HYPERSPECTRAL IMAGE CLASSIFICATION USING DEEP LEARNING

Suba Lakshmi.E
MPhil computer science
manonmaniam sundaranar university

ABSTRACT

The traditional unsupervised loss function like mean square error (MSE) calculates the distance between the predicted value and the original input. However, it is difficult to guarantee the effectiveness of the features only by optimizing their construction error. In order to make the learned features more effective for classification tasks, we optimize the contrastive loss function to make the features from different views of the same sample consistent. This makes the features of the same class aggregate with each other, and the features of different classes are far away from each other. Therefore, the features obtained by optimizing the contrastive loss function of different views could effectively improve the classification accuracy. We use a deep CNN as the base feature extractor. We call this proposed method deep multiview learning. Therefore, the proposed method belongs to the category of unsupervised learning, which could alleviate the lack of labeled training samples. Finally, a conventional machine learning method (e.g., support vector machine) is used to complete the classification task in the learned latent space. To demonstrate the effectiveness of the proposed method, extensive experiments are carried on four widely used hyperspectral data sets. The experimental results demonstrate that the proposed method could improve the classification accuracy with small samples.
CHAPTER 1

INTRODUCTION

1.1 GENERAL

RemotesensingisthescienceofacquiringinformationabouttheEarth'ssurface without actually being in contact with it. This is done by sensing and recording reflected or emitted radiation from different bodies. Objects having different surface features reflect or absorb the sun's radiation in different ways. The reflectance properties of an object depend on the particular material and its physical and chemical state (e.g. moisture), the surface roughness as well as the geometric circumstances (e.g. incidence angle of the sunlight). The most important surface features are colour, structure and surface texture. These differences make it possible to identify different earth surface features or materials by analysing their spectral reflectance patterns or spectral signatures. These signatures can be visualized in so called spectral reflectance curves as a function of wavelength.

The primary prerequisite for remote sensing is to have an energy source to light up the target (unless the sensed energy is being radiated by the target). This energy is known as electromagnetic radiation. All electromagnetic radiation has key properties and carries on in unsurprising routes as indicated by the fundamental laws of wave hypothesis.

The electromagnetic spectrum ranges from the shorter wave lengths (including gamma X-ray) to the more extended wave lengths (including microwaves and telecast radio waves). There are few areas of the electromagnetic range which are helpful for remote sensing. Ultraviolet or UV portion of the spectrum has the shortest wavelength that can be used for remote sensing for most purposes. Ultraviolet or UV portion of the spectrum has the shortest wavelength that can be used for remote sensing for most purposes.

This radiation extends beyond the violet portion of the visible wavelengths and is therefore called. Some earth surface materials, essentially shakes and minerals, fluoresce or radiate certain light when lit up by ultraviolet radiation. Electromagnetic waves utilized as a part of remote sensing is demonstrated in Figure.1.1.
In remote sensing, the process involves an interaction between incident radiation and the targets of interest. This is exemplified by the use of imaging systems where the following seven elements are involved.\[1\] Note, however, that remote sensing also involves the sensing of emitted energy and the use of non-imaging sensors.

**Figure 1.1 Electromagnetic Spectrum**

Energy Source or Illumination (A) - the first requirement for remote sensing is to have an energy source which illuminates or provides electromagnetic energy to the target of interest.

Radiation and the Atmosphere (B) - as the energy travels from its source to the target, it will come in contact with and interact with the atmosphere it passes through. This interaction may take place ease second time as the energy travels from the target to the sensor.

Interaction with the Target (C) - once the energy makes its way to the target through the atmosphere, it interacts with the target depending on the properties of both the target and the radiation.

Recording of Energy by the Sensor (D) - after the energy has been scattered by, or emitted from the target, it is recorded by the sensor.
Transmission, Reception, and Processing (E) - the energy recorded by the sensor has to be transmitted, often in electronic form, to a receiving and processing station where the data are processed into an image (hardcopy and/or digital).

Interpretation and Analysis (F) - the processed image is interpreted, visually and/or digitally or electronically, to extract information about the target which was illuminated.

Application (G) - the final element of the remote sensing process is achieved when the information is able to extract from the imagery about the target in order to better understand it, reveals some new information, or assists in solving a particular problem.

1.2 Types of Remote Sensing

The sun is a source of energy or radiation, which provides a very convenient source of energy for remote sensing. The sun's energy is either reflected, as it is for visible wavelengths, or absorbed and then reemitted, as it is for thermal infrared wavelengths. The remote sensing system can be classified into two types depending on the source of energy: passive remote sensing and active remote sensing.

Passive Remote Sensing: Passive sensors can only be used to detect energy when the naturally occurring energy is available. For all reflected energy, this can only take place during the time when the sun is illuminating the Earth. There is no reflected energy available from the sun at night. Energy that is naturally emitted (such as thermal infrared) can be detected day or night, as long as the amount of energy is large enough to be recorded. Examples of passive remote sensors include film photography, infrared, and radiometers. Passive remote sensing is illustrated in Figure 1.3a.

Active Remote Sensing: Active sensors, on the other hand, provide their own energy source for illumination. The sensor emits radiation which is directed toward the target to be investigated. The radiation reflected from that target is detected and measured by the sensor. Advantages of active...
sensors include the ability to obtain measurements anytime, regardless of the time of day or season. Active sensors can be used for examining wavelengths that are not sufficiently provided by the sun, such as microwaves, or to better control the way a target is illuminated. However, active systems require the generation of a fairly large amount of energy to adequately illuminate targets. Examples of active sensors are laser fluoro sensor and Synthetic Aperture Radar (SAR). Active remote sensing is illustrated in Figure 1.3.b

1.3 Types of Optical Remote Sensing Systems

Depending on the number of spectral bands used in the imaging process, optical remote sensing systems are classified into the following types:

1. **Panchromatic image:** The sensor is a single channel radiation sensitive detector within a wide range of wavelengths. If the wavelength range coincides with the visible range, then the resulting image resembles a black-and-white photograph taken from space. The physical quantity being measured is the apparent brightness of the targets. The spectral information or color of the targets is lost.

2. **Multispectral image:** The sensor is a multichannel detector with a few spectral bands. Each channel is sensitive to radiation within a narrow wavelength band. The resulting image is a multilayer image which contains both the brightness and spectral (color) information of the targets being observed.

3. **Superspectral Image:** It has many more spectral channels (typically >10) than a multispectral sensor. The bands have narrower bandwidths, enabling the finer spectral characteristics of the target to be captured by the sensor.

4. **Hyperspectral Image:** A hyperspectral image consists of hundreds or more contiguous spectral bands forming a three-dimensional (two spatial dimensions and one spectral dimension) image cube.

5. **Ultraspectral Image:** It contains thousands of spectral bands offering the capability to extend spectral imaging to a high level.

1.3.1 Advantages of Remote Sensing

The major advantages of remote sensing are:

1. **Synoptic view:** Remote sensing process facilitates the study of Earth's various features in their spatial relation to each other and helps to trace the required features and circumstances.

2. **Accessibility:** Remote sensing process makes it possible to accumulate information about the unreachable area when it is not possible doing ground survey like in mountainous regions.
3. Time: The information areas or foreign lands about a large area can be gathered quickly, the techniques save time and efforts of human beings/machine.

4. Cost savings: The costs are relatively small when compared with the benefits, which can be obtained from interpretation of satellite imagery.

5. Coverage: With the use of high-altitude sensor platforms, it is now possible to record extensive areas on a single image. The advent of high-flying aircraft and satellites, single high-quality image covers thousands of square miles.

1.4 Limitations of Remote Sensing

The disadvantages of remote sensing are:

- Requires cross verification with ground (field) survey data
- Data analysis and interpretation problems
- Cost of data collection and data purchase
- Possibilities form misclassification or confuse of objects
- Potential limitations with the different sensors' spatial, spectral and temporal resolutions.

1.4.1 Applications of Remote Sensing

Satellite data enables our renewable and non-renewable resources to be properly managed as it provides timely and detailed Earth surface information. Remote sensing finds applications in the following fields.

- Urban Planning
- Geographic information
- Weather and agricultural forecasts and assessment of environment and natural disaster
- Image processing
- Aerial traffic control, Interferometric synthetic aperture radar
- Laser and Radar altimeters
- Precision geo-referencing
- Ultrasound (acoustic) and radar tide gauges
- Light detection and ranging
- Radiometers and photometers
- Stereographic pairs of aerial photographs
- Mineralogy, Biology, Defense, and Environmental measurements
1.4.2 Hyperspectral Imaging

The word “hyper” in hyperspectral means “over” as in “too many” and refers to the large number of measured wavelength bands. Hyperspectral images are spectrally overdetermined; they provide ample spectral information to identify and distinguish between spectrally similar (but unique) materials. Consequently, hyperspectral imagery provides the potential for more accurate and detailed information extraction than is possible with other types of remotely sensed data. A Hyperspectral Image (HSI), in general, has hundreds of spectral bands in contrast to a normal digital image which has three spectral bands (blue, red, and green) and thus offers a more complete part of the light spectrum for viewing and analysis. In general, hyperspectral sensors measure bands at 10 to 20 nm intervals. A regular digital image can be viewed as a collection of three-dimensional spectral vectors, each representing the information for one pixel. Similarly, a HSI can be viewed as a collection of multidimensional spectral vectors, each representing the information for one pixel.

Hyperspectral remote sensing images acquire many, very narrow, contiguous spectral bands throughout the visible, near-infrared, mid-infrared and thermal infrared positions of the electromagnetic spectrum. Hyperspectral sensors typically collect 200 or more bands enabling the construction of an almost continuous reflectance spectrum for every pixel in the scene. Contiguous narrow bandwidths characteristic of hyperspectral data allows for in-depth examination of earth surface features which would otherwise be lost within the relatively coarse bandwidths acquired with multispectral scanners.[3] Over the past decade, extensive research and development has been carried out in the field of hyperspectral remote sensing. With commercial airborne hyperspectral imagers such as Compact Airborne Spectrographic Imager (CASI) and Hymap and the launch of satellite-based sensors such as Hyperion HSI, hyperspectral technology is fast moving into the mainstream of remote sensing and applied remote sensing research studies. Hyperspectral images have found many applications in water resource management, agriculture, and environmental monitoring. It is important to remember that there is not necessarily a difference in spatial resolution between hyperspectral and multispectral data but rather in their spectral resolutions.

Hyperspectral images typically include spectral bands representing the ultraviolet (200-400 nm), visible (400-700 nm), near infrared (700-1000 nm), and short-wave infrared (1000-4000 nm). Thus, HSI are favoured over regular images for some applications such as forestry and crop analysis, mineral exploration, and surveillance. Hyperspectral image cube structure is illustrated in Figure 1.4. Each pixel has intensity values corresponding to all spectral bands as shown in Figure 1.5.
Analyzing hyperspectral data becomes a difficult task. Important factors are making it too complex such as atmospheric corrections, huge size and large volume of the image, curse of dimensionality, spatial/spectral signatures variability, few labelled samples, exploring the spatial correlation among pixels and adding contextual information along with spectral information during classification.

General processing of hyperspectral data involves the following steps:

- Data pre-processing
- Correction of data by using atmospheric correction
- Dimensionality reduction
- Pure end-member selection using pixels

Perform classification using selected end-members. Hyperspectral imaging is commonly referred to as spectral imaging or spectral analysis. The distinction between hyper- and multi-spectral is sometimes based on an arbitrary "number of bands" or on the type of measurement, depending on what is appropriate to the purpose.
Multispectral imaging deals with several images at discrete and somewhat narrow bands. Being “discrete and somewhat narrow” is what distinguishes multispectral in the visible from color photography. A multispectral sensor may have many bands covering the spectrum from the visible to the longwave infrared. Multispectral images do not produce the “spectrum” of an object. Landsat is an excellent example of multispectral imaging.

**Hyperspectral Image Sensors**

Hyperspectral and multispectral sensors are based on the same physical technology. They both record radiance in the Visible to Near-Infrared (VNIR) and Short-Wave Infrared (SWIR) of the spectrum, VNIR spanning 400–1000 nm and SWIR 1000–2400 nm. Unlike multispectral sensors, such as Landsat-8 (11 bands), recording in a fairly limited number of discrete spectral bands (4–20 bands), Hyperspectral sensors include a very large number of contiguous and narrow spectral bands of 5–15 nm (Kaufmann et al. 2009). Airborne Hyperspectral sensors provide promising results for many applications as they combine a high spectral resolution with a high spatial resolution and are not so affected by atmospheric perturbation (Lu et al. 2013; Wang et al. 2010). These platforms have played a key role in the development of Hyperspectral science and applications (Kruse et al. 2003; Guanter et al. 2012). With the availability of emblematic sensors such as HyMAP, CASI, Airborne Visible/InfraRed Imaging Spectrometer (AVIRIS), Digital Airborne Imaging Spectrometer (DAIS), Reflective Optics System Imaging Spectrometer (ROSIS), Airborne Imaging Spectrometer for Applications (AISA), Hyperspectral Digital Imagery Collection Experiment (HYDICE), Multispectral Infrared Visible Imaging Spectrometer (MIVIS), etc., Hyperspectral research quickly expanded the number of Hyperspectral applications in vegetation monitoring, water resources management, geology, and land cover (Govender et al. 2007; Van der et al. 2012). However, they do not allow regular and synoptic coverages over large areas as spaceborne sensors. Moreover, spaceborne sensors produce images with lower angular effects due to their much smaller field of view. Fig. 1.6 illustrates the timeline of high-spatial-resolution (≤ 30 m) hyperspectral sensors. These hyperspectral sensors have been implemented on a number of experimental airborne platforms, including the HYDICE and the AVIRIS. Earth Observation-1 (EO-1) carries a hyperspectral sensor called Hyperion.

**Figure 1.6 Timeline highlighting hyperspectral Imaging Sensors**
Table 1.1 Details of Hyperspectral Sensors

<table>
<thead>
<tr>
<th>Type of Sensors</th>
<th>Name of the Sensors</th>
<th>No. of bands</th>
<th>Spectral Range (μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satellite Sensor</td>
<td>FTHSI on MightySatII</td>
<td>256</td>
<td>0.35 to 1.05</td>
</tr>
<tr>
<td></td>
<td>Hyperion on EO-1</td>
<td>242</td>
<td>0.40 to 2.50</td>
</tr>
<tr>
<td></td>
<td>AVIRIS (Airborne Visible Infrared Imaging Spectrometer)</td>
<td>224</td>
<td>0.40 to 2.50</td>
</tr>
<tr>
<td></td>
<td>HYDICE (Hyperspectral Digital Imagery Collection Experiment)</td>
<td>210</td>
<td>0.40 to 2.50</td>
</tr>
<tr>
<td></td>
<td>PROBE-1</td>
<td>128</td>
<td>0.40 to 2.50</td>
</tr>
<tr>
<td>Airborne Sensor</td>
<td>CASI (Compact Airborne Spectrographic Imager)</td>
<td>Over 228</td>
<td>0.40 to 1.00</td>
</tr>
<tr>
<td></td>
<td>HyMap</td>
<td>100 to 200</td>
<td>Visible to Thermal Infrared</td>
</tr>
<tr>
<td></td>
<td>EPS-H (Environmental Protection System)</td>
<td>VIS/NIR (76)</td>
<td>VIS/NIR (0.43 to 1.05)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SWIR1 (32)</td>
<td>SWIR1 (1.50 to 1.80)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SWIR2 (32)</td>
<td>SWIR2 (2.00 to 2.50)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TIR (12)</td>
<td>TIR (8.00 to 12.50)</td>
</tr>
<tr>
<td></td>
<td>DAIS 7915 (Digital Airborne Imaging Spectrometer)</td>
<td>VIS/NIR (32)</td>
<td>VIS/NIR (0.43 to 1.05)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SWIR1 (8)</td>
<td>SWIR1 (1.50 to 1.80)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SWIR2 (32)</td>
<td>SWIR2 (2.00 to 2.50)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MIR (1)</td>
<td>MIR (3.00 to 5.00)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TIR (6)</td>
<td>TIR (8.70 to 12.30)</td>
</tr>
<tr>
<td></td>
<td>(AISA) Airborne Imaging Spectrometer for Applications</td>
<td>Over 288</td>
<td>0.43 to 1.00</td>
</tr>
</tbody>
</table>

The sensors are typically measured in spectral resolution, which is the width of each band of the spectrum that is captured. If the scanner detects a large number of fairly narrow frequency bands, it is possible to identify objects even if they are only captured in a handful of pixels. In the hyperspectral field, there are two types of systems that take images: on aircraft and on satellites. Most hyperspectral sensors are mounted on aerial platforms than on the satellite.
1.5 HYPERSONICIMAGECLASSIFICATION

Hyperspectral image classification is the process in which individual values (objects/patterns/image regions/pixels) are grouped based on the similarity between the value and the description of the group. Hyperspectral image classification can be done by either based on pixel information or based on the use of training samples. Based on pixel information, images can be classified as Per-Pixel, Sub Pixel, Per-field, Knowledge based, Contextual and multiple Classifiers. Based on the use of training samples, images can be classified as Supervised Classification, Unsupervised Classification and Semisupervised Classification. Hyperspectral image classification is based on the detection of the spectral response pattern of land cover classes. The major objective of the image classification procedure is to automatically categorize all pixels in the image into appropriate land cover classes. The intent of the classification process is to categorize all pixels in an image into one of several land cover classes or “themes”. This categorized data is then used to produce thematic maps of the land cover present in an image. One of the major problems in Hyperspectral remote sensing is a high amount of data that is available for processing. Due to the huge amount of data, the processing time and classification accuracy are decreased. To deal with this huge data problem, the valuable information and more processing are required to increase the classification accuracy. Therefore, classification of HSI data without losing important information about objects of interest is important.

1.6 DEEP LEARNING IN REMOTE SENSING

Deep Learning (DL) is a type of machine learning in which a model learns to perform classification tasks directly from images, text, or sound. DL is usually implemented using a neural network architecture. The term “deep” refers to the number of layers in the network—the more layers, the deeper the network. Traditional neural networks contain only 2 or 3 layers, while deep networks can have hundreds.

In recent years, DL has become an emerging learning method in big data analysis and has been extensively used in numerous fields, such as natural language processing (Ronan & Weston, 2008), image classification, speech enhancement, due to its exceptional performance compared to other conventional learning algorithms.

Recent advances in Artificial Intelligence (AI) and machine learning, especially the emerging field of deep learning, have changed the way we process, analyse and manipulate geospatial sensor data. This is largely driven by the wave of excitement in deep machine learning, as a new frontier of AI, where the most representative and discriminative features are learnt end-to-end, hierarchically. DL methods have achieved huge success not only in classical computer vision tasks, such as target detection, visual recognition, and robotics, but also in many other practical applications (Hu et al., 2015). They have made considerable improvements beyond the state-of-the-art records in a variety of domains, and have attracted great interest in both academia and industrial communities.

DL offers a different outlook on feature learning and representations, where robust, abstract and invariant features are learnt end-to-end, hierarchically, from raw data (e.g., image pixels) to semantic labels, which is a key advantage in comparison with previous state-of-the-
Many deep learning-based methods have been proposed, including deep belief networks (DBNs), deep Boltzmann machines (DBMs), Stacked Autoencoder (SDE), and deep convolutional neural networks. Amongst them, the CNN model represents the most well-established method, with impressive performance and great success in the field of computer vision and pattern recognition, such as for visual recognition (Krizhevsky et al. 2012), image retrieval and scene annotation.

1.6.1 Deep Learning for HSI Classification

Classification is the task of labelling pixels (or regions in an image) into one of several classes. The DL methods outlined as follows use many forms of DL to learn features from the data itself and perform classification at state-of-the-art levels.

As DL has emerged as one of the well-known machine learning techniques, it is widely used in the field of computer vision and image processing, with applications such as image classification (He et al. 2014; Krizhevsky et al. 2012), object detection (Girshick et al. 2014), and super-resolution restoration (Dong et al. 2016). In recent past, DL is used for remote sensing image classification, and a good number of relative papers are discuss it the literature. As a part of this survey, in this section, it is presented the pixel-wise and scene-wise remote sensing image classification approaches that are based on DL, supported with comparative experimental analyses.

HSI data classification is of major importance to RS applications, so many of the DL results reviewed were based on HSI classification. HSI processing has many challenges, including high data dimensionality and usually low numbers of training samples. Chen et al. based HSI classification framework. The input data are converted to a one-dimensional (1-D) vector and processed via a DBN with three RBM layers, and the class labels are output from a two-layer logistic regression NN. A spatial classifier using Principal Component Analysis (PCA) on the spectral dimension followed by 1-D flattening of a 3-D box, a three-level DBN, and two-level logistic regression classifier. A third architecture uses combinations of the 1-D spectrum and the spatial classifier architecture. He et al. (2016) developed a DBN for HSI classification that does not require SGD training. Nonlinear layers in the DBN allow for the nonlinear nature of HSI data and a logistic regression classifier is used to classify the outputs of the DBN layers. A parametric depth study showed that nine layers produced the best results of depths from 1 to 15, and after a depth of nine, no improvement resulted by adding more layers.

Some of the HSI DL approaches use both spectral and spatial which integrates spatial information. Small training sets are mitigated by a collaborative, representation-based classifier and salt-and-pepper noise is mitigated by a graph-cut-based spatial regularization. Their method is more efficient than comparable kernel-based methods, and the collaborative representation-based classification makes their system relatively robust to small training sets. Yang et al. (2016) use a two-channel CNN to jointly learn spectral and spatial features. Transfer learning is used when the number of training samples is limited, where low-level and mid-level features are transferred from other scenes. The network has a spectral CNN and spatial CNN, and the results are combined in three fully connected layers. A softmax classifier produces the final class labels. Pan et al. (2017) proposed the so-called rolling guidance filter and vertex component analysis network (R-VCANet), which also attempts to solve the common problem
of lack of HSI training data. The network combines spectral and spatial information. The small details from imagery. The VCANet is a combination of vertex component rolling guidance filter, an edge-preserving filter used to remove noise and analysis, which is used to extract pure endmembers, and PCANet. A parameter analysis of the number of training samples, rolling times, and the number and size of the convolution kernels is discussed. The system performs well even when the training ratio is only 4%. Lee & Kwon (2016) designed a contextual deep fully convolutional DL network with 14 layers that jointly exploit spatial and HSI spectral features. Variable size convolutional features are used to create a spectral–spatial feature map. A feature of the architecture is the initial layers use both $[3 \times 3 \times B]$ convolutional masks to learn spatial features, and $[1 \times 1 \times B]$ for spectral features, where $B$ is the number of spectral bands. The system is trained with a very small number of training samples (200/class).

**Objective of the research**

The main objectives of the research are to:

- Develop deep learning techniques for the analysis and classification of remote sensing hyperspectral images.
- Investigate the behavior and performance, in terms of overall accuracy, average accuracy and kappa coefficient of the newly developed techniques with standard hyperspectral data sets.
- Produce accurate classification maps that are suitable to meet the practical requirements for the applications of interest.

**Motivation of the research**

HSI classification plays an important role in the earth observation technology using data from Remote Sensing (RS), which has been extensively used in both military and civil fields. However, RS image classification performance faces major scientific and practical challenges due to the characteristics of RS data such as high dimensionality and relatively small quantities of available labelled samples. In recent years, as new DL techniques emerge, approaches to RS image classification with DL have achieved significant breakthroughs, offering novel opportunities in classification. Specifically, focus is on unsolved challenges and opportunities as they relate to (i) inadequate datasets, (ii) human-understandable solutions for modelling physical phenomena, (iii) big data, (iv) transfer learning, (v) DL architectures and learning algorithms for spectral, spatial, and temporal data, (vi) non-traditional heterogeneous data sources, (vii) better understanding of DL systems theoretically, (viii) high barrier to entry, and (ix) training and optimizing the DL.

**Problem definition**

Hyperspectral classification becomes a difficult task because of high dimensionality, few labelled samples, spatial variability of the spectral signature, spatial correlation among pixels and addition of contextual information with spectral information during classification. It uses distinctive features like spectral, spatial multi temporal and multi sensor information. Important factors in classification accuracy are uncertainty and error propagation chain. [6] For achieving significant improvement in accuracy, weakest links in the chain need to be identified and then the uncertainties are reduced.
Based on the literature survey, it is inferred that still there is a scope for new Hyperspectral Image Classification Algorithm in deep learning area to improve the classification accuracy. Thus the problem identified for this research work is improvement in Hyperspectral Image Classification with respect to overall accuracy, average accuracy and kappacoefficient.

Hyperspectral Image Classification Algorithm using Multiscale Convolutional Neural Network

In this work, a novel hyperspectral image classification system that uses a Multiscale Convolutional Neural Network with Gaussian Kernel (MCNN-GK) has been highlighted. Although CNNs have successfully been used in remote sensing scene classification, the scale of the objects can change greatly between images. When the scale of the image changes a lot within the dataset, it is very difficult to achieve a good classification of remote sensing data. To solve this problem, a novel MCNN framework with Gaussian convolution kernel function has been proposed as it is the only correct kernel function to approximate scalespace.

In MCNN structure, three fully connected layers are added in front of the output layer, which have been abandoned in many current CNN structures. The weights of convolution layer are initialized as Gaussian convolution kernels. The Gaussian smoothing layer adjusts the size of the scale by training, so that the entire learning process is carried out in a stable scale space. Different learning rates are employed in each octave in the training process and adjusted according to the change of the scales. In the training process of the traditional CNNs, the weights in the front hidden layers are more difficult to train than the hidden layers behind it. Hence, a larger learning rate is used on the small scale front hidden layer, and a smaller learning rate corresponding to the large scale hidden layer. For training, as in CNN, the loss function is chosen as cross-entropy, and mini-batch gradient descent is used to find the best parameters of the network. Training an neural network to find the best parameters (weights of the network) to minimize the loss function, which in a classification task measures the compatibility between a prediction (e.g., the class scores) and the ground truth label. The experiments show that the proposed method outperforms in terms of overall accuracy, average accuracy and kappacoefficient.

DATASETS

The various benchmark data sets of HSI are generally utilized for assessing the performance of the proposed methods for diverse fields of application. The data set include Indian Pines, Pavia University, Pavia Center, Kennedy Space Center, Botswana, Salinas, Salinas-A and Washington images. In this work, the experimental results are exposed on three hyperspectral airborne images recorded by the AVIRIS and the ROSIS sensors, with different contexts (agricultural and urban areas), different spatial resolutions. These three datasets are detailed in the following:

**Indian Pines:** It was acquired by the AVIRIS sensor in Indiana in June 1992. It is the first data set with 20-m resolution image taken over the Indian Pines test site in June 1992. The image size is 145 × 145 pixels and contains 220 spectral bands. Twenty water absorption bands have been removed (Tadjudin & Landgrebe, 1999), and a 200-band image was used for the experiments. It contains two-thirds agriculture, and one-third forest are the natural perennial vegetation. A ground survey of 10366
pixels, distributed in 16 crop types classes, is available. This dataset is a classical benchmark to validate model accuracy and is known to be very challenging because of the strong mixture of the classes’ signatures, since the image has been acquired shortly after the crops were planted.

**Pavia University:** The Pavia University image was collected by the Reflective Optics System Imaging Spectrometer (ROSIS) sensor over the urban area of the University of Pavia, Italy. It consists of 103 spectral bands with a spectral range from 430 nm to 860 nm. The images' spatial resolution is 1.3 m, and the total image size is 610 × 340 pixels. The reference data contain nine classes of interest.

**Salinas:** The Salinas image was acquired via the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) over Salinas Valley, California, and the image size is 512 × 217, with the spatial resolution of 3.7 m. It contains 224 spectral bands. Like the Indian Pine scene, the 20 water absorption bands were discarded and there maining 204 bands were utilized for the experiments. The ground reference data for the Salinas image entails 16 classes. It includes vegetables, bare soils, and vineyard fields.

### CHAPTER 3

**METHODOLOGY**

### 3.1 Classify Hyperspectral Images Using Deep Learning:

Hyperspectral image classification is the process in which individual values (objects, patterns/image regions/pixels) are grouped based on the similarity between the value and the description of the group. Hyperspectral image classification can be done by either based on pixel information or based on the use of training samples. Based on pixel information, images can be classified as Per-Pixel, Sub-Pixel, Per-field, Knowledge based, Contextual and multiple Classifiers. Based on the use of training samples, images can be classified as Supervised Classification, Unsupervised Classification, and Semi-supervised Classification.

Hyperspectral image classification is based on the detection of the spectral response pattern of land cover classes. The major objective of the image classification procedure is to automatically categorize all pixels in the image into appropriate land cover classes. The intent of the classification process is to categorize all pixels in an image into one of several land cover classes or "themes". This categorized data is then used to produce thematic map of the land cover present in an image. One of the major problems in Hyperspectral remote sensing is a high amount of data that is available for processing. Due to the huge amount of data, the processing time and classification accuracy are decreased. To deal with this huge data problem, the valuable information and more processing are required to increase the classification accuracy. Therefore, classification of HSI data without losing important information about objects of interest is important.

Hyperspectral imaging measures the spatial and spectral features of an object at different wavelengths ranging from ultraviolet through long infrared, including the visible spectrum. Unlike color imaging, which uses only three types of sensors sensitive to the red, green, and blue portions of the visible spectrum, hyperspectral images can include dozens or hundreds of channels. Therefore, hyperspectral images...
can enable the differentiation of objects that appear identical in an RGB image.

**Load Hyperspectral Data Set:**
This example uses the Indian Pines dataset, included with the Image Processing Toolbox™ Hyperspectral Imaging Library. The data set consists of a single hyperspectral image of size 145-by-145 pixels with 220 color channels. The data set also contains a ground truth label image with 16 classes, such as Alfalfa, Corn, Grass-pasture, Grass-trees, and Stone-Steel-Towers.

**Preprocess Training Data:**
Reduce the number of spectral bands to 30 using the hyperpca function. This function performs principal component analysis (PCA) and selects the spectral bands with the most unique signatures.

**Specify Training Options:**
Specify the required network parameters. For this example, train the network for 100 epochs with an initial learning rate of 0.001, a batch size of 256, and Adam optimization.

**Train the Network:**
By default, the example downloads a pretrained classifier for the Indian Pines data set. The pretrained network enables you to classify the Indian Pines data set without waiting for training to complete.

To sum up, the successful mapping of TSS distribution in mulberries suggested that the application of hyperspectral imaging to realize the visualization of mulberry fruits’ internal quality is feasible and promising. The PLSR and LS-SVM model based on 23 and 11 wavelengths had a good performance to predict TSS of mulberries, which indicated that RF algorithm was effective in reducing three-dimensional data. PLSR-RF based on 23 important wavelengths provided the optimal visualization results. It could be revealed that PLSR was feasible to map chemical component concentration (TSS) distribution of mulberry fruits. This research provided a theoretical basis for developing the instrument for measuring the internal quality of fruits and made it possible to sort mulberries based on TSS spatial distribution.

3.2 PCA Features
PCA is an important method for feature extraction and image representation. In PCA, matrix transformation of the image takes place into high-dimensional vectors, and its covariance matrix is obtained. Consuming high-dimensional vectors, PCA is a dimensionality reduction technique that has four main parts: feature covariance, eigen decomposition, principal component transformation, and choosing components in terms of explained variance. The purpose of this blog is to share a visual demo that helped the students understand the final two steps.

Principal component analysis (PCA) is a technique for reducing the dimensionality of such datasets, increasing interpretability but at the same time minimizing information loss. It does so by creating new uncorrelated variables that successively maximize variance.

The following represents six steps of principal component analysis (PCA) algorithm:

- **Standardize the dataset:** Standardizing / normalizing the dataset is the first step one would need to take before performing PCA. The PCA calculates a new projection of the given data set representing one or more features. Thenew axes are based on the standard deviation of the value of these features. So, a feature / variable with a high standard deviation will have a higher weight for the...
calculation of axis than a variable / feature with a low standard deviation. If the data is normalized / standardized, the standard deviation of all features / variables get measured on the same scale. Thus, all variables have the same weight and PCA calculates relevant axis appropriately. Note that the data is standardized / normalized after creating training / test split. Python’s `sklearn.preprocessing.StandardScaler` class can be used for standardizing the dataset [12].

- Construct the covariance matrix: Once the data is standardized, the next step is to create a covariance matrix where n is the number of dimensions in the dataset. The covariance matrix stores pairwise covariances between the different features. Note that a positive covariance between two features indicates that the features increase or decrease together, whereas a negative covariance indicates that the features vary in opposite directions. Python’s `numpy.cov` method can be used to create covariance matrix.

- Perform eigendecomposition of the covariance matrix: Then the step is to decompose the covariance matrix into its eigenvectors and eigenvalues. The eigenvectors of the covariance matrix represent the principal components (the directions of maximum variance), whereas the corresponding eigenvalues will define their magnitude. `numpy.linalg.eig` or `linalg.eigh` can be used for decomposing the covariance matrix into eigenvectors and eigenvalues.

- Selection of most important Eigenvectors / Eigenvalues: Sort the eigenvalues by decreasing order to rank the corresponding eigenvectors. Select k eigenvectors, which correspond to the k largest eigenvalues, where k is the dimensionality of the new feature subspace. One can use the concept of explained variance to select the k most important eigenvectors.

- Projection matrix creation of important eigenvectors: Construct a projection matrix, W, from the top k eigenvectors.

- Training / test dataset transformation: Finally, transform the d-dimensional input training and test dataset using the projection matrix to obtain the new k-dimensional feature subspace.

Here are the steps followed for performing PCA:

- Perform one-hot encoding to transform categorical dataset to numerical dataset
- Perform training / test split of the dataset
- Standardize the training and test dataset
- Construct covariance matrix of the training dataset
- Construct eigendecomposition of the covariance matrix
- Select the most important features using explained variance
- Construct projection matrix; in the code below, the projection matrix is created using the five eigenvectors that correspond to the top five eigenvalues (largest), to capture about 75% of the variance in this dataset
- Transform the training dataset into new feature subspace

### 3.3 Deep Features Classification

CNN is a neural network that extracts input image features and another neural network classifies the image features. The input image is used by the feature extraction network. The extracted feature signals are utilized by the neural network for classification.

Hyperspectral imaging measures the spatial and spectral features of an object at different wavelengths ranging from ultraviolet through long infrared, including the visible spectrum. Unlike color imaging, which uses only three types of sensors sensitive to the red, green, and blue portions of the visible spectrum, hyperspectral images...
can include dozens or hundreds of channels. Therefore, hyperspectral images can enable the differentiation of objects that appear identical in an RGB image.

**LoadHyperspectralDataSet**

The data set consists of a single hyperspectral image of size 145-by-145 pixels with 220 color channels. The data set also contains a ground truth label image with 16 classes, such as Alfalfa, Corn, Grass-pasture, Grass-trees, and Stone-Steel-Towers.

**PreprocessTrainingData**

Reduce the number of spectral bands to 30 using the `hyperpca` function. This function performs principal component analysis (PCA) and selects the spectral bands with the most unique signatures. Split the hyperspectral image into patches of size 25-by-25 pixels with 30 channels using the `create ImagePatchesFromHypercube` helper function. This function is attached to the example as a supporting file. The function also returns a single label for each patch, which is the label of the central pixel.

**CreateCSCNNClassificationNetwork**

Define the CSCNN architecture.

**SpecifyTrainingOptions**

Specify the required network parameters. For this example, train the network for 100 epochs with an initial learning rate of 0.001, a batch size of 256, and Adam optimization.

**ClassifyHyperspectralImageUsingTrainedCSCNN**

Calculate the accuracy of the classification for the test dataset. Here, accuracy is the fraction of the correct pixel classification overall the classes.
3.4 ARCHITECTURE DESIGN

The high dimensionality of hyperspectral image data and the lack of labeled samples can lead to the Hughes phenomenon. Earlier in the research of hyperspectral image classification, people often focused on spectral information, using only spectral information to achieve image classification, and developed many classification methods, such as support vector machine (SVM), random forest (RF), neural networks, and Polynomial logistic regression. Dimension reduction methods such as feature extraction and feature selection have also been proposed, such as principal component analysis (PCA), independent component analysis (ICA), and linear discriminant analysis (LDA). The other is a nonlinear feature extraction method. For example, in 2000, the local linear embedding (LLE) algorithm published by Science of Roweis and Saul in Science projects local high-dimensional data points into a low-dimensional coordinate system. The overall information is obtained by superimposing local neighborhoods, maintaining the spatial topological relationship, and retaining the overall geometric properties. At the same time, Tenenbaum et al. proposed (Isometric Feature Mapping, ISOMAP) an algorithm based on the classic MDS. It uses geodesic distance to embed high-dimensional data into low-dimensional coordinates. The neighborhood structure between high-dimensional spatial data points is still retained in low-dimensional coordinate space. Belki and Niyogi proposed a similar pull to LLE in 2001. Laplacian Eigenmap (LE), also known as Spectral Clustering
It is worth noting that deep learning has excellent capabilities in image processing. Especially in recent years, image classification, target detection, and other fields have set off a wave of deep learning. Some deep learning network models have been used in remote sensing image processing, such as the Convolutional neural network (CNN), deep belief network (DBN), and recurrent neural network (RNN). Moreover, in order to solve the problem of poor classification results due to the lack of training samples, a new tensor-based classification model was proposed. Experiments confirmed that this method is superior to support vector machines and deep learning when the number of training samples is small.

3.6 PROPOSED SYSTEM

The paramount challenge for HSI classification is the curse of dimensionality which is also termed as Hughes phenomenon. To confront with this difficulty, feature extraction methods are used to reduce the dimensionality by selecting the prominent features. In unsupervised methods, the algorithm or method automatically groups pixels with similar spectral characteristics (means, standard deviations, etc.) into unique clusters according to some statistically determined criteria. Further, unsupervised classification methods do not require any prior knowledge to train the data. The familiar unsupervised methods are principal component analysis (PCA) and independent component analysis (ICA).

Principal component analysis

It is the most widely used technique for dimensionality reduction. In comparative sense, appreciable reduction in the number of variables is possible while retaining most of the information contained by the original dataset. The substantial correlation between the hyperspectral bands is the basis for PCA. The analysis attempts to eliminate the correlation between the bands and further determines the optimum linear combination of the original bands accounting for the variation of pixel values in an image.

CHAPTER 4

EXPERIMENTAL RESULTS

Implementation of a software package refers to the installation of the package in its real environment to the full satisfaction of the users and operating system. In short, implementation constitutes all activities that are required to put an already tested and completed package into operation. The success of any information system lies in its successful implementation.

4.1 DATASETS

The various benchmark data sets of HSI are generally utilized for assessing the performance of the proposed methods for diverse fields of application. The dataset include Indian Pines, Pavia University, Pavia Center, Kennedy Space Center, Botswana, Salinas, Salinas-A and Washington images. In this work, the experimental results are exposed on three hyper spectral air borne images recorded by the AVIRIS and the ROSIS sensors, with different contexts (agricultural and urban areas), different spatial resolutions.

4.2 Functional Documentation

Functional Documentation plays a vital role in describing the various functionalities of the project. Basically, it considers the various forms designed for the project and explains various functions associated with the form. As a matter of fact, each form is an integrated part of the project and has its
In this section, I explain the functional documentation of the project. It considers various blocks of the modules and the associated forms.

Experiment on Indian Pines dataset: The Indian Pines dataset was gathered by AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) sensor over the Indian Pines test site in North-western Indiana in 1992. The Indian Pines (IP) dataset has images with $145 \times 145$ spatial dimension and 224 spectral bands, and the ground truth is designated into 16 classes of vegetation. The experiments were conducted on the Indian Pines dataset with different numbers of training and testing samples. Some other experiments we performed involved observing the effects of different spatial window sizes and the effects of number of PCA components. We found 30 to be the optimal number of PCA components for this particular dataset. It is also observed that the proposed method outperforms most of the state-of-the-art methods.

![Classification Map for Indian Pines](image)

**Figure 4.1** The Classification Map for Indian Pines
Figure 4.1 The Classification Map for Salinas Scene

Figure 4.2 The Classification Map for Pavia University
The three datasets are detailed in the following:

**Indian Pines:** It was acquired by the AVIRIS sensor in Indiana in June 1992. It is the first dataset with 20-m resolution image taken over the Indian Pines test site in June 1992. The image size is 145 × 145 pixels and contains 220 spectral bands. Twenty water absorption bands have been removed (Tadjudin & Landgrebe, 1999), and a 200-band image was used for the experiments. It contains two-thirds agriculture, and one-third forest as natural perennial vegetation. A ground survey of 10366 pixels, distributed in 16 crop types classes, is available.[10] This dataset is a classical benchmark to validate model accuracy and is known to be very challenging because of the strong mixture of the classes’ signatures, since the image has been acquired shortly after the crops were planted.

**Pavia University:** The Pavia University image was collected by the Reflective Optics System Imaging Spectrometer (ROSIS) sensor over the urban area of the University of Pavia, Italy. It consists of 103 spectral bands with a spectral range from 430 nm to 860 nm. The image spatial resolution is 1.3 m, and the total image size is 610 X 340 pixels. The reference data contain nine classes of interest.

**Salinas:** The Salinas image was acquired via the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) over Salinas Valley, California, and the image size is 512 X 217, with the spatial resolution of 3.7 m. It contains 224 spectral bands. Like the Indian Pines scene, the 20 water absorption bands were discarded, and there maining 204 bands were utilized for the experiments. The ground-referenced data for the Salinas image entails 16 classes. It includes vegetables, bare soils, and vineyard fields.
### Table 1.2: CSCNN Performance with TEM Iteration

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1 Score</th>
<th>Time Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>83.52</td>
<td>80.36</td>
<td>91.84</td>
<td>85.71</td>
<td>0.8351</td>
</tr>
<tr>
<td>2</td>
<td>81.32</td>
<td>80.12</td>
<td>85.11</td>
<td>82.47</td>
<td>0.8131</td>
</tr>
<tr>
<td>3</td>
<td>76.92</td>
<td>73.08</td>
<td>84.44</td>
<td>78.35</td>
<td>0.7692</td>
</tr>
<tr>
<td>4</td>
<td>79.12</td>
<td>75.44</td>
<td>89.58</td>
<td>81.9</td>
<td>0.7912</td>
</tr>
<tr>
<td>5</td>
<td>80.22</td>
<td>78.33</td>
<td>90.38</td>
<td>83.93</td>
<td>0.8021</td>
</tr>
<tr>
<td>6</td>
<td>85.71</td>
<td>84.62</td>
<td>89.8</td>
<td>87.13</td>
<td><strong>0.8571</strong></td>
</tr>
<tr>
<td>7</td>
<td>83.52</td>
<td>86.27</td>
<td>84.62</td>
<td>85.44</td>
<td>0.8351</td>
</tr>
<tr>
<td>8</td>
<td>81.32</td>
<td>86.79</td>
<td>82.14</td>
<td>84.4</td>
<td>0.8131</td>
</tr>
<tr>
<td>9</td>
<td>83.52</td>
<td>78.95</td>
<td>93.75</td>
<td>85.71</td>
<td>0.8351</td>
</tr>
<tr>
<td>10</td>
<td>76.92</td>
<td>71.43</td>
<td>88.89</td>
<td>79.21</td>
<td>0.7692</td>
</tr>
</tbody>
</table>

The above table shows the performance of the CSCNN such as accuracy, precision, recall, and F1-score with time complexity. The overall accuracy of the CSCNN is up to 85.71%, precision of the CSCNN is 84.62%, recall of the CSCNN is 89.8%, F1-score is 87.13%, and time complexity is 0.8571 milliseconds.

The above chart shows the accuracy-wise chart. The accuracy has obtained up to 96%.
The above chart shows the precision-wise chart. The precision has obtained up to 90%.

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1score</th>
<th>TimeComplexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>92.42</td>
<td>77.78</td>
<td>96.25</td>
<td>85.96</td>
<td>0.8341</td>
</tr>
<tr>
<td>2</td>
<td>91.32</td>
<td>75.47</td>
<td>91.92</td>
<td>82.47</td>
<td>0.5248</td>
</tr>
<tr>
<td>3</td>
<td>94.81</td>
<td>87.27</td>
<td>91.57</td>
<td>82.88</td>
<td>0.8562</td>
</tr>
<tr>
<td>4</td>
<td>91.32</td>
<td>81.13</td>
<td>96</td>
<td>83.5</td>
<td>0.8041</td>
</tr>
<tr>
<td>5</td>
<td>94.62</td>
<td>86.21</td>
<td>99.29</td>
<td>81.67</td>
<td>0.8641</td>
</tr>
<tr>
<td>6</td>
<td>92.42</td>
<td>82.45</td>
<td>90.33</td>
<td>82.96</td>
<td>0.8421</td>
</tr>
<tr>
<td>7</td>
<td>95.91</td>
<td>88.46</td>
<td>91.25</td>
<td>81.46</td>
<td>0.8196</td>
</tr>
<tr>
<td>8</td>
<td>92.42</td>
<td>78.18</td>
<td>91.45</td>
<td>82.45</td>
<td>0.8412</td>
</tr>
<tr>
<td>9</td>
<td>94.62</td>
<td>80</td>
<td>95.69</td>
<td>80.2</td>
<td>0.8259</td>
</tr>
<tr>
<td>10</td>
<td>91.32</td>
<td>86.44</td>
<td>95</td>
<td>85.71</td>
<td>0.8031</td>
</tr>
</tbody>
</table>

The above table shows the performance of the PCA CSCNN such as accuracy, precision, recall and f1-score with time complexity. The overall accuracy of the PCA CSCNN is up to 95.71%, precision of the PCA CSCNN is 84.62%, Recall of the PCA CSCNN is 99.8%, f1-score is 87.13% and time complexity is 0.8571 milliseconds.
The above chart depicts the performance of the PCA CSCNN such as accuracy, precision, recall and f1-score with time complexity.

CONCLUSION

Classification and recognition of hyperspectral images are important content of hyperspectral image processing. This paper discusses several methods of hyperspectral image classification, including supervised and unsupervised classification and semisupervised classification. Although the supervised and unsupervised classification methods described in this article have their respective advantages to varying degrees, there are limitations in the application of various methods. For example, supervised classification requires a certain number of prior conditions, and human factors will affect the classification results have an impact. Therefore, based on different application requirements, combined with the acquisition of hyperspectral images with massive information, multiple methods need to be combined with each other in order to achieve the desired classification effect. With the development of hyperspectral image technology, hyperspectral image classification has been widely used. Existing theories and methods still have certain limitations for more complicated hyperspectral image classification. Therefore, researching more targeted hyperspectral image classification methods will be an important research direction in the future.