A REVEW OF IMAGE CLASSIFICATION TECHNIQUES BASED ON PERFORMANCE METRICS

Pavithra, K Kalaivani, K UlagaPriya

Department of Computer Science and Engineering

Vels Institute of Science, Technology and Advanced Studies, Chennai, India

ABSTRACT:

The objective is to classify satellite images and highlight the differences either based on object, type of land cover or land usage. Image classification is the essential phase of digital image analysis. The main purpose of image classification is to automatically classify each and every pixel in an image into multiple classes. Typically for remote sensing applications, multispectral data is utilized to perform the classification, and furthermore the spectral pattern present in the image is the basis for categorisation. This paper presents the fundamental ideas of image classification, performance metrics and comparison of various image classification techniques used by researcher in various application.The performance metrics such asAccuracy, Precision, Recall or Sensitivity, specificity and Kappa Statistic are compared and analyzed. Image classification techniqueswere used effectively in many fields such as agriculture, remote sensing, medical imaging and so on.

Keywords: Image Classification, Metrics, Artificial Neural Network, Decision Tree, SVM, Random Forest.

1. Introduction

The purpose of classification is to cluster all pixels in a digital image into one among numerous land cover classes.

Image classification is to categorize pixels within the image such as territories, waterbody, green vegetation, exposed soil, rough regions, cloud, shadow.Image classification is a method of mapping statistics to cryptogram.So as to group a lot of information into various classes, the connection between the information and the classes into which they are characterized must be surely known.

2. Types of Classification :

There are two types of classification based on training samples namely: Supervised classification and unsupervised Classification.

2.1. Supervised classification

It is the process of classifying the pixels in an image based on the training set containing known labels for each category.In this classification the identity and location of some of the land cover(e.g territories, waterbody, green vegetation, exposed soil, rough regions, cloud) are known a priori through a combination of aerial photography, map analysis, and personal experience.The process of using samples of known informational classes (training sets) to classify pixels of unknown identity. Some of the algorithms which are used by the researchers were minimum distance to means algorithm, parallelepiped algorithm, maximum likelihood algorithm

2.2 Unsupervised classification

It is the process of classifying the pixels in an image based on the parameters and the outcome of classification will be in the form of clusters. The number of clusters depends upon the variance in the pixels of a digital Image. The inference from an image is not based upon the prior knowledge about the samples as in the case of supervised classification. This is usually preferred when the ground truth information is not well defined for the required dataset. Normally Kmeans clustering algorithm is used as unsupervised classifier for various applications.

3. Image Classification Techniques:

Some of the image classification techniques vastly used by the researchers for different iimage processing applications were Artificial Neural Network (ANN), Decision Tree(DT), Support Vector Machine(SVM) and Random Forest(RF)

4. Performance Metrics :

The Confusion matrix is one among the essentially unconstrained and simple measurements worn utilized for deciding the exactness and accuracy of the classes. The following are the results between the prediction and the actual.

True Positives (TP): True positives are the situations when the real class of the data point was 1(True) and the anticipated is also 1(True)

True Negatives (TN): True negatives are the instance, anticipated is become zero (false) when the real model of the information point was zero(false)

False Positives (FP): False positives are the instance, predicated is become one (true) when the real class of the information point was zero (false). It is false if the ideal has predicated wrongly and the model has predicated was a positive one when it is positive. (1)

False Negatives (FN): False negatives are the instance, predicated is become zero (false) when the real class of the information point was one (true). It is false if the ideal has predicated wrongly and the model was a negative one when it is negative. (0)[8]

4.1 Accuracy:

In classification issues, Accuracy is the measure of exact computations done by the perfect over all typescalculation made.[4]Accurate calculation in the numerator (True positives and True Negatives) All types of calculation done by the algorithm in the denominator[9].

If the goal variable classes in the information are almoststable is measured by accuracy.

Ex: 70% area in SAR images data ismangrove& 30% -non-mangroves.

A model which predicts whether a new SARpicturebemangrove or non-mangrove, 97% of times effectively is a decent assesswithin this precedent.

If the goal variable classes within the information are mainstream of individual class then accuracy should not be used as a measure.[8]

Ex: In Mangrove vegetation discovery venture with wide territory, just 5% zone has mangrove vegetation. Assume our model is terrible and predicts each case as non-mangrove. In doing all things considered, it has orchestrated those 9% non-mangrove successfully and 5% mangrove as Non-mangrove. Directly in spite of the way that the model is shocking at envisioning vegetation, The exactness of such a horrendous model is moreover 95%.

4.2 Precision:

Precision is a assess that says what percentage of vegetation that detected like mangrove, really had mangrove. The predicted positives (Vegetation predicted as mangrove are TP and FP) and the vegetations really having a mangrove are TP[5].

4.3 Recall or Sensitivity:

Recall or Sensitivity is known as true positive rate. The chance of finding measures the amount of actual positives that are correctly recognized as such (Ex : The proportion of mangrove vegetation who are correctly identified)

4.4 Specificity:

Specificity is otherwise called genuine negative rate. Particularly it is a measure that discloses to us what extent of vegetations that isn't mangroves, were anticipated by the model as non-mangrove.

4.5 Kappa Statistic:

One more accuracy meter is the kappa coefficient. It is a assess of how the classification results evaluate to value assign by option. It is capable of get values from 0 to 1. If kappa coefficient is 0, there is confirmation that there is a no agreement between the classified image and the reference image. If kappa coefficient is 1, then the classified image and the reference image are totally matching.

5. Discussion

Here we are considering two main metrics for comparing various image classification techniques.

1) Overall Accuracy
2) Kappa Statistic

Performance report of various mage classification techniques

Algorithms	Overall Accuracy	Kappa Statistic	Reference
Random Forest	~81.10%	~0.76	[1],[8]
Decision Tree	~90%	~0.87	[2],[8]
Support Vector Machine	~94%	~0.88	[2],[4],[7],[8],[9],[10]
Artificial Neural Network	~99.20%	~ 0.98	[1],[3],[5],[10]

Advantages and Limitations of various image classification techniques

Algorithm	Advantages	Limitations
Artificial Neural Network	• A non-parametric classifier. • An universal functional approximator with arbitrary accuracy. • It is a data driven selfadaptive technique • Can handle noisy images	• Semantically poor. • Time consuming process • Problem of over fitting. • Difficult in finding the best type of network architecture.
Decision Tree	• A non-parametric classifier. • An extensive design and training is not essential • Better inference can be made based on hierarchical associations between input variables	• Calculation becomes complex when there more correlation in various outcome
Support Vector Machine	• Nonlinear transformation • It provides a good generalization capability. • Eliminates over fitting • Less computational complexity. • Easy to handle decision rule complexity and Error frequency.	• Traning prcocess requires more time. • Difficulty in understanding the Structure of algorithm. • Because of non- linear data, it is difficult to find the optional parameters.
Random Forest	• A non-parametric classifier • Very high accuracy	• Time com=nsuming for spectral images

6. Conclusion

This paper has presented an overview of various image classification techniques in supervised and unsupervised methods like Artificial Neural Network, Decision Tree, Random Forest and SVM. The analysis and usage of different classification techniques are elaborated. The various performance metrics, which are used to measure the quality of the classified image were reviewed and analysed.

7. References:

1) "Random forest classification for mangrove land cover mapping using Landsat 5 TM and ALOS PALSAR imageries" -RomieJhonnerie, Vincentius P. Siregar, BismanNababan, Lilik Budi Prasetyo, and Sam Wouthuyzen
2) "Exploring changes of land use and mangrove distribution in the economic area of SidoarjoDistrict,East java using multiple-temporal Landsat images"-NoridaMaryantikaChinsu Lin-INPA 2017
3) "Hyperspectral Data for Mangrove Species Mapping: A Comparison of Pixel-Based and Object-Based Approach"-Muhammad Kamal and Stuart Phinn- MDPI 2011
4) "An Object-Based Classification of Mangroves Using a Hybrid Decision Tree—Support Vector Machine Approach" - Benjamin W. Heumann- MDPI 2011
5) "Retrieval of Mangrove Aboveground Biomass at the Individual Species Level with WorldView-2 Images"-Yuanhui Zhu, Kai Liu , Lin Liu , Shugong Wang and Hongxing Liu- MDPI 2015
6) "Per-pixel and sub-pixel classifications of high-resolution satellite data for mangrove species mapping" - Kasturi Devi Kanniah - Applied GIS 2007
7) "Mapping Mangrove Forest from LiDAR Data Using Object-Based Image Analysis and Support Vector Machine: The Case of Calatagan, Batangas"-Lawrence Charlemagne G. David and Alejandro H. Ballado, Jr. – IEEE 2015
8) "Artificial Mangrove Species Mapping Using Pléiades-1: An Evaluation of Pixel-Based and Object-Based Classifications with Selected Machine Learning Algorithms"-DezhiWang,Bo Wan, PenghuaQiu, QinghuaGuo, XincaiWc – MDPI 2018

9) "Textural–Spectral feature-based species classification of mangroves in Mai Po Nature Reserve from Worldview-3 imagery." -Ting Wang, Hongsheng Zhang, Hui Lin and Chaoyang Fang – MDPI 2015
10) "Satellite images for monitoring mangrove cover changes in a fast growing economic region in southern Peninsular Malaysia" -Kasturi Devi Kanniah,AfsanehSheikhi,Arthur P. Cracknell,HongChingGoh,Kian Pang Tan,ChinSiong Ho and FateenNabillaRasli – MDPI 2015
11) "A Survey on Image Classification Approaches andTechniques" - Pooja Kamavisdar[1], Sonam Saluja[2], SonuAgrawal – IJARCCE -2013