Classification of remote sensing images using CNN

Aswathy K Cherian*, E. Poovammal

Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, India.

*E-mail: aswathyc@srmist.edu.in; poovamme@srmist.edu.in

Abstract. A number of factors can influence the classification of images. The effective outcome or valuable information from multi-source data for better earth exploration becomes a fascinating yet challenging problem when a list of remotely sensed data sources is available. Efficient Remote Sensing Images (RSI) classification has been the foundation for remote sensing applications. A range of classifiers are used to realize the classification of several types of remote sensing image targets. Better classification of targets is critical in both military and civilian sectors. In this paper, the focus is on summarizing the main advanced approaches used to enhance classification precision. The result shows that the Convolutional Neural Network outperforms all the other traditional classification methods.

Keywords: Feature extraction, Remote Sensing Images, Naïve Bayes, Random Forest, Convolutional Neural Networks, Data mining, Support vector machines.

1. Introduction

Remote sensing images have a wide range of applications. Remote data or images are those which are taken by devices and used for advanced learning. Thus, enabling us to understand various aspects in the respective fields. The major drawback of using such images is the noise and blur included in them. Thus, the major and the first step towards the learning of images involve noise removal. There are various types of noises added to the images due to environmental conditions, dust, haze, light attenuation, particles present in the sensor devices etc. The pre-processing of these images is a challenging task to the experts. The different varieties of noises [23] in the images are salt & pepper, Gaussian, Poisson noise etc. Salt & pepper noise is caused due to sudden and sharp variations in the signal. This kind of noise appears because of data transfer errors. It is represented by sparsely occurring white and black pixels. The Gaussian noise is distributed over the signal in equal proportions. It is a statistical noise having a Gaussian distribution. Gaussian distribution is defined as a function having a probability density function equal to that of normal distribution. The vibration at each point is autonomous of pixel-value intensity. Poisson noise is the noise produced when the sensor has not identified sufficient samples to provide observable statistical data. Shot noise appears to exist in discrete amounts such as light and electric current.

Various noise removal algorithms can remove the noises in the image. The noise in the images are normally removed by using the Linear and Non-Linear Filters. The different types of noise removal filters like the mean filters, median filters, wiener filters, Adaptive filters etc. The wiener filter's performance is better compared to mean and median filters to remove noise like poison and Gaussian from an image. In comparison, a mean filter provides the best results to remove salt and pepper noise [22].
Remote sensing (RS) images also include the astronomical images taken from satellites. These are used in various environmental studies on stars, planets and galaxies. Various applications and studies are also being carried out in this area. Classification of galaxies is also a prominent area of study. Astronomers are likely to work on a huge amount of data for their studies. Classification of such data, which includes bars, spirals, disks, etc., will help predict galaxies' evolution. The bars are indicative of galaxies reaching full maturity predicting that the formal years are about to end. Remote sensing images also find applications in underwater images to study the mammals that survive in the deep sea. These images are also used in the study of soil in oceans, seas etc. They also play a vital role in the study of the existence of various species in the water. The galaxy view or astronomical data set is widely available as an open-source from the SDSS (Sloan Digital sky survey) or the Catalina survey. These depict the screenshot of the sky using a telescope. These telescopic images are very blue and greyscale, making a study of the stars and sun a tedious task. Thus various methods for error or noise removal need to be added to gain knowledge from these images. Classification of these images is of high importance to the area of physics. The Deep Neural Network architecture has been used by Sheelu Abraham et al. [21] to classify the barred and non-barred galaxies. The Neural network helped in classifying the images with an accuracy of about 95%. Other advancements in astronomical studies are still advancing and have a wide development area soon using neural networks.

Since the 1960s, Geographical Information Systems (GIS) have been evolved for developments in land use and military needs. Such systems were developed to view and control both vector and raster data. The key feature of GIS was to combine graphical and non-graphical data to render spatial analyses and queries. The two approaches (RS and GIS) were used exclusively with each other until recent years. However, only one, either GIS benefitted from RS or vice versa. With advances in computing, several investigators have used GIS and RS strategies to improve each other's activities. In recent years, several methods are adopted for working on the classification of SAR (Synthetic Aperture Radar) as the image target is enormous. Su et al. [14] suggested SAR image categorization strategy using Markov Random Regions with a multi-scale regional relationship calculus model. Karine et al. [9] used the SIFT keypoint method for accurate identification of the SAR image reference. The Polar Scale-Invariant Transform Function (PSIFT) is proposed by Zeng et al. [16], which renounces the measurement of dominant orientations, and SAR-DM defines the goals. These algorithms will carry out the classification activities to some degree, although the precision is not achieved. Furthermore, with the increasing growth of artificial intelligence, approaches focused on profound learning are being introduced to classify SAR image targets. David et al. [12] proposed Convolutional Neural Networks (CNN) training method for SAR ATR. Ali et al. [7] have attempted to present a CNN-based method with a Convolutional Auto-Encoder. Chen et al. [2] introduced a sparse CNN based auto-encoder solution. Although these techniques can attain a high level of accuracy, they are of low performance about Deep Learning.

Remote sensing also provides background data information, especially when discussing the biophysical framework in which people live and work [10]. Remote sensed data include alternate portrayals of reality. In the opposite maps, remote sensed data shows facts, i.e. the satellites get pictures of what's on Earth. Furthermore, remote sensing can enhance geo-referenced social data by clarifying many aspects of the climate, ranging from land cover to soil moisture or weather and observational data.

2. Remote sensing classification process

Categorizing remote data is a concern since many variables, such as landscape complexity, remote data selected, image processing, and classification methods, can influence a classification's functioning. Remote sensing is a dynamic process that involves multiple image analyses. Determining classification scheme, choice of training objects, preprocessing of the image, features to be extracted, applying suitable classification methods, post-
classification storing, and accuracy assessment may be the key stages in the classification of images.

2.1 Remotely sensed data selection

Remote sensed data differ in geographical, isotopic ratios, statistical and temporal resolution values. Recognizing the advantages and disadvantages of various sensor data types is vital for selecting an appropriate remote sensed image for classification. The most important factors affecting remotely sensed data collection are scale, image quality and consumer needs. Consumer need depends on the type of classification and research field, thereby the collection of adequate remote data is affected.

2.2 Selection of a classification method and sample preparation

A classification system is designed for image resolution of a chosen remote sensed images by integration with prior research, available methodologies for image recognition and time limits based on user prerequisites. In certain instances, a hierarchical classification system may be introduced to satisfy a particular requirement.

2.3 Data preprocessing

Image pre-processing can include identifying and restoring bad lines, geometric re-evaluation, image recording, radiometric adjustment, spectral analysis, and elevation adjustment. When using separate ancillary data, translation of data between different origins or settings is often needed for evaluating the output of data. The precondition for mixing multiple source information in a classifier is called as assessment scale amplification or remote sensing image recording.

2.4 Feature extraction

Choosing suitable variables is a crucial step towards the efficient implementation of the classification of images. Many possible variables, including reflectance spectra, vegetation index, translated images, compositional or contextual data, multi-temporal pictures, multi-sensor images and satellite images may be used in classification tasks.

2.5 Selection of a suitable classification method

Many considerations such as remote sensing spatial resolution, various data origins, classification scheme, and classification software accessibility need to be considered when choosing a classification method. There are various forms of classifications that have their own merits. Various classification methods are discussed in the below sections.

2.6 Post-classification processing

Standard per-pixel classifiers in the classification maps will lead to consequences of ’salt and pepper.’ To minimize noise a filter is sometimes added. Most classification of images is based on reflectance spectra which are remotely sensed. For instance, the spread of forests in
mountain terrain has to do with elevation, slope and facets.

2.7 Evaluation of classification performance

A significant step in the classification is the assessment of the classification tests. Different methods can be applied, ranging from an expert-led quality approach to a numerical accuracy assessment led on survey methods. The classification system's performance assessment is focused on precision, robustness, reliability, ability to fully exploit the data statistical validity, consistent applicability, and objectivity.

3. Advanced classification algorithms

Improving the spatial resolution of remote data does not favor the retrieval of more precise functionality. The methods used for the classification of photographs play a significant role for better precision. Modern classifiers are widely used to classify images and outscore traditional classifiers. They are best suited for integrating non-spectral data into the classifier. Table 1 represents a list of a classification algorithm for remote sensing images.

3.1 Artificial Neural Networks (ANN)

ANN mimics several human brain functions to predict distinct features of the image. ANN-based classification utilizes a non-parametric approach, making it easy to integrate additional data into the classification system. This enhances the accuracy of the classification. An ANN has several layers in it, each with a group of processing units known as neurons. All neurons in a given system are connected to all the other neurons in subsequent layers through weighted connections. ANN algorithms are highly successful when they appear to be repetitive assignment. A well-trained network will be able to classify highly complex data.

3.2 Classification trees

Classification Tree [8] is a multivariate, incremental, and evolutionary pattern recognition system based on an approach to hierarchical law. The preceding elements of a CT: the root, the non-terminal and the terminal. It predicts membership by segmenting a dataset repeatedly using binary dividing laws. Such laws are based on 'impurity' and are derived using statistical methods from the training data. If pixels containing a given node relate to the same group, then the node is ideal and the impurity is 0. The left branch is chosen unless the logical condition is met at a given node; or else the right branch is observed. The cycle goes on till the node is empty, or a terminal node is reached.

3.3 Support Vector Machine (SVM)

SVMs are among the most recent advances in AI and rely on the concepts of statistical theory. Furthermore, SVMs have lived up in classification precision compared to most other image classification algorithms. SVMs are binary classifiers that distinguish two categories to optimize the difference in them by applying appropriate hyper-plane division to multidimensional function space training data.
Analysis of classification and regression problems normally use the support vector machine algorithm. SVM mainly suits for binary classification. SVM involves the representation of points in space which is separated into two classes using a hyperplane. The new feature/point that needs to be mapped is then predicted to belong to any category based on the position of the point in the space. SVM finds its application in remote sensing images, text and hypertext classification, handwritten character recognition etc. The main objective of SVM is to identify a hyperplane that clearly separates the dataset into two different classes. The pictorial representation of SVM is shown in figure 1.

Two processes are frequently used to tackle binary classification with SVMs and address multi-class issues in remote sensing frameworks: One-Against-All (1AA); One-Against-One (1A1). Nonetheless, the 1A1 strategy consists of many binary SVMs, and thus complex calculations than the 1AA method. SVM's original output reflects the distances between each pixel and the optimal hyper-plane separation. Both positive (+1) and negative (-1) votes for a particular class. They are scaled up, and a clear majority vote determines a particular pixel's final class membership.

| Table 1. Remote Sensing Data Classification Algorithms |
|-----------------|-----------------|-----------------|
| Strategy        | Characteristics | Example         |
| Unsupervised    | Every pixel is believed to be pure and is usually classified as a single form of land use cover | k-means, ISODATA, SOM, k-nearest Neighbors [1], random forests, support vector machine, genetic algorithms |
| Supervised      | Each pixel is regarded mixed, and the areal proportion is estimated for each class | Fuzzy classification, neural networks, regression modeling, regression tree analysis [15], spectral mixture analysis, fuzzy-spectral mixture analysis |
| Object based    | Geographical structures are regarded the basic unit in place of individual pixels | E-cognition, ArcGIS Feature Analyst |
3.4 Fuzzy Classifiers

Fuzzy classifiers convey each pixel’s membership in each class. The Fuzzy set membership is determined using a particular algorithm, focused on standardized Euclidean Distance from the signature mean. The mean of a signature is the optimal level for the class where there is a Fuzzy set membership. As distance rises, the membership of fuzzy sets reduces, until it reaches the user-defined distance where the membership of fuzzy sets decreases to 0. Due to the absence of a criterion for selection at hand, it is often difficult to categorize the classifier with the best results for a particular dataset.

3.5 Random Forest Classifier

Random forest is a supervised learning algorithm that operates on constructing multiple decision trees during the dataset's training. The output class is predicted based on the individual classes predicted by all the decision trees. Random forest reduces overfitting of the data. These individual decision trees act as an ensemble. Each individual tree produces an output class. The final result of the random forest is based on the maximum class predicted by the ensemble. Random forest finds its application in feature extraction, feature ranking, image processing, text characterization etc. There are mainly two prerequisites for random forest to perform well. Firstly, to avoid random guessing, there needs to be true signals of the features during the training of dataset. Secondly, the correlation between the outputs of the individual decision trees should be minimum. Random forest finds its application in feature extraction, feature ranking, image processing, text characterization etc.

3.6 Naïve Bayes Classifier

Naïve Bayes Algorithm is based on Bayes theorem. According to this theorem the features extracted from the data set is considered to be independent of each other. Naïve Bayes classifier is scalable, thus requiring large number of parameters linear in the number of variables (features/predictors) in a learning problem. Bayes’ Theorem provides the probability that an event would occur concerning another event that has already occurred. Bayes’ theorem can be mathematically expressed as:

\[ P(A/B) = \frac{P(B/A)P(A)}{P(B)} \]  

(1)

where A and B are events.

For a Naïve Bayes classifier, two matrix main matrix required are the feature matrix and response vector matrix. Feature matrix consists of all dependent feature values of the dataset. The response vector contains the class value for prediction. This algorithm assumes that all features are independent and make equal contribution to the prediction.

3.7 Convolutional Neural Networks

CNN is a deep learning network with following benefits over models: CNNs uses a convolution procedure on the pixels of an image to retrieve the image's features. CNN can display image data, and gather image information from large data easily.

3.7.1 Structure of CNN

CNNs structure can effectively handle nonlinear problems (e.g., rotation and image
translation). Compared with ordinary images, remote sensing images re-emphasize richer spectral information and spatial details representing their design, color and size [4][3]. CNNs specifically learn to analyze two-dimensional (2D) shapes that transform the original input into the output. Inside a CNN, each neuron is connected within the preceding layer neurons, thereby reducing weights in the network. As shown in Figure 2, this network consists of the convolutional layer, pooling layers and a fully connected output layer. The results of the classification are generated from the final output layer. A CNN's core layer is the convolutional layer [18]. The convolution kernel/filter is an array of weight matrix. The characteristic map formed by convolving the input matrix and filter can be the input for the convolutional neural network's successive stages. Furthermore, the neurons in each input image convolve with the kernel in every convolutional layer. A CNN can have any number of convolutional layers.

![Figure 2. Structure of CNN.](image-url)

### 3.7.2 Pooling layer
The pooling layers have average, median, and unpredictable arrays for pooling. There are different forms of pooling: superimposing pooling and pooling of spatial pyramids. Regardless of the type of pooling layer utilized, pooling layers tend to capture unique characteristics. Thus, the network's successful features can still be discovered, even if a small amount of input data changes occur.

### 3.7.3 Fully connected layer and output layer
Multiple hidden layers form an entirely interconnected network. Each concealed layer includes correlated neurons. Fully connected layer refers to the One-dimensional (1D) feature vectors that are generated by straightening feature maps after the convolutional and pooling operation. A fully-connected layer aims to transform these characteristics into a lower dimensional space and align with the output layer. The Soft-max feature and the SVMs are widely used in CNNs today [19].

### 3.7.4 Activation and loss functions
The activation function is usually nonlinear which helps the network can learn on nonlinear layer-wise mapping. Sigmoid, Rectified Linear Unit (ReLU) [20] and Max-out functions...
[21] are common activation functions. A loss function, also called a cost function or an objective function, is used to depict the extent of prior discrepancies between the model's predicted value and the actual value.

4. Remote sensing data trends

In 1972, Landsat-1 was the first spacecraft to collect Earth reflection at a resolution of 60 meters. During this time the dual image classification methods[11] available were the unsupervised [6] and supervised detection. These algorithms were appropriate for the spatial resolution. Later, the OBIA (Object-Based Image Analysis) [13] was evolved for the processing of digital images. Over the decades there has been a growing need for remotely sensed data. There are hundreds of remote sensing data criteria, such as for food, climate, and public health safety. Satellite imagery aims at greater spatial resolution and a broader range of frequencies to meet their requirements. However, photos with higher resolution will not guarantee greater land coverage. The algorithm used for the classification of photographs are a very important consideration for better precision.

Remote sensing strategies are important for socio-economic preparation and ecological initiatives such as deriving soil cover information for land use. These spectral classifiers are still the prevalent techniques for categorizing remote sensing imagery due to their computational simplicity and deployment.

![Figure 3. The architecture of CNN used for classification proposed by Sheelu [17].](image)

5. Experimental results

A sample experiment was conducted to validate the above-stated algorithms. The experiment is carried out using the galaxy image from the Sloan Digital Sky (Figure 3). The data set consists of images of 9346 galaxies. The classification is carried out based on two-class labels barred and unbarred images. The data set consists of 3864 barred galaxies and the rest being unbarred. A significant fraction of disk galaxies near the Universe have bars [24]. Also, hydro-dynamical simulations indicate that bars have a clear impact on driving the host galaxy's evolution by transporting material between the disc and the bulge, thereby redistributing the angular momentum of baryonic and dark matter components of disc galaxies. As a consequence, bars play a significant role in the secular evolution of the host disc galaxy.
The standard method of identifying bars in a galaxy is through human intervention, but this could create a high human error range. For automation, a machine learning algorithm using a convolutional neural network was introduced. The data set was split as 80:20 for training and testing purposes. The architecture of the network proposed by Sheelu Abraham et al. [17] is shown in Figure 3. Initially, the training was carried out for 25 epochs which produced an accuracy of 79.98%. The training was extended to 100 epochs which helped to achieve the topmost accuracy of 94%. The detailed accuracy of different classification algorithms is shown in Table 2.

Table 2. Accuracy of different classification algorithms applied in Sloan Digital Sky Data set.

| Supervised Algorithm          | Accuracy  |
|-------------------------------|-----------|
| Convolutional Neural Network  | 94%       |
| Support Vector Machine        | 82.33%    |
| Naïve Bayes Algorithm         | 79.14%    |
| Random Forest                 | 74.88%    |
| K- Nearest Neighbor           | 73.01%    |

Figure 5. Percentage of accuracy for different classification algorithms.
The same data set from Sloan Digital Dataset was used to validate the different algorithms of supervised learning. Figure 5 plots the graph showing the accuracy level of different classification algorithms used for the experiment. The maximum accuracy was achieved by the CNN algorithm of 94% with 100 epochs. KNN recorded the least with 73.01%. Other algorithms performed at an average level SVM with 82.33%, Naïve Bayes recorded an accuracy of 79.14, and Random Forest with 74.88. The application of these algorithms differs in terms of application and data set. An experiment with another dataset may not record a similar trend in the accuracy level. Thus, based on the type of dataset, the best-supervised classification algorithm needs to be selected.

6. Conclusions

This paper provides a complete outline of the different remote sensing images and the research areas where they are applied. Vast knowledge on how the remote sensing images are prepared for the learning process and the different algorithms involved in the classification and clustering of the remote sensing images are also discussed. A detailed study on applying Convolutional Neural Network on the remote sensing images to extract features from these images is also explained.

7. References

[1] Blanzieri E., Melgani F. (2008) - Nearest neighbor classification of remote sensing images with the maximal margin principle. IEEE Transactions on Geoscience and Remote Sensing, 46: 1804-1811. doi: http://dx.doi.org/10.1109/TGRS.2008.91609 0.
[2] Chen, S., Wang, H.: ‘SAR target recognition based on deep learning’. Int. Conf. on Data Science and Advanced Analytics, Paris, France, 2015, pp. 541–547.
[3] Chen, Y., Lin, Z., Zhao, X., et al.: ‘Deep learning-based classification of hyperspectral data’, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 2014, 7, (6), pp. 2094–2107.
[4] Chen, Y., Zhao, X., Jia, X.: ‘Spectral-spatial classification of hyperspectral data based on deep belief network’, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 2015, 8, (6), pp. 2381–2392.
[5] Feng, J., Jiao, L.C., Zhang, X., et al.: ‘Bag-of-visual-words based on clonal selection algorithm for SAR image classification’, IEEE Geosci. Remote Sens.Lett., 2011, 8, (4), pp. 691–695.
[6] Gonçalves M.L., Netto M.L.A., Costa J.A.F., Zullo Júnior J. (2008) - An Unsupervised Method of Classifying Remotely Sensed Images using Kohonen Self organizing Maps and Agglomerative Hierarchical Clustering Methods. International Journal of Remote Sensing, 29:3171-3207.doi:http://dx.doi.org/10.1080/014311607014421 46.
[7] Housseini, A.E., Toumi, A., Khenchaf, A.: ‘Deep learning for target recognition from SAR images’. 2017 Seminar on Detection Systems Architectures and Technologies, Algiers, Algeria, 2017, pp. 1–5.
[8] Jiang H., Zhao D., Cai Y., An S. (2012) - A Method for Application of Classification Tree Models to Map Aquatic Vegetation using Remotely Sensed Images from Different Sensors and Dates. Sensors, 12: 12437-12454.doi:http://dx.doi.org/10.3390/s120912437.
[9] Karine, A., Toumi, A., Khenchaf, A., et al.: ‘Saliency attention and SIFT keypoints combination for automatic target recognition on MSTAR dataset’. IEEE Int. Conf. on Advanced Technologies for Signal and Image Processing – Atsip’2017, Fez, Morroco, 22–24 May 2017
[10] Lillesand T.M., Kiefer R.W., Chipman J.W. (2004) - Remote Sensing and Image Interpretation. Ed. 5. John Wiley & Sons Ltd.
[11] Lu D., Weng Q. (2007) - Survey of Image Classification Methods and Techniques for Improving Classification Performance. International Journal of Remote Sensing, 28: 823-870.doi: http://dx.doi.org/10.1080/014311606007464 56.
[12] Malmgren-Hansen, D., Engholm, R., Pedersen, M.O.: ‘Training convolutional neural networks for translational invariance on SAR ATR’. Proc. of EUSAR 2016: European Conf. on Synthetic Aperture Radar, Hamburg, Germany, 2016.

[13] Myint S.W., Gober P., Brazel A., Grossman-Clarke S., Weng Q. (2011) - Per-pixel vs. Object-based Classification of Urban Land Cover Extraction using High Spatial Resolution Imagery. Remote Sensing of Environment, 115: 1145-1161. doi: http://dx.doi.org/10.1016/j.rse.2010.12.017

[14] Su, X., He, C., Feng, Q., et al. ‘A supervised classification method based on conditional random fields with multiscale region connection calculus model for SAR image’, IEEE Geosci. Remote Sens. Lett., 8, (3), pp. 497–501

[15] Xu M., Watanachaturaporn P., Varshney P., Arora M. (2005) - Decision Tree Regression for Soft Classification of Remote Sensing Data. Remote Sensing of Environment, 97: 322-336. doi: http://dx.doi.org/10.1016/j.rse.2005.05.008.

[16] Zeng, L., Zhou, D., Liang, J., et al.: ‘Polar scale-invariant feature transform for synthetic aperture radar image registration’, IEEE Geosci. Remote Sens. Lett., 2017, 14, (7), pp. 1101–1105

[17] S Abraham, AK Aniyan, AK Kembhavi, NS Philip, K Vaghmare ‘Detection of bars in galaxies using a deep convolutional neural network’, Monthly Notices of the Royal Astronomical Society 477 (1), 894-903

[18] Jia Song, Shaohua Gao, Yunqiang Zhu & Chenyan Ma, ‘A survey of remote sensing image classification based on CNNs’, BIG EARTH DATA 2019, VOL. 3, NO. 3, 232–254.

[19] Liu, W., Wen, Y., Yu, Z., & Yang, M. (2016). Large-margin softmax loss for convolutional neural networks. ICML, 2(3), 7.

[20] Hara, K., Saito, D., & Shouno, H. (2015). Analysis of function of rectified linear unit used in deep learning. International Joint Conference on Neural Networks, Killarney, Ireland.

[21] Goodfellow, I. J., Warde-Farley, D., Mirza, M., Courville, A., & Bengio, Y. (2013). Maxout networks. arXiv preprint arXiv:1302.4389.

[22] Pawan Patidar, Sumit Srivatsava,” Image De-noising by Various Filters for Different Noise”, International Journal of Computer Applications, Volume 9– No.3, November 2010

[23] Sukhjinder Kaur ‘Noise Types and Various Removal Techniques’ International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 4, Issue 2, February 2015.

[24] Knapen J. H., 1999, in Beckman J. E., Mahoney T. J., eds, Astronomical Society of the Pacific Conference Series Vol. 187, The Evolution of Galaxies on Cosmological Timescales. pp 72–8 1148–1156.