Journal of Medical Systems (2019) 43: 216
https://doi.org/10.1007/s10916-019-1347-9

TRANSACTIONAL PROCESSING SYSTEMS

Hyperspectral Image Features Classification Using Deep Learning Recurrent Neural Networks

R. Venkatesan[1] · S. Prabu[1]

Received: 2 March 2019 / Accepted: 20 May 2019 / Published online: 4 June 2019

Abstract
The implementation of Deep learning (DL) techniques, Object detection and classification has achieved remarkable results in remote sensing application. Deep learning with Recurrent Neural Network (RNN) technique on hyper-spectral data has been presented here. The only model which can analyze the hyper-spectral pixels as the sequence of information and also to identify the additional information categories through network reasoning is RNN model. This is first time that the framework of RNN has been introduced for the classification of hyper spectral Image. An activation function is proposed by the DL-RNN and also the parameter rectified functions for analyzing the sequence of data in the hyper-spectral images. Throughout the training procedure, the higher learning rates are fairly used by the activation function which has been proposed by avoiding the risk of divergence. In the proposed system the pixels of hyper–spectral images through the sequential perspective has been processed for capturing the sequence based data. The experimental result also shows that the proposed RNN has produced the improved F- score than the traditional deep learning methods.

Keywords Hyperspectral imaging · Deep learning · Recurrent neural network · Activation functions · Features vectors

Introduction

In order to obtain benefit, hundreds of spectral channels are composed over a single scene in the remote sensing community with the wide usage of Hyper-spectral imagery (HSI). To extract the features of an image, HSI needs the most robust and accurate technique for classification. Due to complexity nature of image scene, having large amount of data, assorted pixels and limited training samples, it has been predominantly a challenging problem to classify those HSI images. In the last few decades, many attempts have been made to address this unease. Remote sensing, image acquisition and processing it has recently became very imperative in Earth's observation problem. Many applications such as environment monitoring and managing, agriculture, defense or security or intelligent issues, problem on observation exhibited in practical. Multispectral and hyper-spectral images are formed by computational and efficient processing on multi-spectral band images. Information are collected from these kind of images consequent to observation on large areas on the Earth's surface, using many contiguous spectral bands, thus a three-dimensional data cube has been significantly created in larger size than the traditional remote sensed images. Meticulous improvement on computation is required by the multi-spectral and hyper-spectral images for storing and advanced processing as an outcome. For classifying the remote sensing information, a semi-supervised method, deep learning has been recently introduced. To classify the complex problems in an effective manner, an artificial neural network is extended and it is considered as the deep network. On the other hand, as the deep network requires a huge number of samples for training, the deep network training is quite expensive. As working out samples are limitedly accessible and the characteristic space dimensionality is huge, applying a deep network for classifying the hyper-spectral imaging is not practically implemented.

The huge spectral decision of HSIs is done by the fee of decreased spatial declaration. This is an immediate outcome of

This article is part of the Topical Collection on *Transactional Processing Systems*

✉ R. Venkatesan
venkatesaneng@gmail.com

S. Prabu
prabu.sevugan@gmail.com

[1] School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

the bodily obstacles of the sensor because the aggregate strength of circumstance luminous for every slender band is too low to permit used by numerous pixels. The magnificent spectral reflectances of multiple materials are combined within the prospect regularly in every pixel of HSI short spatial decision. The reflectance spectra of a consultant material in a prospect are referred as an end member spectrum, and the profusion portion of that end member is comparatively shared in a pixel. The profusion vectors called the abundant fractions are used within each pixel of HSI. Estimating the concurrent end participants in HSI is the mission of hyper spectral un mixing (HSU). Therefore, the pixels spectrum is decomposed right keen into a subjective mixture to give up members. Un blending strategies depends on the underlying blending models, which is probably deployed using the actual mixing of the spectra that takes vicinity [1]. There are two different kinds of mixing fashions such as non-linear and linear models. By using a couple of substance, Non-linear fashions model the bodily connections among the slightly scattered within the scene wherein the interactions can live in a traditional multi-layered degree, or in a microscopic intimate degree. The samples which can be most informative are decided on from the subset of unlabeled samples in an iterative method through Active mastering. This desire is performed primarily based on the computational final results of the version by way of ranking score. The elected candidates are delivered to the training set, and with this new training samples, the classifier is then skilled. Instead of random selected samples, the actively selected samples are more efficient in the course of the training carried out, as it makes use of the maximum suitable samples for training. Therefore, whilst as compared to the traditional semi-supervised mastering strategies [2], the deep community is speedy trained by means of energetic studying approach with fewer education samples. At first, multi-label class problem is taken into consideration, in which each pixel in satellite TV for pc image is annotated with a couple of labels which encodes mixing of different substances inside a unmarried pixel. Moreover, to searching for “extremely good representations” on satellite TV for pc information to symbolize the spectral and temporal functions of authentic records below a actual-international state of affairs, a sort of superior synthetic network known as deep mastering framework of Recurrent Neural Network is employed successfully by that specialize in a specific unsupervised function gaining knowledge of method.

The neural network construction framework is shown in Fig. 1.

Literature review

Qi Wang, et.al,… [3] projected a new-fangled method for selection of band based on MR. To assess the structure of hyper-spectral data, the assorted structure is suitably considered instead of similarity rating in the Euclidean space. Ranking operation is performed on the given related measurements which fabricate the ranking score and this score helps for the selection of subsequent data. This new technique of ranking problem has been formulated which alternates the hyper-spectral band selection. The distance between the inter-bands is estimated in a batch mode. The distance between two individual bands is mostly computed in existing techniques. The calculated results then hand out as supervision for band assortment. This approach may not be appropriate for the band assortment in a sequence, as the selected one might resemble the previous one. The batch of already selected bands is preferred as query and is comparatively examined with the whole batch in the implementation. Additional selected band can be ensured distinctly with the previous selected ones.

Gustavo Camps-Valls, et.al,…. [4] permitted to characterize the interested objects (for example class of land-cover) among exceptional correctness, in addition to to remain inventories up-to-date. For advancement in signal processing and exploitation in algorithms, spectral resolutions have been improved. Recently, classification on hyper-spectral images has been mainly focused in popular as a challenging problem and have fascinated the attention of further technical restraint such because machine learning, image processing and computer revelation. The process of assigning single pixel to set of class is denoted as ‘classification’, in remote sensing community while the process of aggregating the pixels into objects, then assigning to a class is termed as segmentation. In earlier days, linear discriminant analysis or Gaussian maximum likelihood are parameters used for classification in remote sensing than HSI. Based on the estimation of the covariance matrix, these two methods images comprises of four to ten bands in dimension, which was successful when dealing with early multi-spectral images. When the dimensionality of the pixels is increased in hundreds, the rules are changed by HSI. Due to unavailability of labeled samples, standard parametric method for the estimation of covariance matrices becomes unfeasible or unreliable. For the unavailability of labeled samples regularization has been included in the implicated covariance matrices either explicitly through Tikhonov’s terms, or by performing classification in a reduced dimension of a subspace.

Wei Li, et.al,… [5] Analyzed a work of fiction classification framework and has applied a deep CNN based on studying pixel-pair capabilities (PPFs). Any two samples are categorized as 0 if is selected from extraordinary training or categorized as no exchange if from equal magnificence and are paired for training the samples. The newly categorized paired samples are fed into nicely designed architecture of deep CNN for the training manner; during trying out the every pixel in the checking out procedure, neighboring pixel-pairs are built the

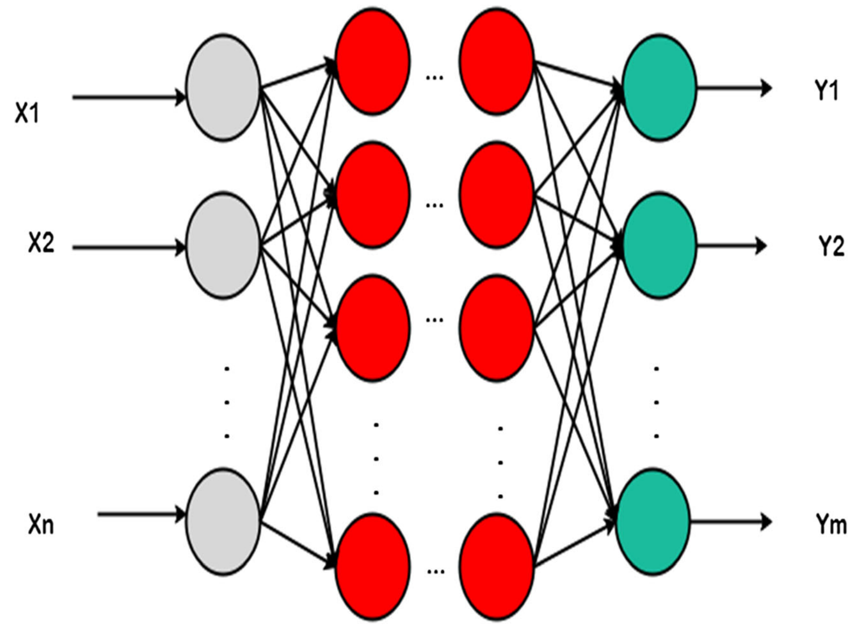


Fig. 1 Neural network construction

use of its environment that are classified through the educated CNN, and the final label is then resolute via a joint category vote casting approach outcomes. The grounds deep CNN is selected due to the reality that CNN has been proved to classify hyper-spectral statistics successfully after building suitable layered structure. This framework operates at the version of pixel-pairing to resolve this trouble, wherein the brand new statistics combination is built through pairing any selected samples from the to be had categorized facts and people records are relabeled. During schooling, the enter statistics grows exhibit in quadratic amount wherein the nicely tuned parameters are being ensured. In this technique HSI absolutely utilizes the inner correlation of buddies which might be ignored with the aid of authentic CNN.

Mingming Xiong, et.al,… [6] has implemented weighted JCR (WJCR) classifier, a next version of Joint Collaborative Representation(JCR) classifier. The same weight is being adopted by JCR classifier to the pixel of neighbors while the spatial and spectral features are extracted. When extracting spatial and spectral features since surrounding pixels JCR adopts the same weights. Unlike JCR, WJCR tries to consume the suitable weights of the pixel by similarity consideration of the pixel at centre and its surroundings. The centre pixel and the neighboring pixel vary in the heterogeneous region as it belongs to different classes. It is suboptimal to choose the pixels with the same weights by JCR in heterogeneous regions. The neighboring pixels that are associated with the centre one is considered under such circumstances. Yet it is not easy to remove the irrelevant pixels when the superfluous computational complexity increases. Therefore, neighboring pixel has been involved with an adaptive weight was effectively implemented. WJCR determines the most suitable weights of pixel using the function named Gaussian kernel. In heterogeneous images, the spatial and spectral features are efficiently extracted by WJCR which provides an advantage of more accuracy.

Lianru Gao, et.al,… [7] has deployed dimensionality reduction technique on the inputted subspace with making use of the available limited training samples. As an illustration, subspace-projection has been developed based on the multinomial logistic regression technique. In this technique the hyper-spectral data are analyzed and are characterized into mixed pixels. For classifying the images of remote sensing, Support vector machine (SVM) has been added to the subspace-projection concept. The subspaces which are connected to each class are constructed using SVM nonlinear functions. The hyper-spectral data set of National Aeronautics and Space Administration's Airborne are collected using Evident /Infrared image spectrometer and that data set are validated practically using SVM sub method [8]. It is difficult to linearly separate the noise from slack variables using Soft margin classification which may allow fault classification. This soft margin method is combined with kernel called the SVM classifier. This SVM classifier is preferably used for high dimension space of class separation using a nonlinear transformation method. When the kernel functions are included with the nonlinear functions of SVM, many different types of SVMs are formed. A binary classifier, SVM has been extended to unravel the multiclass problems. The distributions of classes are estimated consistently by incorporating SVM method to a subspace-projection based approach. SVM sub technique vigorously classifies the occurrences of mixed pixels, noise in the training samples available.

C. Wu, et.al,… [9] proposed a new SFA transform recognition set of rules to discover the existence of actual adjustments in bi-temporal multispectral photographs. In sort to comply SFA with the bi-temporal change recognition problem, the mathematical technique of SFA has been reformulated to detect alternatively. The most invariant component has been extracted from bi-temporal photos and thus the changes in SFA been detected using the encompassed set of rules. The unaffected pixels in the distorted feature space acquire very tiny difference values, and subsequently, the modified areas are quite highlighted. Three variations of the SFA exchange detection set of regulations has been implemented. Firstly an unsupervised SFA (USFA) is designed and implemented to the whole image. To compute the competencies in transformation, supervised SFA (SSFA) uses the unchanged education samples. The iterative SFA (ISFA) has designed finally which accredits immoderate weights towards the unaffected pixels inside the iterative technique, to higher find out modifications. SFA has been implemented successfully to the invariant objects recognition in spatial transformation, nonlinear canopy supply separation, and human motion reputation. On intensive study, the mathematical technique SFA has been reformulated to permit the pattern recognition. The unaffected pixels difference values are not equal to 0 in the unique space, and the reparability between the adjustments and no modifications is not especially authentic. When compared to the modified one, it is really worth noting that the spectral changes of the unchanged pixels are normally small. In the era, one among a type extraordinary weights are assigned to the pixels, as the larger weights to unchanged pixels and smaller weights to changed pixels. The unaffected pixels play the maximum essential position within the mastering approach, and the modified pixels could have a vulnerable impact. It is therefore an automated unchanged pattern preference system.

A. Ertürk,et.al,… [10] has comparatively analyzed Hyper spectral alternate detection by means of spectral immixing, which has the ability to offer approximate interpretable facts without any difficulty based on the nature of the alternate, and also offer sub pixel-stage change facts. Many natural disasters may trade the nature such as a flood, which would amount to a great increase within the abundance of "water" end member in the course of the scene, exchange in the cultivated crop in a farmland, which could bring about the exchange of the distinguished quite member within the area, settlement encroachment, i.e., boom within the abundance of synthetic materials, or goals of hobby, which might suggest new and anomalous give up members inside the scene. However, despite the fact that hyper spectral alternate detection through unmixing has the essential benefit of offering subpixel-degree effects and statistics on the nature of the alternate, it's miles nonetheless in an early degree inside the literature, with most studies being restricted to case research. A preferred framework has been proposed to detect the change via spectral unmixing. An observation is made on the utilization of unmixing pixel and mapping of land-coverage to detect the change in Subpixel-degree. Change detection has been with the aid of spectral unmixing and the subpixel-stage information about the nature of the exchange are presented in it [11]. In this work, sparse unmixing is utilized as the primary factor for trade detection in multi temporal hyper spectral datasets. Sparse unmixing, used along with spectral libraries, circumvents the drawbacks of everyday spectral unmixing, at the same time as nevertheless supplying sub pixel-level detection and records on the nature of the change by means of obtaining a separate change map for each stop member.

B. Demir, et.al,… [12] proposed a unique trade-detection-driven switch approach to replace land-cowl maps by means of classifying far off sensing pics obtained at the identical place in special period (i.e., picture instance cycle). To obtain effective sudden shots from the supplied domain within the instance cycle, a set of consistent preparation is required. The objective area which are labeled are not used in some other photographs. Unlike different literature transport mastering strategies, no extra acceptance on both the resemblance among magnificence dissemination and the existence of the equal set of land-covers instructions inside the two domains are essential. The planned approach aims at significant a dependable instruction set for the objective domain, enchanting expand of the previously accessible expertise at the resource domain. This is executed through applying an unconfirmed revolutionize revealing approach to target and supply domain names, and transferring elegance detected labels of untouched education samples commencing the foundation to the goal area to initialize the objective area instruction set. A new procedure named Active Learning (AL) is used for subsequent optimization of preparation set. Those modified samples are detected and are given precedence for labeling in the early iterations of AL. At the same time, the maximum edifying samples are decided as changed and unchanged unlabeled samples in the final one. The goal pictures are classed finally. Due to extended supply of snap shots, updation of land-cowl maps through magnificence of far off sensing images is an crucial problem. In the identical areas, precise instances such as time series of remote sensor images and temporally shifted pixels are acquired through satellite TV for pc-borne sensors. In landsat achieve and destiny ESA Sentinel missions, information are available unfastened manner and also there is an issue of time series which can be accessed scientifically by every capability customer. Direct supervision on the time series in each image would helps for the updation of Land-cowl mapping. To train the classifier perfectly, a statistical reference method for the ground reliability is required for all the available temporal pix. In conditions based on operations, accumulating an enough wide type of categorized schooling samples for classification of each single photo isn't practical because of the excessive fee and the associated time consuming machine of this project.

J. Meola, et.al,… [13] deployed a complete rate function, in which the avoidance of community minima are ensured through the model parameters record estimation. Parameters that are close to enough to the right values are chosen which may decrease the performance of detection. However, there comes a full estimation problem and convexity evaluation problem due to the occurrence of thousands of parameters consequent in a hessian. Preliminary estimation on sensitivity detection is done to improve the overall performance in the troublesome computation of convexity on hard miles. MB approach shows the high performance in the estimation of initial parameters and it is proven for its degree of robustness. Actual international statistics reality parameters such as illumination situations, reflectance, and shadows are unknown and difficult for collection and estimating it. In an operational state of affairs, a large set of MODTRAN realizations (masses) is probably generated in offline mode which simulates a vast majority of anticipated jogging situations to result the best sensor gathering records. Sizeable realization can be developed by large amount of time as similar to MODTRAN recognition can require 10–20 s to generate it. However, the era of those simplest realizations is wanted as quickly as at the beginning of an exacting venture. Based upon the altitude and sun's position, the idea vectors are created for the sensed data collected and the decisions are made on those subset realizations. Further MB change detection technique finds the proper changes and provides the notion into obstacles and strength. Airborne testing of hyper spectral imagery gives similar real-international outcomes in an eminent geometry of remote sensing. HYDICE airborne imagery are calibrated as a complete statistical model with a body of radiance gadgets. Moreover, the un-calibrated statistics version is accomplished to tower imagery amassed the usage of an AFRL in-house spectrometer. Change aim ground fact does no longer exist in this records set. Changes in target subpixels are precluded that may exist within the scene. Goal truth map which are manually created are recognized in visual and the consequent changes present in it are restricted.

F. Hu, et.al,… [14] proposed scenarios for generating image talents through extracting CNN capabilities from one-of-a-kind layers. In the number one state of affairs, extraction of the activation vectors from absolutely-associated layers are appeared in the very last image functions; in the second state of affairs, dense talents from the last convolutional layer at multiple scales are extracted in which characteristic coding technique is used to encode the dense abilities into international image capabilities. Experimental analysis on two public landscape class datasets in an extensive manner show off that the picture competencies received through the two proposed situations, not withstanding a simple linear classifier, can result in excellent performance and decorate the ultra-modern-day by a massive margin. The convolutional (conv) layers and pooling layers collect the first few tears, and a median degree is found. The convolution layers outputs feature maps, each element of that are calculated by a dot product computation of the neighborhood location number (receptive field), as it's far related to within the given feature maps and a hard and fast weights which are called as filters (or kernels). In massive, an element realistic non-linear activation function is done to those characteristic maps. The pooling layers carry out a down sampling operation alongside the characteristic of the spatial dimensionality maps thru the maximum computation on a neighborhood region. The completely-related (FC) layers eventually have a look at several stacked convolutional and pooling layers, and the remaining absolutely-related layer is a Soft max layer that computes the rankings for every defined magnificence. CNNs transform the input photograph from original pixel values to the very last elegance rankings through the community in a feed beforehand way. Based on the lower back propagation algorithm, the weights parameters in the convolution layer and completely related layers of CNN are well-informed with classic stochastic incline descent.

Exsisting methodologies

Various Neural network approaches are used to classify Hyper-spectral images. Back Propagation neural network and Convolutional neural network are the existing neural networks. These two networks are described as follows:

Back propagation neural network

In the prevailing framework, a pioneering method, Back Propagation Neural Network (BPNN) is carried out to stumble on the exceptional varieties of HSI pixel. To recognize the styles of the facts, to forecast destiny occasions, find answers and to classify the data, Artificial Neural Networks (ANN) can be learnt and therefore be trained. Neural Networks behavioral getting to know is depending on the manner how the individual elements are connected based totally on the computation and by using the strengths of those connections or weights. According to a special mastering rule these weights may be adjusted robotically by way of education the community until it performs the favored mission acceptably. ANN is a supervised studying method, a machine studying algorithm that makes use of recognized dataset; trained dataset. These acknowledged parameters help ANN to formulate predictions. The fundamental additives of a schooling dataset are the enter records alongside their response values. Based at the predictions made on the unusual and untaught dataset response values, supervised gaining knowledge of set of rules is assembled. A take a look at dataset is hired to validate the version. Larger education datasets are preeminently used that allows you to have better predictive power and the capacity to generalize for numerous new datasets, the exceptional manner is to use larger education datasets.

To classify the HSI pixels back propagation set of rules is used. To diminish the objective feature, a not unusual approach to educate the artificial neural networks is Back propagation. It is a supervised getting to know method that is fashioned by means of generalizing many policies. For the development of training set, it requires many inputs to supply the preferred output. This is ordinarily very beneficial in feed-ahead networks (the networks which has no remarks, in simple phrases, it'll don't have any connections within the loop). This term is also an abbreviation for "backward propagation of errors".

Steps in BPP algorithms:

Step 1: The weights and biases are randomly initialized.
Step 2: The training sample is feed.
Step 3: Promulgate the inputs forward; the net input and output of each unit are computed in the hidden layers and output layers.
Step 4: The error is back propagated to the hidden layer.
Step 5: Weights and biases are updated based on the reflection of propagated errors.
To adjust the network's weights and biases without human intervention, Training and mathematical learning functions are used.
Step 6: Terminate the condition.

For pixel classification in land cover detection, BPNN provides less accuracy and need large number iterative steps to classify the pixels. The back propagation layer is shown in Fig. 2.

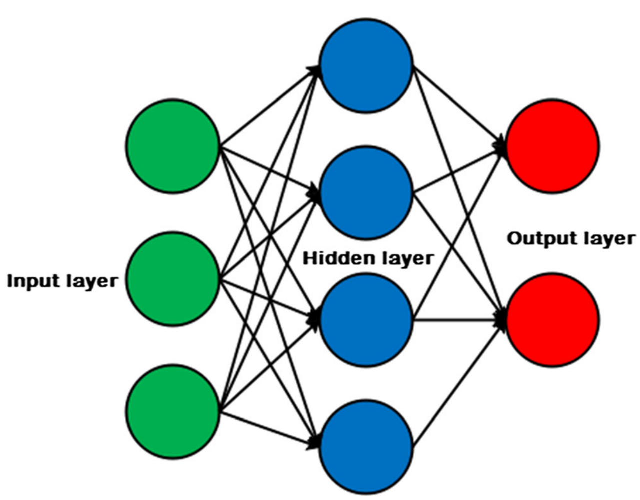


Fig. 2 Back propagation layer

CNN based classification

CNN feed-forward neural community is represented by means of combining the completely related convolutional layers and max. This takes a bonus of spatial nearby correlation with the aid of imposing the patterns of neighborhood connectivity among the adjacent layers neurons. The complex characters of the convolutional layers and easy cells in mammalian seen cortex are mimicked and exchanged with max pooling layers. A complete neural network CNN is framed by means of which includes many pairs of convolution layers and max pooling layers. CNN hierarchical structure is efficiently established by using the analysis of observable representations in an efficient way. Appearance of inter-elegance and the items of shape variations modeling is the primary project in such visual duties [15]. The hyper-spectral statistics are illustrated as 2D curves with hundreds of spectral channels. Differentiating the relative lessons with human eye is difficult, as it is seen that every class curve has its very own visible gravels and self-block adding bricks form which is different from other lessons. CNNs can complete aggressive troubles and provide higher overall presentation rather than human by the use of the spectral signatures; its functionality conjures up to take a look at the possibility of making use of CNNs for HSI class. The CNN varies primarily based on education of the network and on how the convolutional layers and max pooling layers are realized.

Figure 3 normally has five different tiers with weights, the first layer is the input layer and the output layer is the final layer and the layer called convolution layer, which is represented as C1, and the connection layer (Full), which is represented as F1, and finally the pooling layer (Max), which is represented as M2. Let us assume that θ is represented as the parameter, which will be used for training. And in other words it will be represented as $\theta = \{\theta_i\}$ where i is the positive integer and the parameter θ_i will be set between the i^{th} and $(i-1)$th layer. Each and every pixel in the Hyper Spectral Imagery will be treated as the 2D image and the weight of the above said image 2D would be 1(This is considered as similar to the weights of the audio inputs of the speech recognition which has the 1D representation). By considering 2D image weight as 1, the input layer will be having $(n1,1)$ where $n1$ represents the total number of bands. Now, the first hidden layer C1 (convolutional layer) along with its bands $n1$ and weights as 1 filters the input along with 20 kernels. Which has the size 1 per each pixel. It is also represented as $k1 \times 1$ and in total the layer C1 will have 20 $(n2 \times 1)$ and $n2$ is calculated using $1+n1-k1$. The total parameter, which is of trainable kind, that is present in between c1 and input layer is $20 \times (k1 + 1)$. The next layer M2, which is also called as max pooling layer is the secret layer and the size of the kernel is $(k2, 1)$. The total nodes in this hidden layer M2 contains $20(n3 \times 1)$

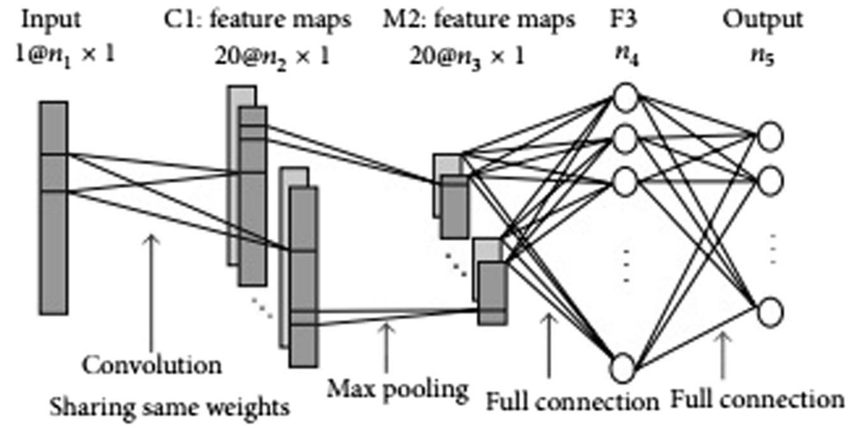


Fig. 3 Convolutional neural network for HSI classification

and here $n3$ is obtained by $n2$/k2. No parameter will be available in this M2 layer. The next part of this tier is F3, which is indicated as connected layer (Fully) and it has $n4$ nodes and it has been represented as $(20 \times n3)$ +1. In between the connected layer (Fully) F3 and the layer M2 there were around $n4$ trainable parameters available. And finally the output node with $n5$ nodes has $(n4 \times 1)$ $n5$ training parameters, which is, found in between F3 and output layer. Accordingly, the design of the proposed CNN classifier will be totally having $20 \times (k1 + 1) + (20 \times n3 + 1) \times n4 + (n4 + 1) \times n5$ trainable parameters. For categorizing a stated HSI pixel requires the analogous CNN along with the aforesaid factors, where $n1$ denotes the channel size of the spectral image and $n5$ denotes the number of output classes present in the data set. In the tryouts, $k1$ is better to be $\lceil n1/9 \rceil$, and the node $n2$ is obtained with the help of $n1-k1+1$. $n3$, which is obtained by $n2$/k2, can be any between 34 and 39, and $k2$ is obtained by using $\lceil n2/n3 \rceil$. $n4$ trainable parameter that is in between F3 and M2 is set to 100. The choices that were taken here, as the tryouts may not considered as the finest but will have the effect on general HSI data. In the design, the hidden convolutional layer C1 and max-pooling layer M2 will be noticed for the input of HSI data as the feature extractor and the connected layer (Fully) F3 is used as a trainable classifier. The real feature of the original data will be extracted from the output received from the subsampling. 20 features are being extracted from each original hyper-spectral in the proposed CNN structure and among those the $n3$ dimensions will be present in each and every feature extracted.

Training of the CNN model

Initialize the followings: 훼- Learning rate, ITERmax- number of maximum iteration, ERRmin-minimum error, BATCHEStraining- training batches, SIZEbatch- batch size.

These networks varies based on channel size of the spectral data and also based on the output classes that has been formed based on the input received from the HSI data. Irregular boundaries separation has been overcome by the proposed work by extracting the spatial and spectral features in hyperspectral image classification. The basic features construction of CNN can be shown in Fig. 4.

Hyper spectral image classification using recurrent neural network

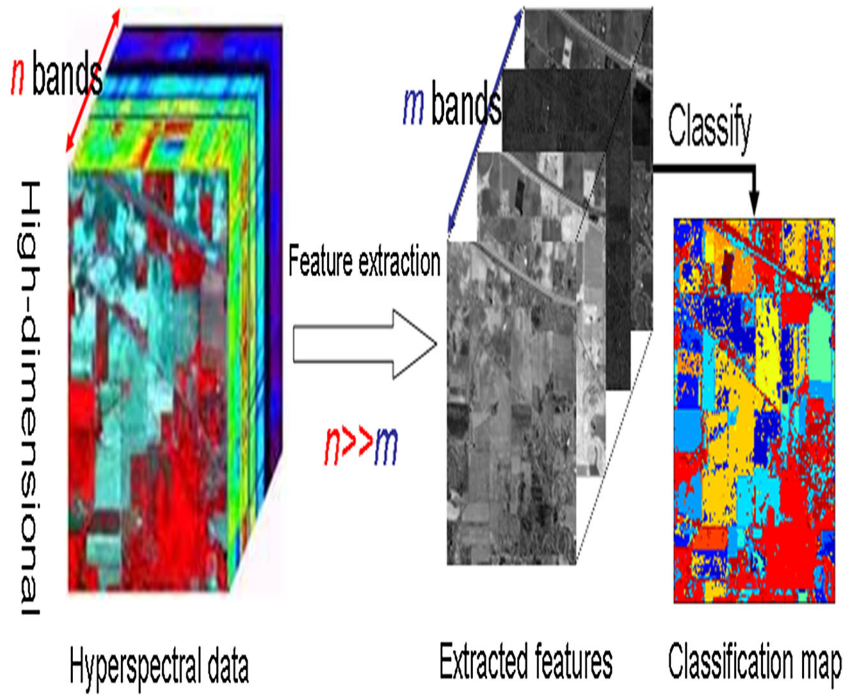
The conventional feed-forward neural network with loops in associations has been extended by an RNN, a class of artificial neural network. Contrasting a feed forward neural network, the sequential inputs are being processed by an RNN in which the activation at each step of the hidden state depends on the previous state. By this approach, the dynamic temporal behavior of the network exhibits. Given a sequence of data, $x = (x_1, x_2, ..., x_T)$, where x_i is the data at i^{th} time step. The recurrent hidden state h_t of the RNN is updated as,

$$h_t = \begin{cases} 0\, if\ t = 0 \\ \varphi(h_{t-1} x_t,)\ Otherwise \end{cases} \tag{1}$$

Where φ is a hyperbolic tangent, nonlinear, logistic sigmoid function. Here, $y = (y_1, y_2, ..., y_T)$ will be considered as output for RNN as an option. For some tasks, only one output, i.e., y_T is needed for classifying an hyper-spectral image. In the conventional RNN model, the implementation of recurrent secret state rule is updated as below:

$$h_t = \varphi(W_{x_t} + U h_{t-1}) \tag{2}$$

Fig. 4 Features extraction in CNN



Where U stands the coefficient matrices for the stimulation of persistent concealed units at the preceding step and where W stands for the coefficient matrices of the current step input. In circumstances, the probability distribution in the sequential data over the next element can be done exemplary by RNN, by giving its present state ht. and also by taking the sequence data of variable length and capturing its distribution. Let the sequence probability can be decomposed into be $p(x_1, x_2,...,x_T)$

$$p(x_1, x_2, \ldots, x_T) = p(x_1), \ldots p(x_T|x_1, x_2, \ldots, x_{T-1}) \tag{3}$$

Then, the conditional probability distribution will be modeled for each and every recurrent network.

$$p(x_t|x_1, \ldots, x_{t-1}) = \varphi(h_t) \tag{4}$$

Where h_t is obtained from (1) and (2). The apparent motivation is to adopt the recurrent network to model the spectral sequence in which the hyper-spectral pixel exploits as sequential data as a substitute of a feature vector. RNN, as an imperative division of the deep learning household, have revealed encouraging results recently in several machine learning and computer vision tasks. Still, it has been witnessed that training

Fig. 5 Recurrent neural network framework

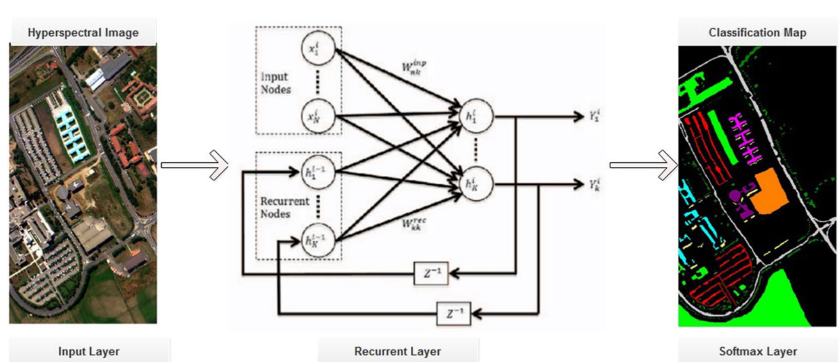


Table 1 Performance table

Algorithm/ Performance measures	Precision	Recall	F- measure
BPNN	42	80	55
CNN	44	82	57
DL-RNN	46	88	60

the RNNs to deal with long-term sequential data has been a difficult task, as the gradients inclined to disappear. One usual approach is to design a new high-level recurrent unit to address this issue. The recurrent result is shown in fig. 5.

Experimental results

In experimental results, from Pavia datasets, the hyper-spectral facts are acquired which are employed to measure the usefulness of the suggested method. Using F-measure, Recall and Precision the performance of the system is being evaluated.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$
$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$
$$\text{F measure} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

The performance evaluation result is shown in following Table 1 and in fig. 6.

Conclusion

Deep learning RNN method for HSI classification has been proposed. Each pixel is considered as sequential data from

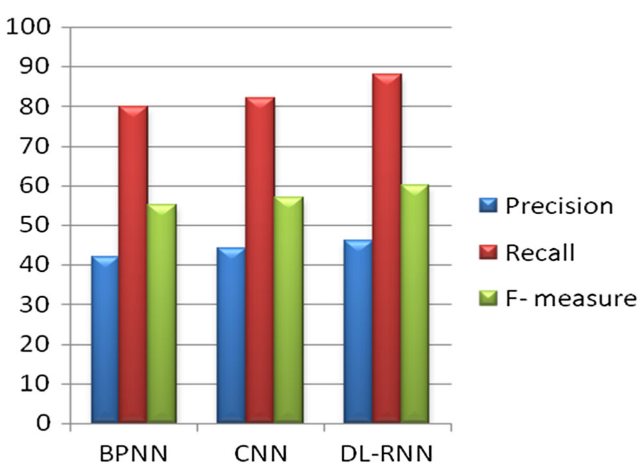


Fig. 6 Performance chart

the observation on hyper-spectral image classification. High learning rates are predicted by this proposed method than the existing neural network approaches. For the first time, the inherent successive data structure of hyper-spectral pixel are considered which represents the tale approach for enhanced understanding, processing and modeling of hyper-spectral data. To authenticate the features of deep RNN profusely for hyper-spectral image processing, in the future, broaden experiments may be steered for providing more precise analysis on transfer learning in HSI data analysis and also in the change detection in the remote sensing applications. In Future, by means of implementing various deep learning approaches it can be extended to improve the accuracy in classification.

Compliance with Ethical Standards

Conflict of Interest This paper has not communicated anywhere till this moment, now only it is communicated to your esteemed journal for the publication with the knowledge of all co-authors.

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Moran, M. S., Inoue, Y., and Barnes, E. M., Opportunities and limitations for image-based remote sensing in precision crop management. *Remote Sens. Environ.* 61(3):319–346, 1997.
2. Lu, X., Li, X., and Mou, L., Semi-supervised multitask learning for scene recognition. *IEEE Trans. Cybern.* 45(9):1967–1976, 2015.
3. Wang, Q., Lin, J., and Yuan, Y., Salient band selection for hyperspectral image classification via manifold ranking. *IEEE Trans. Neural Netw. Learn. Syst.* 27.6:1279–1289, 2016.
4. Camps-Valls, G. et al., Advances in hyperspectral image classification: Earth monitoring with statistical learning methods. *IEEE Sign. Process. Mag.* 31.1:45–54, 2014.
5. Li, W. et al., Hyperspectral image classification using deep pixel-pair features. *IEEE Trans' Geosc. Remote Sens0'* 55(2):844–853, 2017.
6. Xiong, M. et al., Hyperspectral image classification using weighted joint collaborative representation. *IEEE Geosci. Remote Sensing Lett.* 12.6:1209–1213, 2015.
7. Gao, L. et al., Subspace-based support vector machines for hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.* 12.2:349–353, 2015.
8. Olmanson, L. G., Brezonik, P. L., and Bauer, M. E., Airborne hyperspectral remote sensing to assess spatial distribution of water quality characteristics in large rivers: The Mississippi River and its tributaries in Minnesota. *Remote Sens. Environ.* 130:254–265, 2013.
9. Wu, C., Du, B., and Zhang, L., Slow feature analysis for change detection in multispectral imagery. *IEEE Trans. Geosci. Remote Sens.* 52(5):2858–2874, 2014.
10. Ertürk, A., Iordache, M.-D., and Plaza, A., Sparse unmixing-based change detection for multitemporalhyperspectral images. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* 9(2):708–719, 2015.

11. Li, X., Mou, L., and Lu, X., Scene parsing from an MAP perspective. *IEEE Trans. Cybern.* 45(9):1876–1886, 2015.
12. Demir, B., Bovolo, F., and Bruzzone, L., Updating land-cover maps by classification of image time series: A novel change-detection-driven transfer learning approach. *IEEE Trans. Geosci. Remote Sens.* 51(1):300–312, 2013.
13. Meola, J., Eismann, M. T., Moses, R. L., and Ash, J. N., Application of model-based change detection to airborne VNIR/SWIR hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* 50(10):3693–3706, 2012.
14. Hu, F., Xia, G.-S., Hu, J., and Zhang, L., Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sens.* 7(11):14680–14707, 2015.
15. Zhu, Q., Zhong, Y., Zhao, B., Xia, G.-S., and Zhang, L., Bag-of-visualwords scene classifier with local and global features for high spatial resolution remote sensing imagery. *IEEE Geosci. Remote Sens. Lett.* 13(6):747–751, 2016.