Modeling Earth Systems and Environment (2023) 9:1977–1987
https://doi.org/10.1007/s40808-022-01608-y

ORIGINAL ARTICLE

Crop type classification with hyperspectral images using deep learning : a transfer learning approach

Usha Patel[1,2] · Mohib Pathan[1] · Preeti Kathiria[1] · Vibha Patel[3]

Received: 22 September 2022 / Accepted: 6 November 2022 / Published online: 22 November 2022

Abstract

Crop classification plays a vital role in felicitating agriculture statistics to the state and national government in decision-making. In recent years, due to advancements in remote sensing, high-resolution hyperspectral images (HSIs) are available for land cover classification. HSIs can classify the different crop categories precisely due to their narrow and continuous spectral band reflection. With improvements in computing power and evolution in deep learning technology, Deep learning is rapidly being used for HSIs classification. However, to train deep neural networks, many labeled samples are needed. The labeling of HSIs is time-consuming and costly. A transfer learning approach is used in many applications where a labeled dataset is challenging. This paper opts for the heterogeneous transfer learning models on benchmark HSIs datasets to discuss the performance accuracy of well-defined deep learning models—VGG16, VGG19, ResNet, and DenseNet for crop classification. Also, it discusses the performance accuracy of customized 2-dimensional Convolutional neural network (2DCNN) and 3-dimensional Convolutional neural network (3DCNN) deep learning models using homogeneous transfer learning models on benchmark HSIs datasets for crop classification. The results show that although HSIs datasets contain few samples, the transfer learning models perform better with limited labeled samples. The results achieved 99% of accuracy for the Indian Pines and Pavia University dataset with 15% of labeled training samples with heterogeneous transfer learning. As per the overall accuracy, homogeneous transfer learning with 2DCNN and 3DCNN models pre-trained on the Indian Pines dataset and adjusted on the Salinas scene dataset performs far better than heterogeneous transfer learning.

Keywords Hyperspectral images (HSIs) · Transfer learning (TL) · Homogeneous transfer learning · Heterogeneous transfer learning · Pre-trained models · Deep neural network

Introduction

Crops play a vital role in human diets, biofuel, and climate change in the environment which in turn affect the national/global economic development and social stability (Stergioulas et al 2022; Xie et al. 2021). The temporal variation in the crops and their spatial distribution information plays an important role in better agriculture development. Each country's state and national government requires agriculture statistics for the human diets and food security for their country (Xie et al. 2021). So, crop classification and periodic monitoring help the government's decision-making process of the government (Stergioulas et al 2022), (Woldemariam et al. 2022). Traditional and manual methods for gathering this information are time-consuming and laborious (Ajadi et al. 2021). With the help of available satellite data from various hyperspectral remote sensors, crop classification and monitoring can be easier in terms of time, cost, and labor (Ajadi

✉ Usha Patel
ushapatel@nirmauni.ac.in

✉ Preeti Kathiria
preeti.kathiria@nirmauni.ac.in

Mohib Pathan
20mcec15@nirmauni.ac.in

Vibha Patel
vibhadp@gmail.com

[1] CSE Department, Institute of Technology, Nirma University, Ahmedabad, India

[2] Gujarat Technological University, Ahmedabad, India

[3] IT Department, Vishwakarma Government Engineering College, Gujarat Technological University, Ahmedabad, India

et al. 2021). Hyperspectral sensors capture the target's image over a wide spectrum range. Due to this, the hyperspectral image (HSIs) dataset contains spatial-spectral information in abundance. Hyperspectral image data is widely being used in several remote sensing applications (Khan et al. 2018; Transon et al. 2018) such as agricultural (Adam et al. 2010), medicine (Carrasco et al 2003), environmental monitoring and forestry (Ismail et al. 2016), (Zimmer et al. 2022). Due to the advanced remote sensors, high-resolution HSIs images are available for land cover classification. The list of widely known HSIs sensors, along with their various spectral features, is provided by Paoletti et al. in their paper (Paoletti et al. 2019). As a result of the abundance of spatial-spectral information, HSIs images can have hundreds of channels or even more at high resolution. Due to this feature of HSIs images is very useful in identifying different substances and discriminating them from each other in crops.

Some of the Unsupervised (Lopez et al. 2021), supervised machine learning (Cai and Shang 2021), and deep learning methods were proposed for the HSIs classifications. Deep learning(DL) models outperform in the field of HSIs classification and have recently outperformed almost every existing machine learning model. As deep learning models must be trained on a large dataset to achieve the best performance, Hyperspectral image datasets generally have very few samples available in the collection. Also, the DL model's performance depends on features extracted from the labeled dataset. However, labeling HSIs images for training is costly and time-consuming, so a transfer learning approach is used in many applications where the labeled dataset is challenging. After the commencement of the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) (Krizhevsky et al 2012), many researchers have proposed transfer learning models.

Transfer learning is an approach that allows the knowledge acquired from one task to be used for the other similar task. It is broadly categorized into Homogeneous transfer learning and Heterogeneous transfer learning approach. In the homogeneous transfer learning approach, the target and source datasets have the same attributes, labels, and dimensions, but both differ in marginal distribution. The target dataset features and labels differ from the source dataset in the Heterogeneous transfer learning approach.

Let us define (D_s), a dataset using which the DL model was trained. (X_s) refers to the attributes of (D_s) and (Y_s) is its corresponding labels. (D_s) contains enough labeled samples to train the deep network. Another dataset with a limited number of labeled samples is (D_t). (X_t) refers to the attributes of (D_t) and its corresponding label is (Y_t). Generally, the marginal distribution of source and target are not the same $P(X_s) \neq P(X_t)$. Thus, the transfer learning process uses the knowledge acquired from source data (D_s) to classify the labels of the target data (D_t) samples. During the fine-tuning, (D_s) is used to train the model, and (D_t) is used to adjust the model. Here, we do not train a model from scratch, so the number of labeled samples required to train the DL model can reduce.

DL models contain various parameters, and if directly trained on target data (D_t) with these features, it might overfit the data and give poor performance. So, to reach a nontrivial state, fine-tuning the parameters may provide better performance. After that, even a small number of samples from target data can be used to tune the model and achieve the optimal classification result (Mughees and Tao 2016).

Various researchers have proposed extended Convolutional Neural Network (CNN) models for HSIs classification using transfer learning. He et al. (2019) have proposed the heterogeneous transfer learning based on CNN (HT-CNN) method applied to two heterogeneous datasets using transfer learning, while Liu et al. (2020) have proposed shuffled group convolutional neural network (SG-CNN) to improve the HSIs classification accuracy. Again, the same authors Liu et al. (2020) have customized the 1D-CNN architecture and applied it to HSIs classification using transfer learning. So, motivated by this, we concluded to use 2DCNN and 3DCNN for homogeneous transfer learning and well-defined pre-trained models for heterogeneous transfer learning in crop classification. Also, the authors Patel et al. (2022b) have used the VGG16 pre-trained model and heterogeneous transfer learning approach for HSI classification and achieved reasonable accuracy. Based on that, to extend the same work, homogeneous and heterogeneous transfer learning is applied to HSIs datasets for crop classification. The issue with using transfer learning models for HSIs classification is the ILSVR (ImageNet) dataset on which these models were trained only three channels, but hyperspectral image data contains hundreds of channels. Therefore, we need to take care of these issues before using a pre-trained model for the HSIs classification task. So, this paper presents a hyperspectral image classification method with transfer learning to solve the problem of limited training samples and simultaneously boost performance for crop type classification.

The major contribution of the paper is as follows:

1. The heterogeneous and homogeneous transfer learning is proposed for HSIs classification, combining pre-train models, HSIs datasets, and customized CNN models.
2. To mitigate the channel difference between the two data sets, the dimensionality reduction method—Principle Component Analysis (PCA) is used to reweight the feature maps.
3. Three widely used HSIs datasets and two fine-tuning strategies (Homogeneous and Heterogeneous) are used to comprehensively evaluate the proposed approach for crop classification.

The rest of the paper is distributed as follows: Section "Related work" throws light on the related work done by other researchers in the area of HSIs classification by transfer learning and deep learning models. Section "Methodology" gives the pathway used for HSIs dataset classification in two ways: heterogeneous transfer learning and homogeneous transfer learning. The experiment discussion for this work is spoken in Section "Eperiments discussion", and the result discussion of the work is highlighted in Section "Results Discussion". Section "Conclusion" concludes the paper.

Related work

Whenever we talk about HSIs classification till now, the availability of the number of training samples is limited. To deal with this issue, several researchers are working in this field. From the available remote sensing techniques, Kureel et al. have used hyperspectral band information to model the health analysis for HSI data of the Lunar forest of Maharashtra using ENVI software. They have applied various vegetation models to compare and identify the index to select the best accuracy. The vegetation health analysis is done with the help of six different vegetation indices (Kureel et al. 2022).

Deep learning-based approaches for HSIs classification

Recently, many researchers have used deep learning methods for HSIs classification with transfer learning and even, and some authors have proposed variations in deep learning methods to provide promising results. The deep learning model DenseNet is used by the authors Patel et al. for classifying crop images into two classes: droop and healthy (Patel et al 2022a).

HSIs images must have rich spectral-spatial information through their widely available narrow bands. The selection of key features is challenging due to noise and band correlation. In the paper (Roy et al. 2020), Roy et al. have suggested an attention-based adaptive spectral-spatial kernel improved residual network (A2S2K-ResNet) with spectral attention to capturing discriminative spectral-spatial features for HSIs classification, improving the classification performance. The Autoencoder model plays a prominent role in extracting spectral-spatial key features. The author Banerjee et al. explained the basic understanding of the various Autoencoder-based model in their paper (Banerjee et al 2020). Liang et al. (2018) have presented a deep multi-scale spectral-spatial feature fusion method using VGG-16 and cooperative sparse autoencoder models for extracting spectral-spatial key features. Then using, SVM classifier classification was done. For the same work, the author Lin et al. (2018) have explored the correlation of the features between the source and target level domains used for HSIs classification. To find out the correlation, Auto-encoder is fine-tuned on both domains. The other authors, Munishamaiaha et al. (2022), have employed SE-AB-DenseNet to discriminate spatial-spectral features. The AdaBound optimizer is opted to improve the stability and classification accuracy. So, Feature extraction from the HSIs dataset pays more attention before using classification.

CNN for HSIs classification

In recent years many methods have been proposed for HSIs classification. Across all, CNN-based methods have shown higher performance than many state-of-the-art models and methods. Zhang et al. (2019) have proposed lightweight 3DCNN, which has a dense network, requires fewer parameters and less computational cost, and gives competitive performance compared to 3DCNN models. To deal with the "limited number of samples" problem of HSIs classification, they proposed two strategies—cross sensor and cross model-applied to three publically available datasets captured using different sensors. For the same cross-sensor transfer learning HSIs classification, a 3-D separable ResNet(3-D-SRNet) is used by Jiang et al. (2019). The proposed model requires less computation and parameters and provides better classification. Liu et al. (2020) have also employed lightweight shuffled group CNN to address the problem of limited samples in HSIs classification.

In the paper, Yang et al. (2018) have applied 1D DenseNet for extracting spectral features and 2D DenseNet for spatial domain. They have achieved good HSIs classification by applying three deep neural network methods on two benchmark datasets. He et al. (2019) have used CNN-based heterogeneous transfer learning for HSIs classification. The ImageNet dataset is used for training, and the attention mechanism is utilized for feature adjustment of four heterogeneous datasets.

VGG16 for HSIs classification

A lightweight model of VGG16 was offered by Ye et al. (2021), which selects some needed features from remote sensing images and accurately classifies the images. For offering the lightweight model, they modified the VGG16 and combined the complete convolution network with it, and finally, after fine-tuning the decoding part of the network, the integrated network is used for classification. Transfer learning is not used for it. Patel et al. (2022b) have used the VGG16 pre-trained model and heterogeneous transfer learning approach for HSI classification and achieved reasonable accuracy.

Methodology

In this research, the transfer learning approach is used in two ways for the HSIs classification—heterogeneous transfer learning and Homogeneous transfer learning. For heterogeneous transfer learning, well-defined pre-trained models VGG16, VGG19, ResNet, and DenseNet are used for HSIs classification, while in a homogeneous way, 2DCNN and 3DCNN models are customized and used for classification.

Transfer learning involves two steps: First, the model will be trained on the source data with plenty of labeled samples to make the model parameters arrive at a good state, and then the model is tuned for the target data distribution.

Heterogeneous transfer learning on HSIs dataset for classification

The ImageNet dataset contains labeled RGB images of around 1000 classes used for the computer vision task. Using the ImageNet dataset, the selected DL models are pre-trained, and then heterogeneous transfer learning is applied to the HSI dataset. A classification task has been performed with the help of well-defined pre-trained DL models. The trained DL models on the ImageNet dataset contain input as three-channel, wherein the HSIs image comprises hundreds of channels in the input image (Vaddi and Manoharan 2020).

Some well-defined and widely used DL models in image processing are VGG-16, VGG-19, ResNet, and DenseNet. VGG-16 and VGG-19 are the first models proposed at the ILSVR competition, and these models are also less complex (in terms of the number of parameters) than other models. However, they both suffer from "The vanishing gradient problem." ResNet was proposed after VGG-16 and VGG-19. The number of parameters to learn is more in ResNet than in VGG-16 and VGG-19. ResNet overcame the "Vanishing gradient problem" using a skip connection. DenseNet is the extension of the ResNet model. The number of parameters to learn in DenseNet is even more prominent than in ResNet. Hence, these four models with complex architecture can provide detailed input image features. Thus, VGG-16, VGG-19, ResNet, and DenseNet were considered for the experiments. The brief discussion about each model is given as follows:

- **VGG-16**
  VGG-16 was first presented in 2014 and was the winner of the ILSVR (ImageNet) Competition. It is considered one of the best convolution neural network (CNN) architectures. The architecture of VGG-16 does not have numerous hyper-parameters. Instead, the researchers used a 3×3 filter CONV layer with stride one and a max pool layer of 2×2 for down-sampling and the same padding throughout the network. For classification, at last, two fully connected hidden layers followed by a dense layer for output (Breiman 2001).
- **VGG-19** The architecture of VGG-19 has a CONV layer of 3×3 filter and a max pool layer of 2×2 for down-sapling. For classification, at last, a layer has two fully connected layers followed by a softmax output of 1000 units. Each output unit represents one of the classes of images from the ImageNet dataset. To use VGG-19 as an effective feature extractor, we focused on the first five blocks of the architecture (Kavitha and Arivazhagan 2010).
- **ResNet** The problem with VGG-16 and VGG-19 is that they both suffer from vanishing gradient issues. ResNet can solve the vanishing gradient problem with several chain rule applications. When the network is too deep, the gradients do not shrink too much as the gradients can directly flow backward to initial filters through the skip connections (Xu et al. 2014).
- **DenseNet** DenseNet architecture is the extension of ResNet architecture. In ResNet architecture, consider the difference between previous and current layers to force the model to learn residuals. Contrary to that, the DenseNet model concatenates the outputs from previous layers rather than using the summation (Mou et al. 2017).

Table 1 contains a comparative analysis of the pre-trained models based on various used parameters of all the opted models for crop classification.

Figure 1 reflects the pathway used for heterogeneous transfer learning on the HSIs dataset using well-defined pre-trained models.

As shown in Fig. 1, VGG-16, VGG-19, ResNet, and DenseNet models are trained using the ImageNet dataset. ImageNet dataset contains three-channel color images, which provides three-channel images as the input dimension. But the HSIs images have hundreds of bands, so Principle component analysis (PCA) is used as dimensionality reduction for the HSIs images, and the first three PCA components are used as the input image for the DL models. Then, HSIs classification is done on Indian Pines(IP), Pavia University (PV), and the Salinas Scene (SL) datasets using the opted models.

Table 1 Comparison of pre-trained models with their used parameters for crop classification using heterogeneous transfer learning

Model	Conv layers	Pool layers	FC layers	Activation function
VGG16	12	5	3	Softmax
VGG19	14	5	3	Softmax
ResNet	48	2	1	Softmax
DenseNet	120	4	1	Softmax

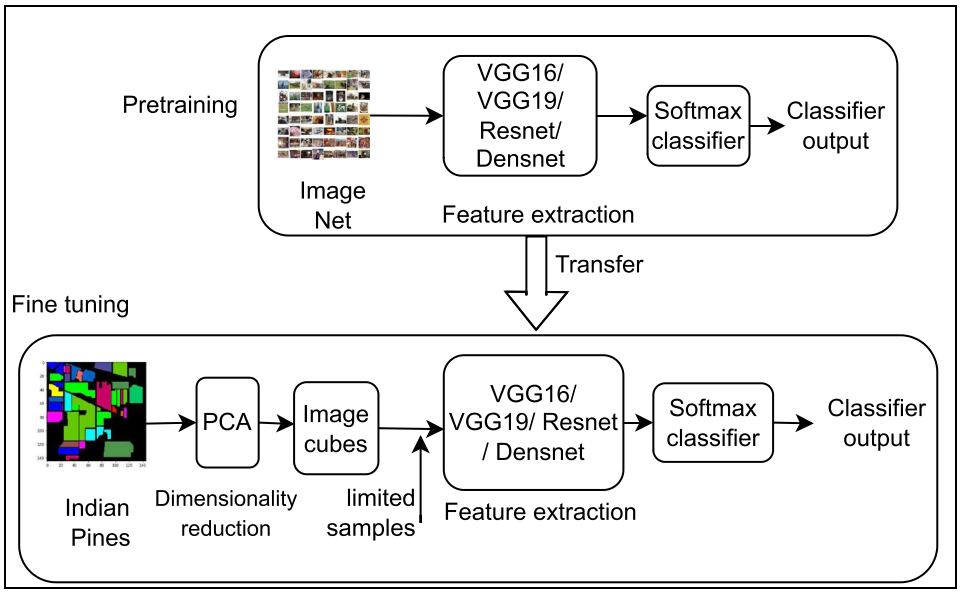

Fig. 1 Flow Diagram for Heterogeneous transfer learning on HSIs datasets using well-defined pre-trained models

Homogeneous transfer learning on HSIs dataset

In Homogeneous transfer learning, the source and target domains have similar features. Figure 2 reflects the pathway for homogeneous transfer learning on the HSIs dataset using customized 2D and 3DCNN models.

The Indian Pines HSIs dataset is used as the source domain, fine-tuned and tested the DL model on the Salinas scene target HSIs dataset. Figure 2 shows the pre-training steps, and Table 2 describes the layer-wise breakdown for CNN models used for crop classification. Batch Normalization, ReLU, and Softmax activation functions are utilized in both models.

Fine tuning algorithm for transfer leaning

Algorithm 1 defined the major steps required to fine-tune the model for the selected target datasets for HSI classification.

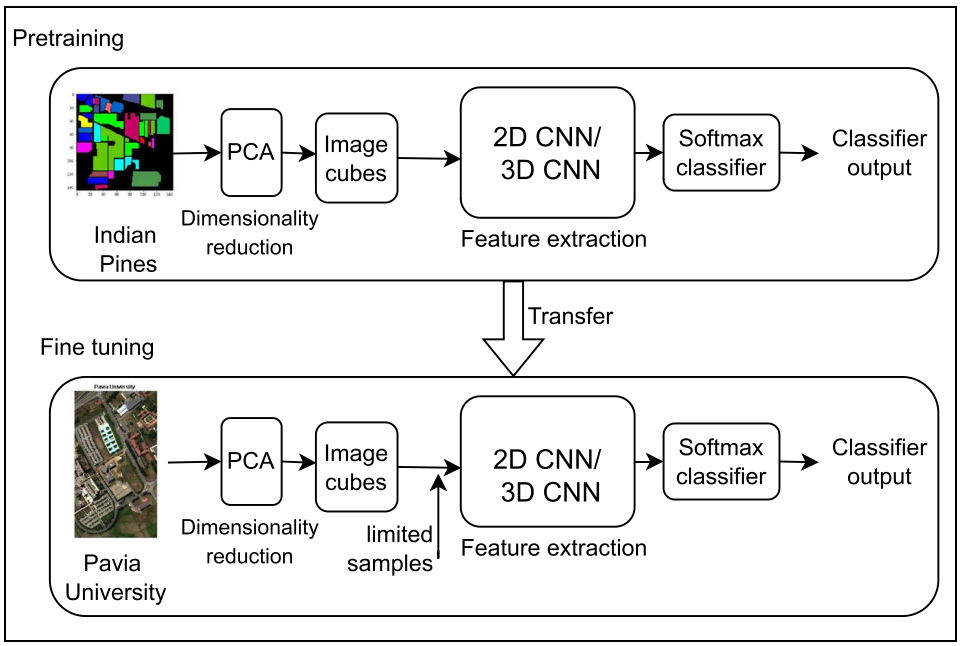

Fig. 2 Flow Diagram for homogeneous transfer learning on HSIs datasets using customized CNN models

Algorithm 1 Algorithm to fine-tune classifier model (M)

Require: Classifier model M(θ) with pre-train weights (w_i)
Dataset $D = \{x_{\mathrm{j}} \in \mathbb{R}^{\mathrm{B}}, j = 1, 2, ...n\}$, where n is number of samples, B is number of spectral band
Apply PCA transformation
$D = \{x_{\mathrm{j}} \in \mathbb{R}^{\mathrm{p}}, j = 1, 2, ...n\}$, Select first p PCA component
Input data cubes ($x_{\mathrm{k\ x\ k\ x\ p}}$), for each $x_{\mathrm{j}} \in D$, k is the cube size
Select t image cubes ($x_{\mathrm{k\ x\ k\ x\ x\ p}}, y$) with the associated label for the retrain model(M),

Suppose model M(θ) is pre-trained with exhaustive source datasets, and the learning weights are w_i. The target dataset D carries the B spectral band reflectance. PCA transformation is applied, and select first p components of dataset D are to reduce redundancy and adjust model M's input dimension. Every pixel is considered along with its neighboring pixels to gain the spatial features from spatial correlation, generating an image cube of dimension k × k. The label of the image cube is the central pixel label. Select such a t image cubes with the corresponding label to fine-tune the model M.

Experiments discussion

The experiment is divided into two parts: (1) The first experiment compares heterogeneous transfer learning performance for the HSIs dataset on different pre-trained DL models trained on the ImageNet dataset. (2) The second experiment compares the performance of homogeneous transfer learning on the HSIs dataset with customized DL models. The information regarding the datasets used for experiments is given as follows:

Datasets used for experiments

For the experiment, the ImageNet dataset and the three most popular hyperspectral image datasets—Indian pines, the salinas scene, and the University of Pavia scene are considered.

Table 2 Layer-wise breakdown of CNN models

Model	Layers
2DCNN	Input layer (19 × 19 × nbands)
	CONV (50 × 5 × 5)
	CONV (100 × 5 × 5)
	FC (1OO)
	FC (nclasses)
3DCNN	Input layer (19 × 19 × nbands)
	CONV (32 × 5 × 5 × 24)
	CONV (65 × 5 × 5 × 16)
	FC (3OO)
	FC (nclasses)

- **Indian Pines (IP)**
  Hyperspectral image dataset Indian pines contain 10249 images with 145 × 145 spatial dimensions and around 224 spectral bands in the wavelength 0.4–2.5 µm. The ground truth for this dataset is divided into sixteen classes representing different crop types.
- **Pavia University (PV)**
  A ROSIS sensor acquired two scenes, Pavia center, and Pavia University, from Pavia, northern Italy. The Pavia university scene is considered for the experiments. Pavia university dataset contains 42,776 images with 610 × 340 spatial dimensions, among which few are noisy bands. After discarding these noisy bands, the dataset contains 103 spectral bands from wavelength 0.43–0.86 µm with nine classes.
- **The Salinas Scene (SL)**
  The Salinas scene is collected by AVRIS sensors over Salinas Valley, California. This dataset contains images with 512 × 217 spatial pixels and 224 spectral bands. Among these 224 spectral bands, twenty bands are water absorption bands that are discarded before experiments. Salinas dataset is divided into 16 classes of crops: bare soils, vegetables, vineyard fields, etc.
- **ImageNet dataset** Russakovsky et al. (2015)
  The ImageNet dataset contains 14,197,122 labeled images of around 1000 classes used for the computer vision task.

Experiments on well-defined DL models pre-trained on ImageNet dataset by heterogeneous transfer learning

The ImageNet dataset is the source dataset for heterogeneous transfer learning and IP, PV, and SL as the target datasets. As the ImageNet dataset has three input channels and the target datasets have hundreds of channels, the dimensionality

reduction technique PCA (principal component analysis) is used to overcome the issue of different dimensions. After reducing the number of components to three, the next task is fine-tuning the transfer learning models on IP, PV, and SL datasets.

The models used for experiments are VGG16, VGG19, ResNet, and DenseNet. These models' last fully connected layers are replaced as a custom fully connected dense layer, which helps to predict the class to which a particular hyperspectral image belongs. This last fully connected layer will have neurons equal to the number of classes of a particular dataset. The model was fine-tuned by using 1%, 5%, 10%, and 15% of the IP, PV, and SL HSIs datasets. The model is tested with the test dataset of IP, PV, and SL.

Models pre-trained on the ImageNet will be used for the HSIs classification task. In that case, the number of channels needs to be reduced to three for all the IP, PV and SL datasets, due to which the data loss will be high, and it won't be possible to take full advantage of hyperspectral images.

Due to this shortcoming, very little literature exists where heterogeneous transfer learning models were used for HSIs classification. Also, the ImageNet dataset is not related to the employed HSIs data, and, hence, it is expected that these models would not be able to exhibit their full potential in HSIs classification.

Experiments on customized DL models pre-trained on Benchmark HSIs datasets by homogeneous transfer learning

With this approach, the goal is to use the transfer learning approach to learn from one dataset, which can be helpful for another related dataset. To achieve this goal, we have used the 100% labeled samples of the IP dataset.

The IP dataset has more than 200 dimensions, and the task of training the model becomes both computationally expensive and time-consuming. The first step before training the model using CNN is to reduce the computation cost, achieved through dimensionality reduction. IP dataset, considered the most diverse HSIs dataset, is used for pre-training customized 2DCNN and 3DCNN models. To achieve transfer learning, the weights of these pre-trained models are slightly adjusted by using 32, 64, 128, and 256 samples of the Salinas scene and Pavia university dataset.

Batch normalization speeds up the training process and avoids over-fitting for the model. ReLU is used as an activation function for convolutional layers and Softmax is used with the last FC layer to provide multi-class classification.

Results discussion

This section discusses the performance accuracy of HSIs classification using well-defined DL models and heterogeneous transfer learning. Also, it discusses the performance accuracy of HSIs classification using customized DL models using homogeneous transfer learning.

Results of well-defined DL models pre-trained on ImageNet dataset by heterogeneous transfer learning

This section discusses the performance accuracy of HSIs classification using DL models like VGG-16, VGG-19, ResNet, and DenseNet with the IP, PV, and SL HSIs datasets using heterogeneous transfer learning.

Table 3 Overall accuracies of pre-trained VGG16, VGG19, ResNet, and DenseNet models on IP, SL, and PV datasets

(a)				(b)			
VGG16	IP	PV	SL	VGG19	IP	PV	SL
1%	46.73	93.15	95.31	1%	44.41	89.23	91.57
5%	86.13	96.27	97.38	5%	78.17	97.81	99.08
10%	95.78	96.13	98.65	10%	91.38	99.17	99.62
15%	96.16	98.43	99.41	15%	95.32	99.26	99.88
(c)				(d)			
ResNet	IP	PV	SL	DenseNet	IP	PV	SL
1%	46.73	93.15	95.31	1%	67.94	98.81	94.06
5%	86.13	96.27	97.38	5%	95.62	98.27	99.6
10%	95.78	96.13	98.65	10%	90.81	99.03	99.73
15%	96.16	98.43	99.41	15%	99.13	99.58	99.9

VGG-16 model

Table 3a shows the results obtained when the pre-trained VGG16 model with the ImageNet dataset was fine-tuned with limited labeled data of IP, PV, and SL datasets. One can observe from Table 3a that even when only 1% of the dataset samples are taken for training, the model gives overall accuracy of 93.15% and 95.31% for PV and SL datasets respectively. However, VGG16 does not perform well on the IP dataset for the same amount of samples because it has fewer labeled samples that can be used for training.

When 5% samples of the dataset are taken for training, the model gives overall accuracy of 86.13%, 97.38%, and 96.27% for IP, SL, and PV datasets respectively. Here, it can be observed that the performance of VGG16 on the IP dataset has improved exponentially. When 10% samples of the dataset are taken for training, the model gives overall accuracy of 95.78%, 96.13%, and 98.65% for IP, SL, and PV datasets respectively. With 15% samples of the dataset taken for training, the model gives overall accuracy of 96.16%, 98.43%, and 99.41% for IP, PV and SL datasets respectively. Figure 3a shows the comparative accuracy achieved by the VGG16 model when heterogeneous transfer learning is applied using the ImageNet dataset as source and IP, PV and SL as target datasets. It can be observed from these results that by using only 15% samples of the dataset, overall accuracy greater than 99% is achieved for both SL and PV datasets, and also 96.16% accuracy for IP dataset.

VGG19 model

Table 3b shows the results obtained when the pre-trained VGG19 was fine-tuned with the IP, PV, and SL datasets. The comparative performance of the VGG19 model is shown in Fig. 3b. It can be observed from the graph that even when only 1% of the dataset samples are taken for training, the model gives overall accuracy of 89.23% and 91.57% for PV and SL datasets respectively. However, as VGG16 for the 1% of samples, VGG19 does not perform very well on the IP dataset giving overall accuracy of only 44.41%. VGG16 and VGG19 suffer from the problem of vanishing gradient, another reason for such comparative low performance. By increasing the labeled samples to 15% of the dataset, we can achieve overall accuracy greater than 99% for both SL and PV datasets and also 95.32% accuracy for the IP dataset. It can be observed that VGG19 performs even worse than the VGG16 model on the IP dataset. The result shows that increasing the number of layers in the model does not continuously improve performance.

ResNet model

Table 3c shows the results obtained with a pre-trained ResNet model with IP, PV, and SL datasets. Figure 3c shows the comparative accuracy achieved by the ResNet model, which reflects that the model ResNet performs comparatively better on the SL dataset. SL dataset gives around 95.31% accuracy just using 1% dataset as training dataset,

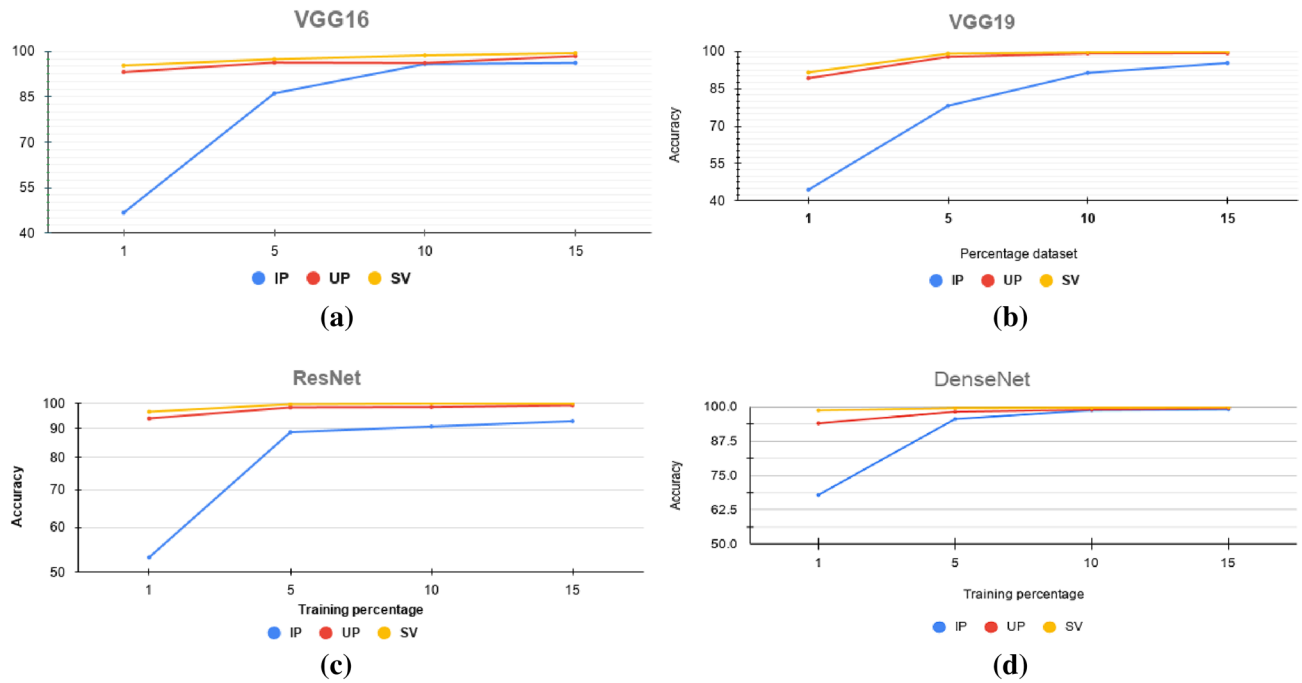

Fig. 3 Performance models pre-trained on ImageNet dataset after heterogeneous transfer learning

which is better than IP and PV datasets. By using 15% of labeled samples of the dataset, the achieved overall accuracy is greater than 98% for both SL and PV datasets and also 96.16% accuracy for the IP dataset. ResNet gives comparatively less accuracy for the IP datasets, which is not very bad considering the number of training samples used.

DenseNet model

Table 3d shows the results obtained when DenseNet was fine-tuning with IP, PV, and SL datasets. The comparative results in Fig. 3(d) show that with only 1% of the dataset samples taken for training, the model gives overall accuracy of 94.06% and 98.81% for SL and PV datasets, respectively. However, like previous models, for 1% of samples, DenseNet does not perform very well on the IP dataset. The pre-trained DenseNet performs comparatively better on the SL dataset and gives around 99% accuracy just using 5% dataset as a training dataset than IP and PV datasets. DenseNet performs better for the IP dataset than all other models. For around 15% training dataset, DenseNet provides accuracy greater than 99% for all three datasets.

Results of customized DL models pre-trained on Benchmark HSIs datasets by homogeneous transfer learning

This section discussed the accuracy of HSIs classification using customized 2DCNN and 3DCNN models trained on the IP dataset and tested on SL and PV dataset samples using homogeneous transfer learning.

2DCNN model

Table 4 displays the overall accuracy (OA), Average accuracy (AA), and Kappa in % for SL and PV datasets when training samples are used as 32, 64, 128, and 256 using the 2DCNN model. Figure 4 displays the comparative performance of PV and SL datasets on the mentioned evaluation measures. As can observe from Table 4, 32 labeled samples of the dataset were used to adjust the weights. The 2DCNN model gives overall accuracy of 83.77%, Average accuracy of 82.63%, a Kappa score of 82.15% for the SL dataset, and overall accuracy of 74.39%, Average accuracy of 73.26%, and a Kappa score of 73.85% for PV dataset respectively. This result is close to the overall accuracy of the best performing model, DenseNet using 1% samples from the first experiment. Due to the lack of similarities between the spectral features, the model does not perform as significantly on PV as on SL. When only 64 samples adjust the weights, the 2DCNN model gives overall accuracy of 88.13% for the SL

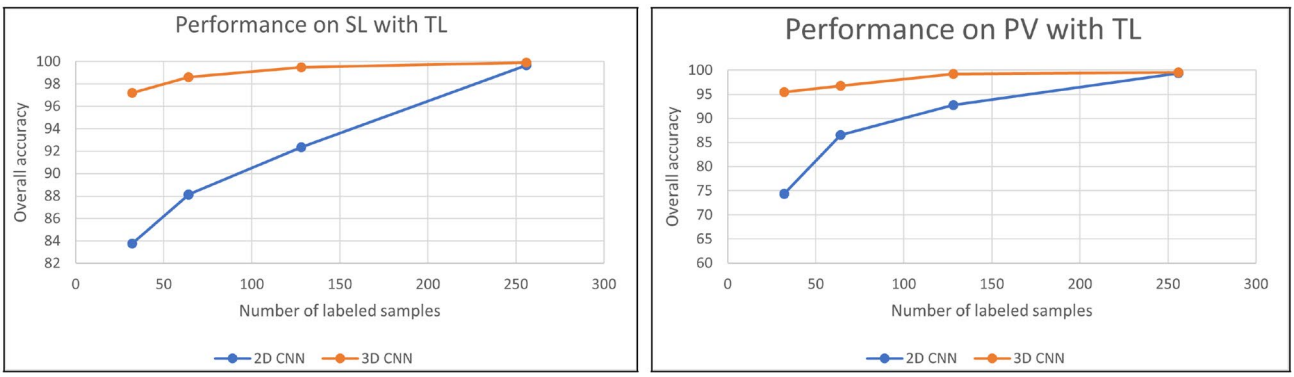

Fig. 4 Performance of SL and PV datasets using 2DCNN and 3DCNN after homogeneous Transfer learning

Table 4 Results for experiments performed on 2DCNN model pre-trained on IP dataset

Training samples	Salinas scene			University of Pavia		
	OA (%)	AA (%)	Kappa (%)	OA (%)	AA (%)	Kappa (%)
32	88.77	82.63	82.15	74.39	73.26	73.85
64	88.13	87.38	86.97	86.53	85.41	85.22
128	92.35	90.54	91.65	92.78	91.59	91.84
256	99.68	97.28	98.21	99.35	98.14	98.79

Table 5 Results for experiments performed on 3DCNN model pre-trained on IP dataset

Training samples	Salinas scene			University of Pavia		
	OA (%)	AA (%)	Kappa (%)	OA (%)	AA (%)	Kappa (%)
32	97.21	96.83	96.52	95.46	94.53	94.17
64	98.61	97.42	96.95	96.76	95.62	95.18
128	99.49	98.19	97.76	99.21	97.74	97.27
256	99.92	99.18	98.71	99.56	99.15	98.51

dataset and 86.53% for the PV dataset. This result is better than the best-performing model ResNet in heterogeneous transfer learning.

2DCNN model gives overall accuracy of 92.35% for the SL dataset and 92.78% for the PV dataset with 128 labeled samples to adjust weights, which is better than the ResNet as received in the first experiment. With 256 labeled samples 2DCNN model was given 99.68% for the SL dataset and 99.35% for the PV dataset accuracy.

3DCNN model

Table 5 displays the overall accuracy(OA), Average accuracy(AA), and Kappa in % for SL and PV datasets when training samples are used as 32, 64, 128, and 256 using the 3DCNN model, and Fig. 4 displays the comparative performance of the 2DCNN and 3DCNN models achieved on the datasets SL and PV. As per Fig. 4, the CNN 3D model gives OA 97.21% for the SL dataset and 95.46% for the PV dataset, with 32 labeled samples used for fine-tuning. This performance is better than the CNN 2D model. It also outperforms the pre-trained DenseNet model. With 64 labeled samples, the CNN 3D model gives overall accuracy of 98.61% and 96.76% for the SL and PV datasets, respectively. This performance is better compared with the OA of the best performing model ResNet using 2% samples from the first experiment and outranks 2DCNN, which for 64 samples gave overall accuracy of 88.13% for the SL and 86.53% for the PV dataset.

With 128 labeled samples, 3DCNN gives OA 99.49% for the SL dataset and 99.21% for the PV dataset, which is better when compared with the accuracy of the best performing model ResNet, which uses 2% samples for such performance, and outranks 2DCNN, which for 128 samples gave overall accuracy of 92.35%, 92.78% respectively for the SL and PV datasets. The 3DCNN model gives an OA of 99.92% for the SL datasets and 99.56% for the PV dataset with 256 labeled samples. This is far better when compared with the accuracy of the best performing model ResNet, which uses 2% samples for such performance.

Conclusion

The temporal variation in the crops and its spatial distribution information plays an important role in better agriculture development and decision-making for state/central government. Hyperspectral data with various benchmark datasets are used along with deep learning models for crop classification. With the limited samples of the Hyperspectral dataset, the Deep learning approach can not perform well. To overcome this limitation, the Tranferlearning approach has opted. This paper presented the performance of homogeneous and heterogeneous transfer learning with the HSIs dataset using well-defined DL models and customized DL models. In heterogeneous transfer learning, the pre-train DenseNet model achieved 99% of accuracy for the IP and PV dataset with 15% of labeled training samples. For the SL dataset, VGG19, ResNet, and DenseNet models perform significantly better all three models give 99.9% accuracy. As shown in the results, the homogeneous transfer learning with 2DCNN and 3DCNN models pre-trained on the IP dataset and used again adjusted on the SL and PV dataset provided overall accuracy far better than the models pre-trained on the ImageNet dataset. conclusively, the result of an experiment shows that transfer learning is one of the solutions for limited labeled samples in the HSIs domain.

References

Adam E, Mutanga O, Rugege D (2010) Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. Wetl Ecol Manag 18(3):281–296

Ajadi OA, Barr J, Liang SZ et al (2021) Large-scale crop type and crop area mapping across brazil using synthetic aperture radar and optical imagery. Int J Appl Earth Obs Geoinf 97(102):294

Banerjee R, Kathiria P, Shukla D (2020) Recommendation systems based on collaborative filtering using autoencoders: issues and opportunities. In: The international conference on recent innovations in computing, vol 701. Springer, pp 391–405

Breiman L (2001) Random forests. Mach Learn 45(1):5–32

Cai R, Shang G (2021) Flexible 3-d gabor features fusion for hyperspectral imagery classification. J Appl Remote Sens 15(3):036,508

Carrasco O, Gomez RB, Chainani A, et al (2003) Hyperspectral imaging applied to medical diagnoses and food safety. In: Geo-Spatial and temporal image and data exploitation III, vol 5097. SPIE, pp 215–221

He X, Chen Y, Ghamisi P (2019) Heterogeneous transfer learning for hyperspectral image classification based on convolutional neural network. IEEE Trans Geosci Remote Sens 58(5):3246–3263
Ismail R, Mutanga O, Peerbhay K (2016) The identification and remote detection of alien invasive plants in commercial forests: an overview. S Afr J Geomat 5(1):49–67
Jiang Y, Li Y, Zhang H (2019) Hyperspectral image classification based on 3-d separable resnet and transfer learning. IEEE Geosci Remote Sens Lett 16(12):1949–1953
Kavitha MK, Arivazhagan S (2010) Combined feature based hyperspectral image classification technique using support vector machines. Int J Electron Commun Eng 4(10):1574–1581
Khan MJ, Khan HS, Yousaf A et al (2018) Modern trends in hyperspectral image analysis: a review. IEEE Access 6:14,118–14,129
Krizhevsky A, Sutskever I, Hinton GE (2017) ImageNet classification with deep convolutional neural networks. Commun ACM 60(6):84–90
Kureel N, Sarup J, Matin S et al (2022) Modelling vegetation health and stress using hypersepctral remote sensing data. Model Earth Syst Environ 8(1):733–748
Liang M, Jiao L, Yang S et al (2018) Deep multiscale spectral-spatial feature fusion for hyperspectral images classification. IEEE J Sel Top Appl Earth Obes Remote Sens 11(8):2911–2924
Lin J, Ward R, Wang ZJ (2018) Deep transfer learning for hyperspectral image classification. In: 2018 IEEE 20th International Workshop on Multimedia Signal Processing (MMSP), IEEE, pp 1–5
Liu Y, Gao L, Xiao C et al (2020) Hyperspectral image classification based on a shuffled group convolutional neural network with transfer learning. Remote Sens 12(11):1780
Lopez J, Hinojosa CA, Arguello H (2021) Efficient subspace clustering of hyperspectral images using similarity-constrained sampling. J Appl Remote Sens 15(3):036,507
Mou L, Ghamisi P, Zhu XX (2017) Deep recurrent neural networks for hyperspectral image classification. IEEE Trans Geosci Remote Sens 55(7):3639–3655
Mughees A, Tao L (2016) Efficient deep auto-encoder learning for the classification of hyperspectral images. In: 2016 International Conference on virtual reality and visualization (ICVRV), IEEE, pp 44–51
Munishamaiaha K, Rajagopal G, Venkatesan DK et al (2022) Robust spatial-spectral squeeze-excitation adabound dense network (se-ab-densenet) for hyperspectral image classification. Sensors 22(9):3229
Paoletti M, Haut J, Plaza J et al (2019) Deep learning classifiers for hyperspectral imaging: A review. ISPRS J Photogramm Remote Sens 158:279–317
Patel P, Patel Y, Patel U, et al (2022a) Towards automating irrigation: a fuzzy logic-based water irrigation system using iot and deep learning. Model Earth Syst Environ 8:1–16
Patel U, Patel S, Kathiria P (2022b) Hyperspectral image classification using transfer learning. In: Communication and intelligent systems. Springer, pp 545–556
Roy SK, Manna S, Song T et al (2020) Attention-based adaptive spectral-spatial kernel resnet for hyperspectral image classification. IEEE Trans Geosci Remote Sens 59(9):7831–7843
Russakovsky O, Deng J, Su H et al (2015) ImageNet large scale visual recognition challenge. Int J Comput Vis (IJCV) 115(3):211–252. https://doi.org/10.1007/s11263-015-0816-y
Stergioulas A, Dimitropoulos K, Grammalidis N (2022) Crop classification from satellite image sequences using a two-stream network with temporal self-attention. In: 2022 IEEE International Conference on imaging systems and techniques (IST), IEEE, pp 1–6
Transon J, d'Andrimont R, Maugnard A et al (2018) Survey of hyperspectral earth observation applications from space in the sentinel-2 context. Remote Sens 10(2):157
Vaddi R, Manoharan P (2020) Hyperspectral image classification using cnn with spectral and spatial features integration. Infrared Phys Technol 107(103):296
Woldemariam GW, Tibebe D, Mengesha TE et al (2022) Machine-learning algorithms for land use dynamics in lake haramaya watershed, ethiopia. Model Earth Syst Environ 8(3):3719–3736
Xie Q, Lai K, Wang J et al (2021) Crop monitoring and classification using polarimetric radarsat-2 time-series data across growing season: a case study in southwestern ontario, canada. Remote Sens 13(7):1394
Xu Y, Du J, Dai LR et al (2014) A regression approach to speech enhancement based on deep neural networks. IEEE/ACM Trans Audio Speech Lang Process 23(1):7–19
Yang G, Gewali UB, Ientilucci E, et al (2018) Dual-channel densenet for hyperspectral image classification. In: IGARSS 2018-2018 IEEE international geoscience and remote sensing symposium, vol 8. IEEE, pp 2595–2598
Ye M, Ruiwen N, Chang Z et al (2021) A lightweight model of vgg-16 for remote sensing image classification. IEEE J Sel Top Appl Earth Observ Remote Sens 14:6916–6922
Zhang H, Li Y, Jiang Y et al (2019) Hyperspectral classification based on lightweight 3-d-cnn with transfer learning. IEEE Trans Geosci Remote Sens 57(8):5813–5828
Zimmer SN, Reeves MC, St Peter JR, et al (2022) Earlier green-up and senescence of temperate united states rangelands under future climate. Model Earth Syst Environ 60:1–17