ISSN 2229-5518

Classification of Remote Sensing Image using Deep learning

Muhammad Sanusi, Odulu Lydia, Amina Ajibola, Efron Gajere, Mustapha Lawal

Abstract— Machine learning has witnessed a lot of research, in the area of image classification. Different algorithms have been proposed in the area of natural image classification, the common practice is to pertain a deep learning (DL) model using a data set with a large number of labeled samples, such as ImageNet, and then to fine-tune the model using a data set which contains limited training samples. However, Remote Sensing (RS) data are more complex than natural image, parts of them are typically even acquired by the use of different remote sensors with different design characteristics. The technique that introduce transfer learning to RS image classification presents a major challenge and requires significant research work. This work uses CNN architectures to classify remotely sense image using transfer learning pre-trained work in Alexnet dataset and then fine-tuned them on the RS images. Based on the research conducted we can conclude that the proposed model demonstrates high level of accuracy when compare to the state-of-the-art models. By introducing transfer learning and exploring a different CNN architecture which was a major weakness of the existing system, the proposed CNN was able to improve the classification accuracy achieving 99.99% as against the existing conventional CNN with accuracy of 99.91%, stacked auto-encoders SAE with accuracy of 93.98% and deep belief network DBN with accuracy of 95.91%.

Keywords- Artificial Neural Network (ANN), Convolution Neural Network (CNN), Deep Learning, Machine Learning.

1 INTRODUCTION

Classification of remote sensing images plays an important role in many fields, such as land management, urban planning, environmental exploration and monitoring, and natural disaster detection. Over the past decades, researchers have done a great deal of experiments in the scene classification for satellites and aerial photographs, and have developed many taxonomies [1, 2].

In order to make informed decisions concerning the environment, there are government ministries and agencies that are equipped to observe it. This enables us to make effective changes around us as appropriate or desirable. This process is referred to as earth observation, and has applications in disaster response, resource management and precision farming among others. Earth observation data is gathered by a range of techniques, and can be roughly categorized as remote sensing. It is where "the distance between the object and the sensor far surpasses the linear dimensions of the sensor". In the last decade, manual analyses of satellite imagery were feasible primarily because the volume of images available was quite low - but that is not the case now. Relevant information extraction from images thus becomes a problem with the high volume of data we deal with today. A major component of these problems is annotation (or labeling), wherein one identifies the structures and patterns visible in a satellite image. Over the years, research in the computer vision community has addressed this problem of automating the analysis of large-scale data in different ways. Machine learning techniques have proven to be strong candidates here, especially in the last few years. At the time of writing, the state of the art in the automation of visual labeling tasks is seen in the deep learning research community, and that is where this research picks up at.

Remote sensing classification technology, along its improvement has been unavoidably utilized to earth observation and land-utilize survey to investigate natural planet resources quantity and distribution. It is an intense tool to utilize either aerial or satellite images for selection, investigation, understanding, mapping, evaluation, error adjustment, and accuracy calculation of scene. classification of remote sensing images plays a very significant role in many fields, such as land management, urban planning, environmental exploration and monitoring, and natural disaster detection.

However, with the continuous improvement of science and technology, the spatial resolution of remote sensing images is getting higher and higher, and their spatial and structural patterns are becoming more and more abundant. The phenomenon of the same objects with different spectrum and the foreign ones with common spectrum are more widespread. However, the most of the classical methods are based on artificial or shallow learning algorithms, and the low middle-level semantic features extracted are limited in the description ability, which makes it difficult to improve the classification accuracy further [3].

There is a wide range of classification methodologies. There are two principle types: unsupervised and supervised classification. Supervised or unsupervised classification techniques used to analyze a reliable land-cover map of the geographical area captured in the image [4]. Supervised classification will be done only with valid ground truth points. It is not that much of easy to take those points. Proper planning should be there to go for this classification. Generally, two ways are followed in supervised classification of remote sensing data. First one is ordering the new data according to our requirement and getting the ground truth data next. Second one is selecting the data from database if is available the ground data with us. In the former case of classification, clustering is inherently a tough problem. Given a set of data samples, every clustering result is similarly conceivable with no earlier learning about the hidden likelihood dispersion of the information. The clustering method expect some model to model the data, which, basically, is reflected in the relating cluster results. In unsupervised classification, a algorithm, for example, K-means or Iso-Data takes a remotely sensed data set and examines a pre-indicated quantity of statistical in multi spectral space. In spite of the fact that these kinds of clusters generally not identical to genuine land cover classes, this kind of approach does not require earlier learning of the Land-covering the review site. Prior knowledge is required in case of supervised classification, thus the quantity of the desired classes. With this type of approach like maximum likelihood (MLLH) classifier or support vector machines (SVM) classifier, samples of each thematic class are chosen to train statistically the classifier. Once trained, the algorithm is applicable for overall image to create a final classified image. If ground truth data is available supervised method is better. In both cases (unsupervised and supervised), the acquired classification results should be statistically surveyed relatively to a ground truth (reference date) or by utilizing a numerical measure [5].

Recently, deep learning (DL) has become the fastest-growing trend in big data analysis and has been widely and successfully applied to various fields of computer application successfully including sequential data, processing of natural language, speech recognition and image classification, because of its outstanding performance compared with that of traditional learning algorithm. Standing at the paradigm shift towards data-intensive science, machine learning techniques are becoming increasingly important. In particular, as a major breakthrough in the field, deep learning has proven as an extremely powerful tool in many fields. Shall we embrace deep learning as the key to all? Or, should we resist a "black-box" solution? There are controversial opinions in the remote sensing community [6]. Deep learning is the fastest-growing trend in big data analysis and has been deemed one of the 10 breakthrough technologies of 2013. Such work is inspired by biology stating that for primate visual systems, the brain is organized in deep architecture and the perception is also represented at multiple levels of abstraction. DL architectures are characterized as artificial neural networks, involving usually more than two layers. It is characterized by neural networks (NNs) involving usually more than two layers (for this reason, they are called deep). As their shallow counterpart, deep neural networks exploit feature representations learned exclusively from data, instead of hand-crafting features that are mostly designed based on domain-specific knowledge [6]. As with their shallow counterpart, deep neural networks exploit feature representations learned exclusively from data [7].

In the last decade, manual analyses of satellite imagery were feasible primarily because the volume of images available was quite low - but that is not the case now. Relevant information extraction from images thus becomes a problem with the high volume of data we deal with today. A major component of these problems is annotation (or labeling), wherein one identifies the structures and patterns visible in a satellite image. Machine learning techniques have proven to be strong candidates here, especially in the last few years.

Despite its great potential, in general, the use of DL in RS image classification brings forward significant new challenges. There are several reasons for this, many RS data, especially hyperspectral images (HSIs), contain hundreds of bands that can cause a small patch to involve a really large amount of data, which would demand a large number of neurons in a DL network [7]. The performance of DL-based RS classification techniques has shown their effectiveness in solving real-world problems, although such performance does not reflect the full potential of DL yet. A number of challenges still affect the classification accuracy of the existing models including lack of unified feature representation for different source images [8], How to introduce transfer learning to RS image classification [9] and relatively higher time consumption [10].

Recently, an increasing number of novel deep networks have been proposed. These networks can often achieve excellent performance in performing their dedicated task. For instance, U-net [11] can obtain an impressive performance in segmentation, ResNet [12] can have an outstanding accuracy in applicable image classification and object detection. However, almost of such networks are aimed at coping with natural image processing. As we mentioned previously, RS images are generally different from natural images. Exploring appropriate network structures for a given RS image classification problem is still an open topic [7]. Moreover, for natural image classification, the common practice is to pretrain a DL model using a data set with a large number of labeled samples, such as ImageNet, and then to fine-tune the model using a data set which contains limited training samples. However, RS data are more complex than natural image, parts of them are typically even acquired by the use of different remote sensors. How to introduce transfer learning to RS image classification therefore, presents a major challenge, which needs significant further research [7], thus this paper is aim at improving the classification accuracy of the RS image in this research work using a deep learning convolutional neural network (CNN) to classify remote sensing image by introducing transfer learning to the RS image.

The remainder of this paper is organized as follows: Section one contains background of the research. Section two presents concept review and related review. Section Three present methodology including data collection and preprocessing, proposed algorithm, and data analysis. Section four contain the results and discussion, while Section five present summary, conclusion and recommendations.

2 LITERSTURE REVIEW

2.1 Review oF Related Work

Detailed This section gives the review of literatures that utilized machine learning techniques for the classification of remote sensing imagery. Machine learning research stems from the idea that a computer can be given the ability to learn, as a human would do, without being explicitly programmed. Deep learning is a subset of machine learning, and refers to the application of a set of algorithms called neural networks, and their variants. In such methods, one provides the network (or model) with a set of labeled examples which it learns, or trains on and labeling these examples is done in many ways. Application of machine learning methods to large databases is called data mining. The analogy is that a large volume of earth and raw material is extracted from a mine, which when processed leads to a small amount of very precious material; similarly, in data mining, a large volume of data is processed to construct a simple model with valuable use, its application areas are abundant. But machine learning is not just a database problem; it is also a part of artificial intelligence. To be intelligent, a system that is in a changing environment should have the ability to learn. If the system can learn and adapt to such changes, the system designer need not foresee and provide solutions for all possible situations. Collaborative platforms such as Open Street Map and crowd sourcing market places are ideal for the annotation of images, and this existing volume can already be leveraged.

Recently, deep learning (DL) has become the fastest-growing trend in big data analysis and has been widely and successfully applied to various fields of computer application successfully including sequential data, processing of natural language, speech recognition and image classification (Abdel-Hamid, Mohamed, Jiang & Penn, 2012, March), because of its outstanding performance compared with that of traditional learning algorithms. Standing at the paradigm shift towards data-intensive science, machine learning techniques are becoming increasingly important. In particular, as a major breakthrough in the field, deep learning has proven as an extremely powerful tool in many fields.

Many approaches have been proposed aiming at accurate and robust classification models for example, Moorthi et al., 2011 [11] proposed a Support Vector Machine (SVM) and evaluate it against Maximum likelihood estimation method, The result clearly shows that SVM algorithm have better performance in satellite image classification although The accuracy in the SVM model can be further improved by using other kernel types such as radial basis function and sigmoid function, by fine tuning the parameters such as degree of kernel polynomial, bias in kernel function and gamma in kernel function. similarly, Praveena & Singh [12] proposed a hybrid clustering algorithm and feed-forward neural network classifier for land-cover mapping of trees, shade, building and road. The proposed technique performed better than all the existing algorithms taken for comparison. an effective deep neural network was also proposed by Wang et al., 2018[8] and compare the performance with SIFT, SURF, SAR-SIFT, PSO-SIFT, the experimental results shows that the applied transfer learning further improves the accuracy and reduces the training cost. Jony et al., 2018[9] employs an ensemble classifier to detect water in satellite images for flood assessment and evaluate it against MediaEval 2017, it was found that this approach is capable of producing good classification accuracy for a seen location when bands are used and an unseen location when NDWI is used. Y. Li et al, 2018[7] Present typical DL models that may be used to perform RS image classification and compare the performance against CNNs, SAEs and DBNs, the performance of DL-based RS classification techniques has shown their effectiveness in solving real-world problems but the major challenge with the study was how to introduce transfer learning to RS image classification therefore, presents a major challenge. thus, we want to improve the classification accuracy of the RS image in this research work by proposing a deep learning using convolutional neural network (CNN) architecture to classify remote sensing image by introducing transfer learning to the RS image.

3 METHODOLOGY

This work uses different CNN architectures to classify remotely sense image using transfer learning. we used CNNs pre-trained on Alexnet dataset and then fine-tuned them on the RS images. Some of these pre-trained models are available as part of Matlab exchange files, while others are provided by the deep learning community. The RS Image dataset consists of satellite image and not photographs, yet CNNs pre-trained on Alexnet dataset have shown the ability to transfer to other image domains.

3.1 Convolutional Neural Network

CNNs where widely applied in various computer application successfully including sequential data. Convolutional Neural Networks mainly focus in learning features that are abstract; this is achieved by stacking and alternating pooling layers and convolution layers respectively. These convolutional kernels which are the convolution layers in CNN convolve raw input data with multiple local filters thus producing translation invariant local features together with the subsequent pooling layers' extract features and a fixed-length over sliding windows of the raw input data in steps of several rules including average, max and other parameters accordingly.

The CNN is composed of a series of layers, where each layer defines a specific computation as shown in Figure 1. These parts are: convolution layers, pooling layers, and fully connected layers. From Figure 1 below the convolution layers is the foremost layer in the CNN network. The input image maps are convolved with learnable kernels and are subsequently put through the activation function to form the output feature maps. The learning and working process of CNN can be summarized into two stages: (a) networking training and (b) feature extraction and classification. There are two parts for the first stage: a forward part and a backward part. In the forward part, the input images are fed through the network to obtain an abstract representation, which are used to compute the loss cost with regard to the given ground truth labels. Based on the loss cost, the backward part computes the gradients of each parameter of the network. Then all the parameters are updated in response to the gradients in preparation for the next forward computation cycle. After sufficient iterations of training, in the second stage, the trained network can be used to extract deep features and classify unknown images. Figure 1 shows the Convolutional Neural Network framework used for Remote Sensing classification.

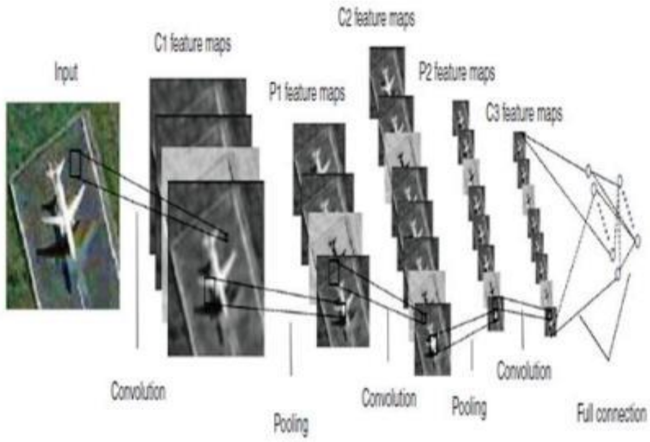

Figure 1: Architecture of The Convolutional Neural Networks Framework used for RS Image Classification

3.2 Transfer Learning

Transfer learning is a deep learning application whereby a fully developed network is pre-trained with a pre- trained network which serves as a starting point to learn a new task for enhanced convergence and performance. Fine-tuning a network with transfer learning is much faster and easier than constructing and training a new network. The features of the pre-trained network are transferred via transfer learning to a new task using a smaller number of training images, features and parameters. The advantage of transfer learning is that the pre-trained network has already learned a rich set of features. These features are applied to a wide range of other similar tasks to reduce the training time with improved accuracy. In this research work, we introduce transfer learning using alexnet to pretrain the network.

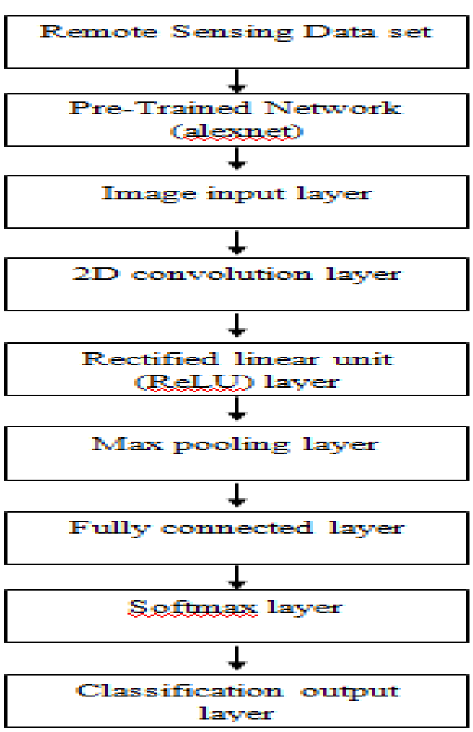

Figure 2: Framework for the proposed Convolutional Neural Network Architecture

This research enhances the accuracy of Remote Sensing Image classification using pre-trained alexnet and CNN algorithms. The pre-trained network with additional modification of the traditional CNN architecture as discussed in the work of [7] were adopted to improve the performance of the system. The working of the proposed system is described below:

The Remote Sensing Data set is resized and preprocessed. They dataset is divided into three (3) main groups which is the train, test and validate set. The train set was given 70% of the dataset while the validate and test set was given 15% each. The train set pass through different stages before a model is built, after the model is built which is tested on the test set and finally the developed model is validated. The training data set is fed to the alexnet pre trained unit for transfer learning. The network is pre-trained via transfer learning whereby the features of the data set pattern are learnt. The output of the unit is fed as input to the network which has multiple convolution layers, Convolution layers, Rectified Linear Unit (ReLU) layer, Max Pool Layers, Fully Connected layer, Softmax layer and Classification layers. The proposed CNN network is a variant of Deep learning network that can learn dependencies between the data and extract patterns for pattern classification applications. The modification was adopted to enhance the classification accuracy in Remote sensing images classification.

3.3 Data Collection

The study makes use of WHU-RS19 Remote Sensing Database (Sheng, Yang, Xu, & Sun, 2012) contains a set of Remote sensing images provided by Google Earth for research purpose. The database was used in the context of aerial scene classification by Google. The Dataset contains 50 images per class, 19 Scene classes, 1,005 images in total. It has about 0.5 Spatial resolution (m) and the image size is 600 × 600.Figure 3.4, 3.5 and 3.6 shows some sample of an airport, commercial area and a viaduct Remote Sensing images used in this work

.

ISSN 2229-5518

Figure 3: Sample Remote Sensing image of an airport.

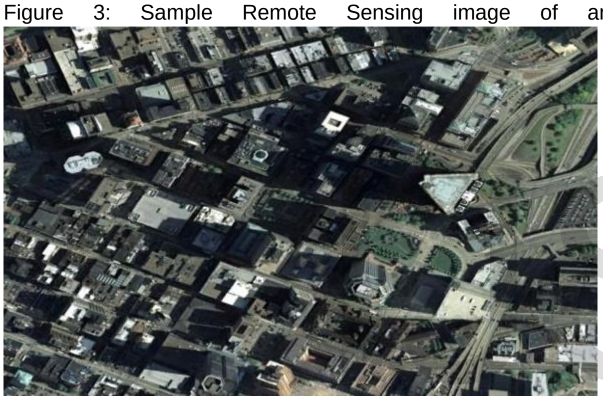

Figure 4: Sample Remote Sensing image of a commercial area.

Figure 5: Sample Remote Sensing image of a viaduct.

3.4 Implementation and Evaluation Setup

The developed CNN classifier model was implemented using MATLAB R2018a. The computer used is a HP laptop running on Windows 10 Operating System with 8GB RAM and Pentium ® Core i7 processor., The ADAM training algorithm is adapted in this work. ADAM is an adaptive learning rate optimization algorithm that's been designed specifically for training deep neural networks. The algorithms leverage the power of adaptive learning rates methods to find individual learning rates for each parameter. It also has advantages of Adagrad training algorithm which works really well in settings with sparse gradients, but struggles in non-convex optimization of neural networks, and RMSprop which tackles to resolve some of the problems of Adagrad and works really well in on-line settings (Duchi, Hazan and Singer, 2011).

the model was evaluated based on the performance obtained using accuracy as the performance metrics which deals with the correct prediction made by the model and this metric can be expressed as:

$$\text{Accuracy} = (TP+TN)/(TP+FP+FN+TN) \qquad (1)$$

4. Results and Discussion

The developed RS image classifier uses different CNN architectures to classify remotely sense image using transfer learning pre-trained on Alexnet dataset and then fine-tuned them on the RS images.

the performance of the model and dataset during training was shown, and the training accuracy was evaluated after 100 iterations. The classification accuracy obtained after training was 99.99% which is a very efficient value that can be obtained in a classification model. The higher the classification accuracy the better the model been built.

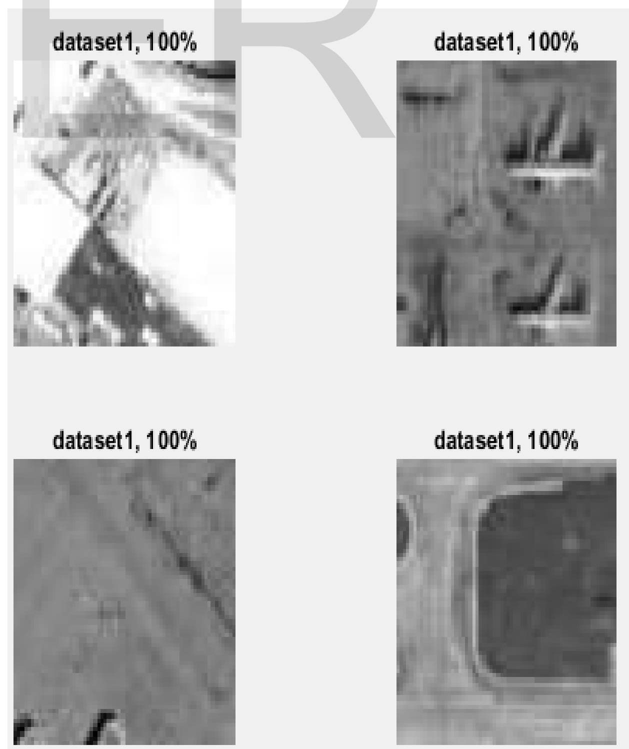

Figure 6: performance of the classifier during training

Figure 6 shows four RS images in grayscale with 99.99% approximated to 100% classification accuracy. They result gotten was analyzed against the existing system Classification results which includes DBN, SAE and CNN respectively. This result is shown in Table 1

Table 1Evaluation of Overall Spectral Features

Models	Classification accuracy in percentage %
CNN	99.91
SAE	93.98
DBN	95.91
Proposed Model	99.99

Based on table 1 above, the proposed system outperforms the existing systems in terms of classification accuracy. By introducing transfer learning and exploring a different CNN architecture which was a major weakness of the existing system, the proposed CNN was able to improve the classification accuracy achieving 99.99% as against the existing work of Conventional CNN with accuracy of 92.28%, stacked auto-encoders SAE with accuracy of 98.42% and deep belief network DBN with accuracy of 96.42%.Although, among the existing system that was used to evaluate the proposed model, the existing conventional CNN used also obtain classification accuracy of 99.91% which was very closed to the classification accuracy of proposed system. This indicated that conventional CNN has demonstrated superiority in terms classification of satellite image as against the stack auto encoders and deep belief network.

The results obtain above is further represented in Figure 7as a graphical chart for more understanding by readers.

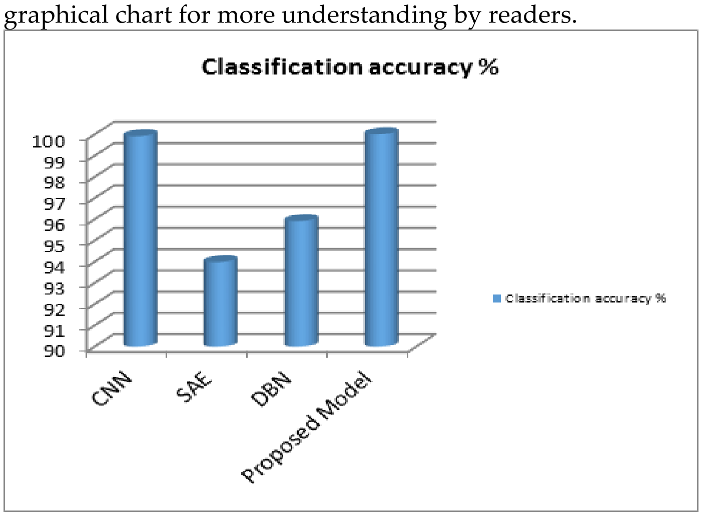

Figure 7: Classification Accuracy of the proposed model Against CNN, SAE and DBN

5. Conclusion

Classification of remote sensing images plays an important role in many fields, such as land management, urban planning, environmental exploration and monitoring, and natural disaster detection. Over the past decades, researchers have done a great deal of experiments in the scene classification for satellites and aerial photographs, and have developed many taxonomies. Machine learning has witnessed a lot of research, in the area of image classification. Different algorithms have been proposed in these areas. This work uses different CNN architectures to classify remotely sense image using transfer learning pre-trained on Alexnet dataset and then fine-tuned them on the RS images. Based on the research conducted we can conclude that the proposed model demonstrates high level of accuracy when compare to the state-of-the-art models, By introducing transfer learning and exploring a different CNN architecture which was a major weakness of the existing system, the proposed CNN was able to improve the classification accuracy achieving 99.99% as against the existing conventional CNN with accuracy of 99.91%, stacked auto-encoders SAE with accuracy of 93.98% and deep belief network DBN with accuracy of 95.91%.

Although, among the existing system that was used to evaluate the proposed model, the existing CNN used in Indian university also obtain classification accuracy of 99.91% which was very closed to the classification accuracy of proposed system algorithms with slides margin of 0.08, thus demonstrate the suitability of application of the proposed model in remotely sense data.

5.1 Recommendation and Future Work

Despite the advancement of DL in remote sensing image classification owing to the increased availability of RS data and computational resources. Still yet, there is still a long way to go in order to realize full potential while coping with many unanswered challenges. Based on our research work, there are several important open issues that needs to be potentially address thus we drive the following recommendation.

i. RS images are generally different from natural images. Exploring appropriate network structures for a given RS image classification problem still need be to explore in the future.

ii. Although DL models can learn high-level abstract features from raw images with excellent performance in dealing with a wide range of problems, we have to pay attention to the observation that such performance heavily relies on large amounts of training samples thus future work should use larger data samples to ensure model reliability.

iii. In RS images, the available labeled samples are rather limited, thereby restricting the DL-based RS images classification approaches to obtain better performance. How to build an efficient network and train it with a small number of training samples is both challenging and interesting. Investigating into novel models that can exploit unlabeled samples is clearly a desirable direction for further work

ACKNOWLEDGEMENT

This study was supported by the Tertiary Education Trust Fund (TETFund) Institutional Based Research (IBR) Fund, through the directorate of Research and Innovation of University of Abuja

References

1. Cheng, G., et al., *Multi-class geospatial object detection and geographic image classification based on collection of part detectors.* ISPRS Journal of Photogrammetry and Remote Sensing, 2014. **98**: p. 119-132.
2. Li, T., J. Zhang, and Y. Zhang. *Classification of hyperspectral image based on deep belief networks*. in *2014 IEEE international conference on image processing (ICIP)*. 2014. IEEE.
3. Yang, Z., X.-d. Mu, and F.-a. Zhao, *Scene classification of remote sensing image based on deep network grading transferring.* Optik, 2018. **168**: p. 127-133.
4. Tyagi, M., et al., *A context-sensitive clustering technique based on graph-cut initialization and expectation-maximization algorithm.* IEEE Geoscience and Remote Sensing Letters, 2008. **5**(1): p. 21-25.
5. Radhika, K. and S. Varadarajan. *Satellite image classification of different resolution images using cluster ensemble techniques*. in *2017 International Conference on Algorithms, Methodology, Models and Applications in Emerging Technologies (ICAMMAET)*. 2017. IEEE.
6. Zhu, X.X., et al., *Deep learning in remote sensing: A comprehensive review and list of resources.* IEEE Geoscience and Remote Sensing Magazine, 2017. **5**(4): p. 8-36.
7. Li, Y., et al., *Deep learning for remote sensing image classification: A survey.* Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2018. **8**(6): p. e1264.
8. Wang, S., et al., *A deep learning framework for remote sensing image registration.* ISPRS Journal of Photogrammetry and Remote Sensing, 2018. **145**: p. 148-164.
9. Jony, R.I., et al. *Ensemble Classification Technique for Water Detection in Satellite Images*. in *2018 Digital Image Computing: Techniques and Applications (DICTA)*. 2018. IEEE.
10. Huang, F., Y. Yu, and T. Feng, *Automatic Building Change Image Quality Assessment in High Resolution Remote Sensing Based on Deep Learning.* Journal of Visual Communication and Image Representation, 2019: p. 102585.
11. Moorthi, S.M., et al. *Kernel based learning approach for satellite image classification using support vector machine*. in *2011 IEEE Recent Advances in Intelligent Computational Systems*. 2011. IEEE.
12. Praveena, S. and S. Singh. *Hybrid clusteing algorithm and Neural Network classifier for satellite image classification*. in *2015 International Conference on Industrial Instrumentation and Control (ICIC)*. 2015. IEEE.