using Hyperspectral Imagery. *Sensors* **2020**, *20*, 6854. [CrossRef] [PubMed]

40. Othman, A.A.; Gloaguen, R. Integration of Spectral, Spatial and Morphometric Data into Lithological Mapping: A comparison of Different Machine Learning Algorithms in the Kurdistan Region, NE Iraq. *J. Asian Earth Sci.* **2017**, *146*, 90–102. [CrossRef]

41. Cheng, G.; Xie, X.; Han, J.; Guo, L.; Xia, G. Remote Sensing Image Scene Classification Meets Deep Learning: Challenges, Methods, Benchmarks, and Opportunities. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 3735–3756. [CrossRef]

42. Tran, D.; Bourdev, L.; Fergus, R.; Torresani, L.; Paluri, M. Learning Spatiotemporal Features with 3D Convolutional Networks. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 11–18 December 2015; pp. 4489–4497.

43. Dong, Y.; Yang, C.; Zhang, Y. Deep Metric Learning with Online Hard Mining for Hyperspectral Classification. *Remote Sens.* **2021**, *13*, 1368. [CrossRef]

44. Roy, S.K.; Krishna, G.; Dubey, S.R.; Chaudhuri, B.B. HybridSN: Exploring 3-D–2-D CNN Feature Hierarchy for Hyperspectral Image Classification. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 277–281. [CrossRef]

45. Cheng, G.; Yang, C.; Yao, X.; Guo, L.; Han, J. When Deep Learning Meets Metric Learning: Remote Sensing Image Scene Classification via Learning Discriminative CNNs. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 2811–2821. [CrossRef]

46. Hu, W.; Huang, Y.; Wei, L.; Zhang, F.; Li, H. Deep Convolutional Neural Networks for Hyperspectral Image Classification. *J. Sens.* **2015**, *2015*, 258619. [CrossRef]