International Journal of Engineering Science Invention(IJESI)
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org ||Volume 6 Issue 12|| December 2017 || PP. 55-70

An Exploration of Classification Approaches For Real Life Image Sets

Rajni Mehta[(1)], Sarbjeet Kaur Bath[(2)], Anuj Kumar Sharma[(3)]
Research Scholar, PTU Jalandahar, Punjab, India
EE Deptt & GZS COET, Bhatinda, Punjab, India
CSE Deptt. & OITM, Hisar, Haryana, India

Abstract: *Classification is the imperative and challenging task in computer vision. Classification is based on description, texture or similarity of items or things. Image classification refers to the labeling of images into one of a number of predefined categories. Pixels are the unit represented in an image. Image classification is a process that understands the image and extracts the information that can be used for other tasks. The image classification process comprised of different phases as image acquisition, image pre-processing and image segmentation. Several classification techniques have been developed for classifying images. This paper presents a overview on various image classification techniques such as K-Nearest Neighbor (KNN), Support Vector Machine (SVM), classification and regression tree (CART),ADABOOST, K-Means, Naive Bayes, Decision Trees, ISODATA, Random Forest and DECORATE. Work done in the field of image classification has been presented in this paper. Various steps involved in the image classification process have also been discussed in the present paper.*

Keywords: *Image classification; Image processing; Image segmentation; Feature Extraction; SVM; CART; K-Means.*

Date of Submission: 11-12-2017 Date of acceptance: 22-12-2017

I. Introduction

Image classification is the most crucial step in all image analysis tasks. Blurred and noisy image content makes image classification more complex and difficult to classify. Image classification assigns pixels in the image to categories or classes of interest and this process maps numbers to symbols. Image classification has become a challenge with respect to the use of methods and techniques in exploiting image processing results, pattern recognition results and classification methods and, subsequently validating the image classification result into expert knowledge. Numerous methods have been developed to classify images that have proven to be good classifiers but when the image contains blurry and noisy content, they fail to provide satisfactory classification result. Image classification is of two types, supervised and unsupervised. In general, this is a final stage of pattern matching. It is a process that describes the percentage of accuracy in pattern recognition. Feature extraction is another vital stage in pattern matching. Pattern matching identifies the regions of grayscale images that match a known reference pattern or a template.

Supervised classification is a classification process in which an image analyst supervises the process of pixel classification, and then user specifies the pixel values such as the pixel represents the specific class or training sites. Algorithms have been developed to use these pixel values from these training sites for classifying the whole image. User selects training sites on the basis of knowledge and sets bounds for similar pixels that can be grouped in one class; this process is known as training. Then classifier uses training sites for classifying other images.

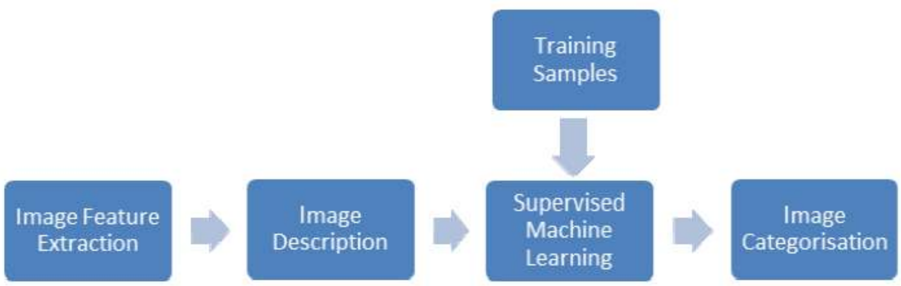

Fig 1: Supervised Classification Process [63]

The steps in supervised classification approach are:
- Analyst identifies the training sites for each informational class, each site having unique identifier.
- Analyze pixels within the training sites and identify pixel signatures.
- Classify each image by comparing pixel signature with the training sites.

Unsupervised classification is a method that examines a large number of unknown pixels and divides it into a number of classes based on natural grouping present in image values. This process categorizes a digital image after processing it on the basis of image statistics without availability of training sites or a-priori knowledge of the area. Training sites are selected based on the knowledge of the user. These groups are known as a cluster and this process is called clustering. In this user decide how many clusters he wants. The unsupervised classification used when no trained pixels are available.

Fig 2: Unsupervised Classification Process [63]

Unsupervised classification approach works as:
- Separates data into groups with clustering.
- Classify all pixels on the basis of clusters.
- Assign a label to each cluster
- All these steps are iterated till user gets the desired results.

Supervised classification generally performs better than unsupervised classification IF good quality training data is available whereas Un-supervised classifiers are used to carry out preliminary analysis of data prior to supervised classification.
Following figure shows the image classification process:

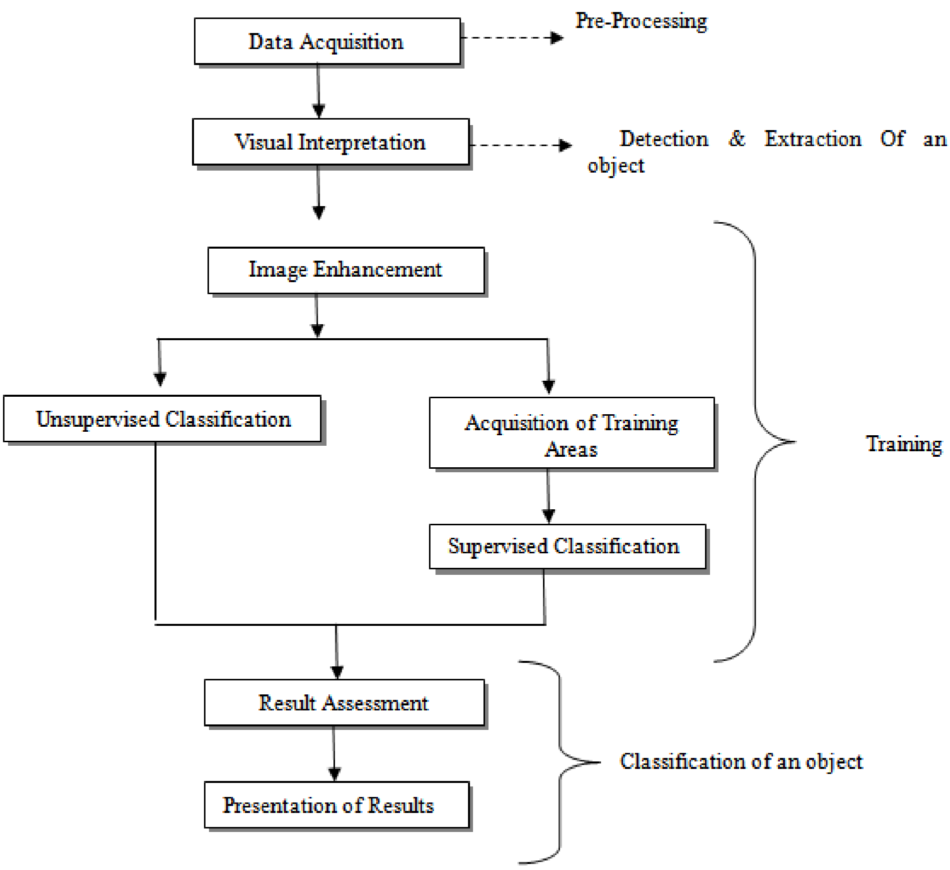

Fig 3: A Typical Image Classification Process

Classification process consists of the following steps:
A. Pre-processing: this step focuses on enhancing the quality of the digital image noise removal, image masking, atmospheric corrections etc. and to identify the data of interest.
B. Detection and extraction of an object: this step detects the locality and other features of moving object image captured from the camera and then extraction stage captures the unique characteristics by observing detected object and guessing the route of the object in the image plane.
C. Training: Selection of the specific attribute, which best defines the pattern.
D. Classification of the object: object classification is a process that assigns pixels in the image to categories or classes of interest and matches the image patterns with the target patterns.

II. Image Classification Functions

An image is considered as a representation of an object with specific properties that are employed in image processing. The medical image analysis tasks consist of three steps namely: Feature extraction and representation, Feature selection for classification and Feature and image classification.

2.1 Feature Extraction and Representation

Features are an important measurement for image understanding. A complete feature representation of the segmented region is an important issue for object classification and analysis [27][26]. Feature extraction and representation techniques are:

2.1.1 Statistical Pixel-Level (SPL) Features

These features provide quantitative information about the pixels within a segmented region. The SPL features include: mean, variance, and a histogram of the gray values of pixels in the region. Additionally, SPL features include the area of the region and information about the contrast of pixels within the region and edge gradient of boundary pixels [27].

2.1.2 Shape Feature

These features provide information about shape characteristic of the region boundary. The shape-based features include circularity, compactness, moments, chain-codes, and Hough transform. Morphological processing methods have also been used for shape description [27]. Shape features are mostly used for finding and matching shapes, recognizing objects or making measurement of shapes. The shape of an object is determined by its external boundary abstracting from other properties such as colour, content and material composition, as well as from the object's other spatial properties.

2.1.3 Texture Features

These features provide information about the local texture within the region or related area of the image i.e. spatial arrangement of the colors or intensities in an image. It is the defining characteristic of regions and critical in obtaining a correct analysis. Texture feature is calculated using the second-order statistical histogram or co-occurrence matrices. In a general sense, texture refers to surface characteristics and appearance of an object given by the size, shape, density, arrangement, proportion of its elementary parts. A basic stage to collect such features through texture analysis process is called as texture feature extraction. Due to the signification of texture information, texture feature extraction is a key function in various image processing applications like remote sensing, medical imaging and content based image retrieval

2.1.4 Relational Features

These features provide information about the relational and hierarchical structure of the regions related to a single object or a group of objects [27]. Relational features are based on the relative positions of different entities, regions, closed contours, or local features. These features usually include distance between features and relative orientation measurements. These features are very useful in defining composite objects using many regions or local features in images.

2.2 Feature Selection for Classification

A feature classification system can be considered as a technical mapping of the input feature related to the specific image to output variable that represents one category or class. Feature selection is the process of finding a most appropriate subset of features correlated to the final feature representation and classification.
Selection of correlated features due to dimensionality reduction in the classification task is able to improve the computational efficiency and classification performance since only well-correlated features are used in the classifier. The final set of features for classification can be determined through data correlation, clustering, and analysis algorithms to explore similarity patterns in the training data [27].

Feature selection for classification techniques are:

2.2.1 Linear Discriminant Analysis

Linear discriminant analysis methods are used to find a linear combination of features that can provide best possible separation among classes of data in the feature space. A linear combination of features can reduce dimensionality for classification and offer better classification performance with a linear classifier [27].

2.2.2 Principal Component Analysis (PCA)

PCA is used to identify patterns in data and highlights their similarities and dissimilarities, as high dimensionality of data makes this identification difficult. It analyzes data table with several dependent variables, which are inter-correlated. It extracts the important information from the data table and identifies a set of new orthogonal variables named principal components.

2.2.3 GA (Genetic Algorithms)-Based Optimization

GA is a robust optimization search method, which is based on the principles of natural selection. GA enhances feature selection for classification performance by utilizing prior information and using competition for survival. A basic feature of GA is the ability to adapt to the specific parameter issues. These parameters are usually encoded as binary strings that are associated with the goodness or fitness measurement [27].

2.3 Feature and Image Classification

Selected features for image representation are applied on object recognition and characterization. In the medical imaging analysis, features and measurements can also be used for region segmentation to extract meaningful structures, subsequently, interpret the result using knowledge-based model and classification methods [26]. Feature and image classification techniques are:

2.3.1 Statistical Classification Methods

Statistical classification methods are generally defined into two categories: unsupervised and supervised approach. The unsupervised methods cluster the data based on their separation in the feature space. Clustering method such as K-means and fuzzy clustering methods are included in this approach. On the other hand, a supervised approach needs training data, test data, and class label to classify the data. Probabilistic methods like nearest neighbor classifier and Bayesian classifier are included in this approach [26].

2.3.2 Rule-Based Systems

A rule-based system analyzes the feature vector using multiple sets of rules that are designed to test specific conditions in the feature vector database to set off an action. The rules consist of two parts: condition premises and actions. They are generated based on an expert knowledge to deduce the action when the conditions are satisfied. Usually, a rule-based system consists of three sets of rules: supervisory or strategy rules, a focus of attention rules, and knowledge rules. The supervisory or strategy rules control the analysis process and provide the control actions include starting and stopping action. The strategy rules determine which rules would be tested during the analysis process. The focus-of-attention rule elaborates the specific characteristics during analysis process by identifying and eliciting the information or features from the database. Subsequently, the rules convey the information from the input (database) into the activity center where the implementations of knowledge rules are scheduled. Finally, the knowledge rules analyze the information related to the required conditions then execute an action that changes the output database [26].

2.3.3 Neural Network Classifiers

Artificial neural network paradigms for feature classification, object recognition and image interpretation namely back-propagation, radial basis function, associative memories, and self-organizing feature maps. At that time fuzzy system-based approaches have been applied in artificial neural networks for better classification and generalization result [26].

2.3.4 Support Vector Machine (SVM) for Classification

The relevance vector machine (RVM) uses regression and classification technique and exploits a Bayesian probabilistic principle. There are other models for pattern classification using theoretical approaches namely kernel-based classifier and linear programming perturbation-based methods [26].

III. Applications Of Image Classification

- Information classes extracted from a multiband raster image from image classification can be used to create thematic maps that emphasize a particular theme or special topic.
- Image classification can be applied in remote sensing for the production of Land Use and Land Cover maps at regional and global scale.
- Image classification is used for the purpose of pattern recognition by identifying the objects in an image and then various machine learning algorithms can be used to train the system for the change in pattern. Pattern recognition is used in computer aided diagnosis, recognition of handwriting, recognition of images etc.
- Image classification plays a vital role in the field of medical for diagnostic and teaching. Major application of classifications in medical field are industrial and biomedical surface inspection, for example finding the defects and disease, ground classification and segmentation of satellite or aerial imagery, segmentation of textured regions in document analysis, and content-based access to image databases etc.

IV. Image Classification Algorithms

4.1 SVM (Support Vector Machine)

SVM is a supervised learning algorithm that analyzes data and recognizes patters which is applied for classification and regression analysis. In this algorithm the local image features can be easily classified for analyzing the data. Support vector machine is formally said to be discriminative classifier. So it scales the high dimensional data. For the best performance of searching the algorithm, SVM kernels are used. SVM also uses the non -traditional data like trees and strings, which are used as input instead of feature factors. So both the small and large datasets can be applicable in SVM algorithm. To classify the MR image classification on the basis of local image features the SVM algorithm is best suitable to that. Also the noisy features are identified while analyzing the data and it is clearly removed by some other classification process.

4.2 K-Nearest Neighbor (KNN)

KNN is one of the simplest and non-parametric supervised algorithms. This is best suited where prior knowledge about the distribution of the data is not available appropriately and this algorithm can be used for both classification and regression predictive problems. The purpose of KNN algorithm is to classify a new object based on attributes and training samples. KNN is a lazy algorithm, as it does not use the training data points to do any generalization. Behavior of this algorithm depends on the value of K which is based on Euclidean distance in feature space (whereby k specifies the number of neighbors to be used.) so this classification is based on a majority vote of the k-nearest neighbors, It does not require a training step to be performed but can be tuned to determine the optimum value of k on which to base the classification. This number decides how many neighbors influence the classification. This algorithm has two approaches, first is if K=1, then the algorithm is simply called the nearest neighbor algorithm and second approach is if K=k where k is a positive integer specified along with a new sample then we find k nearest entries in our database that possess the same characteristics, that gives a new sample.

4.3 CART

The classification and regression tree (CART) algorithm is mainly used for the classification of different tissues in image mining, which is on the basis of several derived parameters. The recursive partitioning method used in the CART algorithm to introduce the tree based modeling which is later converted to the statistical mainstream. To select the optimal tree value the algorithm involves the cross validation scheme from some rigorous approaches. This technique is based on surrogate split that alternatively splits the algorithm and handles the missing values and it uses tree building algorithm that permit prediction or classification of cases.

4.4 ADABOOST

Adaptive Boosting (ADABOOST) was the first successful boosting algorithm developed for binary classification where a boosting method creates a strong classifier from a number of weak classifiers. It works on a sequence of weak learners on different weighted training data. This algorithm predicts original data set and gives equal weight to each observation. If prediction is incorrect using the first learner, then it gives higher weight to observation, which has been predicted incorrectly. Being an iterative process, it continues to add learner(s) until a limit is reached in the number of models or accuracy. Performance of this algorithm bounds on training data, complexity of weak classifiers and number of iterations. It is a flexible algorithm that can be combined with any learning algorithm. ADABOOST can fail if the weak classifier is either too complex or too weak.

4.5 K-Means

K-Means algorithm is said to be an unsupervised clustering algorithm. It works well for numerical data alone. The pixel-by-pixel image classification is possible by defining single and multiple thresholds. So that histogram statistics is used in this algorithm for the pixel-based classification. The main work of this process is to check whether the histogram is bimodal or not. If it is then the gray value will appear, otherwise the images get partitioned into several regions. The threshold of gray value can be determined using the peak values. However, it covers only the local minimum values. So the algorithm involves number of clusters for the optimization.

4.6 Naive Bayes

The Naive Bayes algorithm is the most powerful technique. It does the testing easily and the classification problems can be solved. It can be able to build a model quickly and gives better predictions. To find the missing data the naïve Bayes algorithm plays a major role. The unseen data can be easily predicted by characterizing the problem in naïve Bayes method. During the construction time and prediction time this algorithm separates the attribute values. The probability of each attribute in isolation process needs only the enough data. So, there is no need of more data collection in this algorithm. Finally, if the data has high correlated features the performance will be degraded.

4.7 Decision Tree

Decision tree algorithm is one of the classifier techniques which is in the form of tree structure. For classification and prediction, the powerful tools are available in this algorithm. It has four divisions such as Decision node, leaf node, edge and path. A single attribute is represented in the decision node. Leaf node defines the target attribute. Splitting of one attribute is edge and the path is a final decision. For continuous attribute this algorithm is not applicable.

4.8 ISODATA

Iterative Self-Organizing Data Analysis Technique (ISODATA) is an unsupervised algorithm like K Means algorithm. In the K-means method, the number of clusters K remains the same throughout the iteration, although it may turn out later that more or fewer clusters would fit the data. *ISODATA* Algorithm overcomes the drawbacks of K-Means which allows the number of clusters to be adjusted automatically during the iteration by merging similar clusters and splitting clusters with large standard deviations. The ISODATA represents a comprehensive set of heuristic procedures that have been incorporated into an iterative classification algorithm as it makes a large number of passes through the remote sensing dataset until specified results are obtained, instead of just two passes. It is self-organizing because it needs relatively little human input [1].

4.9 Random Forest (RF)

Random forest algorithm starts with the decision trees in such a manner that it aggregates the results of many randomly constructed classification trees where each decision tree will have different variance. This algorithm spans across a number of decision tree classifiers constructed on different sub-samples of the dataset. This algorithm follows ensemble approach (that works on divide and conquers methodology) and predictive accuracy to improve performance. Here, a group of “weak learners” can come together to form a “strong learner”. Among the different decision trees, ranking is made for the element, which is to be classified. If more number of decision trees point to a particular class, it will suggest that the element belongs to that class and hence the classification is achieved. Random forest tries to build multiple CART model with different sample and different initial variables.

4.10 DECORATE

Diverse Ensemble Creation by Oppositional Relabeling of Artificial Training Examples (DECORATE) is a meta-learner for building diverse ensembles of classifiers by using specially constructed artificial training examples [68]. *DECORATE* uses an existing “strong” learner (one that provides high accuracy on the training data) to build an effective diverse committee in a simple, straightforward manner. This technique is consistently more accurate than the base classifier, Bagging and Random Forests and boosting algorithms both on small training sets and larger training sets. In this algorithm, an ensemble is generated iteratively. Firstly, it acquires a classifier and then adding it to the current ensemble. Subsequently, at each successive iteration, the classifiers are trained on the original training data combined with artificial data. Artificial training examples are generated from the data distribution in each iteration where the number of examples to be generated is specified as a fraction, Rsize, of the training set size [18].

V. Comparison Of Various Image Classification Algorithms

Features	**ANN**	**Decision Tree**	**SVM**	**K-Means**
Algorithm	Supervised		Supervised	Unsupervised
Working	Imitates human mind. Layers of neurons are linked together	Nodes are partitioned then terminal nodes are found and class labels are allocated to these nodes.	The Hyperplane is built in high dimensional space, which is used for classification purpose then.	Various associations are combined into a set to decide traits of an image
Approach Used	Non parametric	Non parametric	Non parametric	Stochastic
Working Precision	Depends on network structure and number of inputs	Depends on hierarchical rule based method.	Depends on hyper plane selection and kernel parameter.	Depends on direction of decision
Overall work rate	Computation rate is high	Computation efficiency is good	Low result transparency sometimes	Depends on a-priori knowledge
Complex Problems	Noisy inputs can be handled with complexity.	When values are undecided computation is complex	Complexity is reduced to higher extent	A-Priori knowledge is required to reduce complexity and have good output
Training	Slow	Quick	Slow	Combining associations requires time and effort
Over Fitting Problem	Exists	Extensive design is not required	Doesn't exist	Uncertainty can be handled efficiently

Table 1: Comparison

Features	**KNN**	**ADABOOST**	**RF**	**ISODATA**	**DECORATE**
Algorithm	Supervised	Supervised	Supervised	Unsupervised	Semi-supervised
Working	Stores all available cases and classifies new cases based on a similarity measure	Works on a sequence of weak learners on different weighted training data.	Build multiple decision trees with different sample and different initial variables, working in an ensemble.	Uses Euclidean distance as the similarity measure to cluster data elements into different classes.	Builds diverse ensembles of classifiers by using specially constructed artificial training examples.
Approach Used	Non parametric	Non parametric	Non parametric	Stochastic	Uncertain
Working Precision	Based on minimum distance from the query instance to the training samples to determine the K nearest neighbors.	Handles lots of irrelevant features very well.	Handles lots of irrelevant features very well unless noise ratio is high	comprehensive set of heuristic procedures that uses the minimum spectral distance to form clusters	Can use any strong learner as a base classifier to build diverse committees.
Overall work rate	Depends on the value of K defined by Euclidean distance in feature space	Best with strong learners.	Depends on the number of trees in an ensemble to grow and the number of features to select randomly at each split.	Depends on number of clusters, and a number of additional user-supplied parameters.	Best when predictions of DECORATE are combined with base classifier.
Features	**KNN**	**ADABOOST**	**RF**	**ISODATA**	**DECORATE**
Algorithm	Supervised	Supervised	Supervised	Unsupervised	Semi-supervised
Working	Stores all available cases and classifies new cases based on a similarity measure	Works on a sequence of weak learners on different weighted training data.	Build multiple decision trees with different sample and different initial variables, working in an ensemble.	Uses Euclidean distance as the similarity measure to cluster data elements into different classes.	Builds diverse ensembles of classifiers by using specially constructed artificial training examples.
Approach Used	Non parametric	Non parametric	Non parametric	Stochastic	Uncertain
Working Precision	Based on minimum distance from the query instance to the training samples to determine the K nearest neighbors.	Handles lots of irrelevant features very well.	Handles lots of irrelevant features very well unless noise ratio is high	comprehensive set of heuristic procedures that uses the minimum spectral distance to form clusters	Can use any strong learner as a base classifier to build diverse committees.
Overall work rate	Depends on the value of K defined by Euclidean distance in feature space	Best with strong learners.	Depends on the number of trees in an ensemble to grow and the number of features to select randomly at each split.	Depends on number of clusters, and a number of additional user-supplied parameters.	Best when predictions of DECORATE are combined with base classifier.
Complex	Sensitive to noise	Depends on	Best algorithm and	Prior knowledge is	Reduces the

Problems	and irrelevant features	training data, complexity of weak classifiers and number of iterations.	can handle large data set with higher dimensionality.	required to reduce complexity and have good output	correlation between ensemble members and generalization errors. Accurate than boosting and bagging algorithms.
Training	Quick	Slow	Slow	Slow if data is unstructured	Quick
Over-Fitting Problem	Using fewer neighbo rs would actually lead to over fitting**.**	Ensemble methods, reduces the likelihood of over-fitting but prone to over-fitting when the training data contains high degree of noise.	Generally caused by over growing the trees.	Uncertain	Doesn't exist

Table 2: Comparison

VI. Related Work

This section presents the literature survey on various classification algorithms as following:

6.1 Support Vector Machines:

Support Vector Machines (SVMs) are investigated by *Ming-Hsuan Yang et. al.[13]* for visual gender classification with low-resolution "thumbnail" faces (21 -by-12 pixels) processed from 1,755 images from the FERET face database. The performance of SVMs (3.4% error) is shown to be superior to traditional pattern classifiers (Lineal; Quadratic, Fisher Linear Discriminant, Nearest-Neighbor) as well as more modern techniques such as Radial Basis Function (RBF) classifiers and large ensemble-RBF networks.

Amit Jain et.al.[33],presented a novel clustering based Short Term Load Forecasting (STLF) technique using Support Vector Machines (SVM). Author proposed that the forecasting is performed for the 48 half hourly loads of the next day. The daily average load of each day for all the training patterns and testing patterns is calculated and the patterns are clustered using a threshold value between the daily average load of the testing pattern and the daily average load of the training patterns. The data considered for forecasting contains 2 years of half hourly daily load and daily average temperature. Here clustering is used as pre-processing step while considering the patterns for training the SVM. Authors have presented the results for both the cases (without clustering and with clustering) and for different threshold values which result in forming different cluster patterns. The proposed method does not require any heavy computational burden and can be easily utilized for forecasting the next day load in energy utilities. Author has concluded that the applicability of clustering techniques for choosing training patterns for SVM based methods for getting better short term load forecasting results.

Ping Fu et.al. [40] proposed an image semantic classification algorithm based on feature subspaces. It is implemented by SVM and AdaBoost algorithm. In every feature subspace, a SVM is trained. According to the error rate of every SVM, the integrating weight of feature subspace is determined, with which different subspace features are concatenated into a feature vector. Later AdaBoost algorithm is employed to train the classifier; whose weak classifiers are SVMs. The experiment shows that this method increases the classifier precision effectively.

A multilayer feed forward, back-propagation ANN model and a SVM model for one hour ahead short term load forecasting on a small island power system of Trinidad and Tobago for three load types has been presented by*Glen Mitchell et. al.[86]*. These load types are batch, continuous and batch-continuous load types which represent three unique industrial customers. A performance comparison between the ANN and SVM showed that the SVM produced repeatability, always yielding the global minimum. Both ANN and SVM were unable to accurately perform load forecasts where there may be erratic load patterns or missing data, yielding deviations greater than 3%. For the continuously varying load, the ANN and SVM load forecasts yielded a maximum deviation of 1.20%.

In a paper by *Liliya Demidova et. al.[88]* a data classification technique, implying the consistent application of the SVM and kNN classifiers in the respective determined experimentally subareas of the characteristic space has been suggested. The recommendations on the use of the kNN classifier near the hyperplane dividing the classes have been provided. The examples of application of the offered technique for the problem of data classification have been given.Author proposed that this technique increases the classification quality, as the application of the kNN classifier to the objects located near the hyperplane dividing the classes and determined by the SVM classifier reduces the number of the mistakenly classified objects as well as this technique allows making the high-precision decisions on classification of the elaborate multidimensional data.

6.2 KNN:

*Pierre Héroux et.al.[8]*presents three classifiers used in automatic forms class identification. A first category of classifier includes the k-Nearest Neighbours (kNN) and the Multi-Layer Perceptron (MLP) classifiers and a second category corresponds to a new structural classifier based on tree comparison. kNN and the MLP classifiers use low level information based on a pyramidal decomposition of the document image whereas high level information represents the form content with a hierarchical structure used by the new structural classifier.

Nitin Bhatia et. al.[39] presented a survey of nearest neighbor techniques with its different variation as Weighted kNN, Model based kNN, Condensed NN, Reduced NN, Generalized NN are Structure Less Techniques whereas k-d tree, ball tree, Principal Axis Tree, Nearest Feature Line, Tunable NN, Orthogonal Search Tree are Structure Based Algorithms developed on the basis of kNN. The nearest neighbor (NN) technique is very simple, highly efficient and effective in the field of pattern recognition, text categorization, object recognition etc. Techniques both Structured Less and Structured Based are improvements over basic kNN technique with respect to speed efficiency as well as space efficiency.

6.3 CART:

Ramana V. Davuluri et. al.[14] prepared a non-redundant database of 2312 full-length human 5′-untranslated regions (UTRs) using state-of-the-art experimental and computational technologies. A comprehensive computational analysis of this data was conducted for characterizing the 5′ UTR features. Classification and regression tree (CART) analysis was used to classify the data into three distinct classes. Class I consists of mRNAs that are believed to be poorly translated with long 5′ UTRs filled with potential inhibitory features. Class II consists of terminal oligopyrimidine tract (TOP) mRNAs that are regulated in a growth-dependent manner, and class III consists of mRNAs with favorable 5′ UTR features that may help efficient translation. The most accurate tree we found has 92.5% classification accuracy as estimated by cross validation. The classification model included the presence of TOP, a secondary structure, 5′ UTR length, and the presence of upstream AUGs (uAUGs) as the most relevant variables. Here, the present classification and characterization of the 5′ UTRs provide precious information for better understanding the translational regulation of human mRNAs. Furthermore, this database and classification can help people build better computational models for predicting the 5′-terminal exon and separating the 5′ UTR from the coding region.

Pavuluri Manoj Kumar et. al.[17],focused on the increased rate of urbanization that has led to haphazard growth, increased infrastructure costs, deterioration of living conditions and degradation of the environment. In order to assist urban planners in more effective urban planning and management strategies, there is a great demand for developing accurate and detailed spatial information, understanding urban land cover changes, and the causes of these changes. This research studies the urban landscape dynamics for the city of Columbia, Missouri USA using multi-temporal and multi-date (1984, 92, and 2000) Landsat TM and ETM satellite imageries. The classification algorithms used to classify these images include traditional classifier (i.e. Maximum Likelihood) and rule-based classifier. Author has compared CART and ML classifications for urban land cover showed that ML performs better than CART in Urban Landscape Dynamics. CART can improve classification accuracy if there is a higher variance in the additional data used.

*Muhammad A. Razi et. al.[21]*have performeda three-way comparison of prediction accuracy involving nonlinear regression, NNs and CART models using a continuous dependent variable and a set of dichotomous and categorical predictor variables. A large dataset on smokers is used to run these models and different prediction accuracy measuring procedures are used to compare performances of these models. Authorsconcluded that NNs and CART models produced better prediction accuracy than non-linear regression model.

*Yang Shao et. al. [56]*have **c**ompared support vector machine with two non parametric image classification algorithms: neural network, and CART algorithms for the land-cover classification using limited training data points. SVM generated overall accuracies ranging from 77% to 80% for training sample sizes from 20 to 800 pixels per class, compared to 67–76% and 62–73% for NN and CART, respectively for 2001 MODIS time-series data. In this paper, classification experiments were conducted with respect to the impact of training sample sizes, training sample variations, and the characteristics of reference data points. This paper concluded that SVM achieved higher overall accuracy and significantly improved Kappa coefficients for the entire range of training sample sizes compared to the NN and the CART algorithms.

6.4 ADABOOST:

Ryuei Nishiiet. al.[22] elaborated the AdaBoost, a machine learning technique for supervised classification of land-cover categories of geo statistical data. Spatial AdaBoost is introduced in order to provide contextual image classification. Authors have defined the classifiers based on the log posterior probabilities on the neighborhoods, and combine the classifiers by the AdaBoost-based method. Authors also proposed a method that can be applied to artificial multispectral images and benchmark datasets and the performance of the proposed method is excellent and is similar to the Markov-random-field-based classifier, which requires an iterative maximization procedure.

Ruihu Wang[54] focussed on AdaBoost algorithm for feature selection, classifier learning and its relation with SVM with brief introduction to the AdaBoost which is used for producing a strong classifier out of weak learners firstly. The original adaptive boosting algorithm and its application in face detection and facial expression recognition are also reviewed.

CAO Ying et. al.[59] elaborated AdaBoost as one of the most excellent Boosting algorithms. It has a solid theoretical basis and has made great success in practical applications. AdaBoost can boost a weak learning algorithm with accuracy slightly better than random guessing into an arbitrarily accurate strong learning algorithm, bringing about a new method and a new design idea to the design of learning algorithm. This paper first introduces how Boosting, just a conjecture when proposed, was proved right, and how this proof led to the origin of AdaBoost algorithm. Second, training and generalization error of AdaBoost are analyzed to explain why AdaBoost can successfully improve the accuracy of a weak learning algorithm. Third, different theoretical models to analyze AdaBoost are given. Many variants derived from these models are presented. Fourth, extensions of binary-class AdaBoost to multiclass AdaBoost are described. Applications of AdaBoost and interested directions which need to be further studied are discussed.

Xiukuan Zhaoa et. al.[62] used the AdaBoost-BP algorithm to construct a new model to predict the critical frequency of the ionospheric F2-layer (foF2) one hour ahead i.e. in short time scale typically 1hr to 24 hrs in advance. Different indices were used to characterize ionospheric diurnal and seasonal variations and their dependence on solar and geomagnetic activity. These indices, together with the current observed foF2 value, were input into the prediction model and the foF2 value at one hour ahead was output. In this paper, twenty-two years' foF2 data from nine ionosonde stations in the East-Asian sector is analyzed, the first eleven years' data were used as a training dataset and the second eleven years' data were used as a testing dataset. The results show that the performance of AdaBoost-BP is better than those of BP Neural Network (BPNN), Support Vector Regression (SVR) and the IRI model.

6.5 K-Means:

Balasubramanian Subbiah et. al. [51] haveproposed two clustering algorithm combinations with integration of K-Means algorithm that can tackle some of the clustering problems. Clustering analysis is a valuable and useful tool for image classification and object diagnosis. A variety of clustering algorithms are available and still this is a topic of interest in the image processing field. But clustering algorithms are confronted with difficulties in meeting the optimum quality requirements, automation and robustness requirements. in this paper, the image segmentation and classification done using K-Means and Laplacian of Gaussian filters or Prewitt Filter. The proposed algorithm has no specific requirement of prior knowledge of any parameters and the mathematical details of the data sets. The integrated novel clustering algorithms forimage classification are tested with different imagesincluding satellite images.

Haval M. SIDQI et. al. [64] investigated the required image operation to extract the features. K-means algorithm is a widely used partition method in clustering techniques. As the dataset's scale increases rapidly, it is difficult to use K-means to deal with massive amount of data. A parallel strategy that incorporated into clustering method and a K-mean algorithm are proposed. For enhancing the efficiency of K-mean, dynamic load balance is also discussed. The K-means method has been shown to be effective in producing good clustering results for many practical applications. The proposed algorithm is applied on a famous 'Lena' gray scale image(256x256). This test is dependent on the number of clusters and they require calculation time. The performance parameter that used to evaluate our algorithm is based on fidelity measure PSNR.

*Chuen-Horng Lin et. al. [69]*stated that image color featureis the most commonly used image feature with a K-means algorithm. Authors have proposed a fast K-Means algorithm for image retrieval. For this purpose, first a level histogram of statistics for the image database is made. The level histogram is used with the K-means algorithm for clustering data. A fast K-means algorithm not only shortens the length of time spent on training the image database cluster centers, but it also overcomes the cluster center re-training problem since large number of images are continuously added into the database. Authors have used gray and color image database sets for performance comparisons and analysis. After analysis, it was concluded that the fast K-means algorithm is more effective, faster, and more convenient than the traditional K-means algorithm and it also overcomes the problem of spending excessive amount of time on re-training caused by the continuous addition of images to the image database.

*Marco Capóa et.al. [80]*authors have proposed an efficient approximation to the K-means problem intended for massive data because there is progressive growth of the amount of data available in a wide variety of scientific fields and it has become more difficult to manipulate and analyze such information. The proposed approach, recursively partitions the entire dataset into a small number of subsets, each of which is characterized by its representative (center of mass) and weight (cardinality), afterwards a weighted version of the K-means algorithm is applied over such local representation, which can drastically reduce the number of distances computed. This approach is called recursive partition based K-Means (RKPM). This paper concluded that RPKM algorithm generates competitive approximations, even at its earlier iterations, while reducing several orders of magnitude of distance computations.

6.6 Naïve Bayes:

A novel algorithm, Self-adaptive NBTree, which induces a hybrid of decision tree and Naive Bayes is proposedby*Li-Min Wang et. al. [23]*. Decision tree is useful to obtain a proper set of rules from a large number of instances but it has difficulty in obtaining the relationship between continuous-valued data points. Therefore, the Bayes measure is used to construct decision tree that can directly handle continuous attributes and automatically find the most appropriate boundaries for discretization and the number of intervals. The Naive Bayes node helps to solve overgeneralization and overspecialization problems which are often seen in decision tree. Authors proposed a hybrid approach, Selfadaptive NBTree, which applies post-discretization strategy to mitigate the negative effect caused by information loss. At the same time, it embodies tradeoff between the accuracy and the complexity of the learned discretization by applying MDL principle.

*Yuguang Huang et. al. [49]*concluded thatNaive Bayes algorithm is one of the most effective methods in the field of text classification, but it can get the more accurate resultonly in a large training sample set. Large training samples require heavy work for manual classification and put forward a higher request for storage and computing resources during the computer post-processing. This paper mainly studies Naïve Bayes classification algorithm based on the small sample data set. By introducing Poisson probability model, each document is regarded as Poisson random variable generated by the multivariate Poisson model, and authors conducted a series of comparative experiments in the Chinese data set using the combination method of Poisson distribution model and Naïve Bayes. The experimental results show that the method not only in the large-scale data set has satisfactory classification results, but also shows a good classification performance in the small sample data set.

*Hugo Jair Escalantea et. al. [74]*described a new methodology for early gesture recognition based on the well known Naive Bayes classifier. The method is extremely simple and very fast, yet it compares favorably with more elaborated state of the art methodologies. The naive baseline is based on three main observations: (1) the effectiveness of the naive Bayes classifier in text mining problems; (2) the link between natural language processing and computer vision via the bag-of-words representation; and (3) the cumulative-evidence nature of the inference process of naive Bayes. Authors evaluated the proposed method in several collections that included segmented and continuous videos. The proposed method takes advantage of Naive Bayes cumulative-evidence property and adapts it to gesture recognition and spotting. In this paper, a comparison with state of the art methods in standard data sets is performed and their result shows that Early Naive Bayes compare favorably with a number of more complex methodologies that have been specifically designed for early gesture recognition.

Artem A. Maksutov et. al. [85] elaborated the Bayesian network (BN), which is the probabilistic model organized in acyclic graph. Artificial neural networks have occupied significant popularity in IT world but suffer from the problem of retraining, which is the real problem for industrial usage of neural networks. Bayesian network classifies the investigated objects and show that this network takes into account the noise and doesn’t suffer from retraining.BN approach shows great results, but still has some problems of application because of huge amount of calculations during the process of learning.

6.7 Decision Trees:

J. R. Quinlant[2] investigated methods for simplifying decision trees without compromising their accuracy with a desire to turn decision trees into knowledge for use in expert systems. Author has discussed four methods: Cost-Complexity Pruning, Reduced Error Pruning, Pessimistic Pruning, and Simplifying to Production Rules, all of which managed to achieve significant simplification when put to test on sets of decision trees from six task domains. This simplification was often coupled with an actual improvement in classification accuracy on unseen cases.

Standard Decision Tree algorithms used for Land Cover and Land Use mapping were evaluated and compared using satellite data by *Dr. P.K. Srimani et. al. [53]*. Classification rules were derived from the spectral image using J48, BFTree, REPTree and Simple Cart algorithms with the same set of training samples. Classification done by using these rules is known as knowledge based classification and the results of these classifiers were compared and evaluated based on True Positive, False Positive, Prediction Accuracy and Learning Time metrics. Among these, J48 performed the best in all aspects and had a Prediction accuracy of 97.34% and Kappa statistics of 0.9685. Further the J48 Decision Tree classified image algorithm based rules has produced an overall accuracy of 87.11% and kappa of 0.8515.

*Lior Rokach, Oded Maimon[89]*presented an updated survey of current methods for constructing decision tree classifiers in a top-down manner using growing and pruning. Author has suggested a unified algorithmic framework for presenting the various algorithms and describes various splitting criteria and pruning methodologies. This paper discusses that splitting criteria categorized in two ways: Univariate Criteria as Impurity-based Criteria, Normalized Impurity Based Criteria, Binary Criteria, Twoing Criterion etc. and Multivariate Criteria. In this paper, Pruning methods as Cost–Complexity Pruning, Reduced Error Pruning, Minimum Error Pruning (MEP), Pessimistic Pruning etc along with their comparison is also discussed. Decision tree algorithms as ID3, C4.5, CART, CHAID and QUEST are also given consideration for research.

6.8 ISODATA:

A novel method of data analysis and pattern classification called ISODATA was described by*GEOFFREY H. BALL et. al. [1]*. The area of research labeled "pattern recognition" consists primarily of efforts to develop techniques capable of dealing with problems of inherently high dimension. Many aspects of the pattern recognition problem are data analysis so this paper proposes that ISODATA fits this description of data analysis and so to pattern recognition as well. The proposed technique has demonstrated its ability to lay bare the structure of the data. ISODATA can be used to evaluate preprocessing by comparing the clustering before the preprocessing with the clustering after the preprocessing. This makes it possible to evaluate the preprocessor with respect to the inherent difficulty of the data.

M. Merzougui et. al. [60] elaborated that the unsupervised classification by ISODATA algorithm has difficulty in adjusting its parameters. These control its convergence. Its result depends strongly on two parameters: distance threshold for the union of clusters and threshold of typical deviation for the division of a cluster. The wrong choice of these parameters may cause the algorithm to spiral out of control leaving at the end of one cluster. Authors have proposed a new approach to overcome this difficulty of the ISODATA algorithm. An evolutionary algorithm is adapted to estimate the two optimal thresholds to be used by the algorithm ISODATA. This approach is validated on simulation examples. The experimental results confirm the favorable convergence speed and good performance of the proposed algorithm. Moreover, authors proposed a new mutation operator to allow the algorithm firstly to avoid local solutions and also to converge to the global solution in a small number of generations.

*Qian Wang et. al. [67]*proposed an improved ISODATA algorithm for hyperspectral images classification. The algorithm takes the maximum and minimum spectrum of the image into consideration and determines the initial cluster center by the stepped construction of spectrum accurately. The classification results show that the improved ISODATA algorithm can determine the initial cluster number adaptively. In comparison to the SAM (Spectral Angle Mapper) algorithm and the original ISODATA algorithm, the improved algorithm have an advantage in finding the subtle difference among different classes, which can be put into wider use in the classification process.

*Sahar A. El_Rahman[70]*developed an unsupervised hyperspectral image classification algorithm, in particular, Iterative Self-Organizing Data Analysis Technique Algorithm (ISODATA) algorithm was used to produce a classified image and extract agricultural information, using ENVI (Environment of Visualizing Images) that is a software application utilized to process and analyze geospatial imagery. Hyperspectral data consists of many bands - up to hundreds of bands - that cover the electromagnetic spectrum. This results in a hyperspectral data cube that contains approximately hundreds of bands - which means BIG DATA CHALLENGE. This study has been applied in Florida, USA. Hyperspectral dataset of Florida was generated by the SAMSON sensor. In this paper, the performance was evaluated on the basis of the accuracy assessment of the process after applying Principle Component Analysis (PCA) and ISODATA algorithm. The overall accuracy of the classification process is 75.6187%. In this paper, the analysis of ISODATA algorithm has been performed to classify pixels. Principle Component Analysis (PCA) is used before the classification process as a technique in data analysis to reduce hyperspectral image dimensions. From the classified image data, Statistical information is calculated.

6.9 Random Forest:

*Leo Breiman's [15]*work states that Random forests are a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. The generalization error for forests converges to a limit as the number of trees in the forest becomes large. The generalization error of a forest of tree classifiers depends on the strength of the individual trees in the forest and the correlation between them. This paper gives detailed description of Random Forest algorithm followed by some theoretical background for random forests and also introduces forests using the random selection of features at each node to determine the split. Here, classifier strength and dependence are computed that works as internal estimates of the generalization error that helps in deciding that how many features to select at each node. This paper shows results for two different forms of random features, first uses random selection from the original inputs and second uses random linear combinations of inputs. Then results are compared to Adaboost which turn out to be insensitive to the number of features selected to split each node.

Vrushali Y Kulkarni et. al.[66] paper explains that the Random Forest is a supervised machine learning algorithm. In Data Mining domain, machine learning algorithms are extensively used to analyze data, and generate predictions based on this data. Being an ensemble algorithm, Random Forest generates multiple decision trees as base classifiers and applies majority voting to combine the outcomes of the base trees. Strength of individual decision trees and correlation among the base trees are key issues which decide generalization error of Random Forest classifiers. Based on accuracy measure, Random Forest classifiers are at par with existing ensemble techniques like bagging and boosting. This paper made an attempt to improve performance of Random Forest classifiers in terms of accuracy, and time required for learning and classification. To achieve this, five new approaches are proposed, proposed approaches are Disjoint Partitioning Approach, Weighted Hybrid decision tree model (WHDT), Optimal Subset of Random Forest, Diversity Based Dynamic Pruning (DBDP) and Parallel Random Forest (PDTPRF). The empirical analysis and outcomes of experiments carried out in this research work lead to effective learning and classification using Random Forest algorithm.

Gerard Biau and Erwan Scornet's [77] work gives detailed description on Random forest classifiers. Authors state that the random forest algorithm, proposed by L. Breiman in 2001, has been extremely successful as a general purpose classification and regression method. This approach combines several randomized decision trees and aggregates their predictions by averaging with excellent performance in settings where the number of variables is much larger than the number of observations. Moreover, it is versatile enough to be applied to large-scale problems, is easily adapted to various ad-hoc learning tasks, and returns measures of variable importance. This article reviews the most recent theoretical and methodological developments for random forests. Emphasis is placed on the mathematical forces driving the algorithm, with special attention given to the selection of parameters, the resampling mechanism, and variable importance measures.

Vishal .T .V et.al. [81] have done a case study of a number of image classification algorithms which include decision trees, k-nearest neighbors, deep neural networks, Convolution neural networks, Support vector machines and random forest. This paper elaborates three main algorithms – SVM, RF and CNN with their varying approaches to classification. They have been juxtaposed with their applications in various domains which have an influence on technology. This paper also discusses the multiple advantages and disadvantages of the aforementioned algorithms.

Random Forest classifier has been considered as an important reference in the data mining area *Joaquin Abellan et. al. [87]*. The building procedure of its base classifier (a decision tree) is principally based on a randomization process of data and features; and on a split criterion, which uses classic precise probabilities, to quantify the gain of information. One drawback found on this classifier is that it has a bad performance when it is applied on data sets with class noise. Very recently, it is proved that a new criterion which uses imprecise probabilities and general uncertainty measures, can improve the performance of the classic split criteria. In this work, the base classifier of the Random Forest is modified using that new criterion, producing also a new single decision tree model. This model is the base classifier that joins the features extracted from the randomization process with a new procedure similar to the Random Forest, called Credal Random Forest. The principal difference between those two models is presented in this paper. Experiments are performed to show that the new method represents an improvement on the Random Forest when both are applied on data sets without class noise. But this improvement is notably greater when they are applied on data sets with class noise.

6.10 DECORATE(Diverse Ensemble Creation by Oppositional Relabeling of Artificial Training Examples)

Prem Melville et. al. [18] state that the Ensemble methods like Bagging and Boosting which combine the decisions of multiple hypotheses are some of the strongest existing machine learning methods. The diversity of the members of an ensemble is known to be an important factor in determining its generalization error. In this paper, a new method for generating ensembles, DECORATE (Diverse Ensemble Creation by Oppositional Relabeling of Artificial Training Examples) is presented. This method directly constructs diverse hypotheses using additional artificially-constructed training examples. The technique is a simple, general meta-learner that can use any strong learner as a base classifier to build diverse committees. This method uses decision-tree induction as a base learner demonstrates that this approach consistently achieves higher predictive accuracy than both the base classifier and Bagging. DECORATE also obtains higher accuracy than Boosting early in the learning curve when training data is limited. Authors proposed to show that DECORATE can also be effectively used (1) active learning, (2) semi-supervised learning, (3) combining active learning with semi-supervision, (4) regression, (5) improving class membership probability estimates and (6) relational learning.

Harith Al-Sahaf et.al.[52] have proposed two genetic programming (GP) strategies, one-shot GP and compound-GP, that plan to develop a program for the errand of paired arrangement in pictures. The two strategies are intended to utilize just a single or a few instances per class to evolve the model. In this study, authors research these two methods in terms of performance, robustness, and complexity of the evolved programs. Authors have used ten data sets that vary in difficulty to find these two strategies. Authors have also compared them with two other GP and six non-GP strategies. The outcomes demonstrate that one-shot GP and compound-GP beat or almost practically identical to contender strategies. Besides, the components extricated by these two strategies enhance the execution of different classifiers with high quality elements and those separated by recently created GP-based strategy by and large.

Bo Sun et. al.'s [68] work analyzes the effectiveness of DECORATE an ensemble learning algorithm. DECORATE algorithm is generated by augmenting the original training set using artificial training examples and has exhibited both stronger robustness to noise than AdaBoost whereas algorithms as bagging and ADABOOST are generated by manipulating the original training set and DECORATE has better resilience to missing features than Bagging and AdaBoost. To better understand the effectiveness of DECORATE; a study has already been conducted from the perspective of the bias–variance theory. However, this theory fails in explaining the effectiveness of AdaBoost which constantly improves the generalization performance by adding more base classifiers but it increases the variance. The margin theory is another important means that can be used to explain the effectiveness of ensemble learning. In this paper, authors have empirically analyzed the effectiveness of DECORATE by conducting experiments on 15 standard data sets and found that the margin theory can also well explain the effectiveness of DECORATE.

VII. Conclusion And Future Scope

Real Life image classification is an interesting research area and it has made great progress in last decades. This paper has provided the detailed review of image classification techniques for real world image classification. Comparison of various image classification techniques as ANN, Decision Trees, SVM, Random Forest, K-Means, KNN, ISODATA, ADABOOST, and Decorate has been presented in this paper. This comparison is conducted with respect to various parameters as algorithm category, training, complexities, working precision etc. Each algorithm has its own strengths and weaknesses. Several researches had been conducted in order to use multiple remote-sensing features, including spectral, spatial, multi temporal, and multi sensor information in different fields. This survey gives theoretical knowledge about different classification methods and provides the advantages and disadvantages of various classification methods.

For the future work, the improvement of image classification techniques will increase accuracy value and subsequently feasible to be employed for computer-aided-classification, and more robust methods are being developed.

References

[1] Geeffrey H. Ba,David j. Hall*"Isodata, A Novel Method Of Data Analysis And Pattern Classification"*, Information Sciences Branch Office of Naval Research Contract, April 1965

[2] J. R. Quinlant, "*Simplifying decision trees*",. Int. J. Man-Machine Studies 1987.

[3] D.R. Hush , "*Classification With Neural Networks: A Performance Analysis*", IEEE 1989.

[4] Etienne Barnard and David Casasent, "*ImageProcessing For Image Understanding With Neural Nets*", IEEE 1989.

[5] John R. Smith and Shih-Fu Chang, "*Transform Features For Texture Classification And Discrimination In Large Image Databases"*, IEEE 1994.

[6] Anil Jain and Douglas Zongker, "*Feature Selection: Evaluation, Application, and Small Sample Performance*", Ieee Transactions On Pattern Analysis And Machine Intelligence, Vol. 19, No. 2, February 1997

[7] Massimiliano Pontil and Alessandro Verri , "*Support Vector Machines for 3D Object Recognition*", Ieee Transactions On Pattern Analysis And Machine Intelligence, June 1998.

[8] Pierre Héroux, Sébastien Diana1, Arnaud Ribert, Eric Trupin, "*Classification Method Study for Automatic Form Class Identification*", 14th International Conference of Pattern Recognition, Aug 1998.

[9] V.N. Vapnik, "*Statistical Learning Theory"*, Wiley-Interscience Pub., 1998.

[10] Trygve Randen, Member, John Håkon Husøy, *"Filtering for Texture Classification: A Comparative Study*", Ieee Transactions On Pattern Analysis And Machine Intelligence, Vol. 21, No. 4, April 1999.

[11] Anil K. Jain, Robert P.W. Duin, Jianchang Mao, "*Statistical Pattern Recognition: A Review*", Ieee Transactions On Pattern Analysis And Machine Intelligence, Vol. 22, No. 1, January 2000

[12] Aleksander KoÃ lcz, "*N-tuple Network, CART, and Bagging*", Neural Computation Volume: 12, Issue: 2, Feb. 1, 2000

[13] Ming-Hsuan Yang , Baback Moghaddam, "*Support Vector Machines for Visual Gender Classification*", IEEE 2000. CART

[14] Ramana V. Davuluri, Yutaka Suzuki, Sumio Sugano, Michael Q. Zhang, *"Classification of Human 5' UTR Sequences*", November 8, 2000, Genome Res. 2000. Pg 1807-1816.

[15] Leo Breiman, "*Random Forests*", Statistics Department University of California Berkeley, CA 947, 20 January 2001.

[16] R.L. Lawrence and A. Wright, "*Rule-Based Classification Systems Using Classification and Regression Tree (CART) Analysis*". Photogrammetric Engineering and Remote Sensing, vol. 67, pp. 1137-1142, 2001.

[17] Pavuluri ManojKumar, Ramanathan Sugumaran, and Daniel Zerr, "*A Rule-Based Classifier Using Classification and Regression Tree (CART) Approach For Urban Landscape Dynamics*", IGARSS, 2002.

[18] Prem Melville, R. J. Mooney, "*Constructing Diverse Classifier Ensembles Using Artificial Training Examples*". In: Eighteenth International Joint Conference on Artificial Intelligence, 505-510, 2003

[19] Yang, C. C., Prasher, S. O. , Enright, P., Madramootoo, C., Burgess, C., Goel, P. K., Callum, I., *Application of decision tree technology for image classification using remote sensing data,* in Agricultural Systems 76, pp. 1101-1117, 2003

[20] Khan, S. S. and Ahmed, "*A.: Cluster center initialization for Kmeans algorithm*". Pattern Recognition Letters, vol. 25, no. 11, pp. 1293-1302, 2004.

[21] Muhammad A. Razi, Kuriakose Athappilly, "*A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models"*, Expert Systems with Applications 29 (2005) 65–74.

[22] Ryuei Nishii, *Member,* Shinto Eguchi, *"Supervised Image Classification by Contextual AdaBoost Based on Posteriors in Neighborhoods",*IEEE Transactions On Geoscience And Remote Sensing, Vol. 43, No. 11, November 2005.

[23] Li-Min Wang , Xiao-Lin Li, Chun-Hong Cao, Sen-Miao Yuan, "*Combining decision tree and Naive Bayes for classification*", Knowledge-Based Systems 19 (2006) 511–515

[24] Rich Caruana, Alexandru Niculescu-Mizil "*An Empirical Comparison of Supervised Learning Algorithms*" 23 rd International Conference on Machine Learning, Pittsburgh, PA, 2006.

[25] D. Lu, Q. Weng, "*A survey of image classification methods and techniques for improving classification performance*",pages 823-870 | 2007

[26] A. P. Dhawan and S. Dai, "*Clustering and Pattern Classification*", in Principles And Advanced Methods in Medical Imaging And Image Analysis. Singapore: World Scientific Publishing Co. Pte. Ltd, 2008, pp 229-265.

[27] A. P. Dhawan, "*Image Segmentation and Feature Extraction*", in Principles And Advanced Methods in Medical Imaging And Image Analysis. Singapore: World Scientific Publishing Co. Pte. Ltd, 2008, pp 197-228.

[28] A. Palaiahnakote Shivakumara, b Deepu Rajan and b Suresh A. Sadananthan, "*Image Classification: Are Rule-based systems effective when classes are fixed and known?*", 19th International Conference on Pattern Recognition, 2008.

[29] B. Krishnamurthya, T. Malikb , S. Stamatisa, V. Venkatasubramaniana, J. Caruthersa, "*Rule-based Classification Systems for Informatics*", Fourth IEEE International Conference on eScience, Dec 2008.

[30] E. Ibrahim, S. Adam, J. Monbaliu, *"Assessment of unsupervised classification techniques for intertidal Sediments"*, 28th EARSeL Symposium and Workshops, Remote Sensing for a Changing Europe, *Istanbul, 348:355, 2008.*

[31] C. Chandra Mouli, P. Jyothi , K. Nagabhushan Raju, "*Comparative Study of Supervised Classification Algorithms for WOSRAS*",International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering.\Vol. 3, Issue 2, February 2014

[32] Lei Shi," *Image Semantic Classification algorithm Research On Kernel PCA support vector machine*", 978-1-4244-3531-9/08@2008 IEEE.

[33] Amit Jain, B. Satish, "*Clustering based Short Term Load Forecasting using Support Vector Machines*", IEEE 2009.

[34] M. Janaki Meena , K. R. Chandran, "*Naive Bayes Text Classification with Positive Features Selected by Statistical Method*", International Conference on Cloud Computing and Intelligence System, 2011.

[35] Manik Varma and Andrew Zisserman, "*A Statistical Approach to Material Classification Using Image Patch Exemplars*", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 31, No. 11, November 2009

[36] Xiaojuan Ban," *Color Image Retrieval and Classification Using Fuzzy Similarity Measure and Fuzzy Clustering Method*", Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference 978-1-4244-3872-3/09©2009 IEEE

[37] A.P. Dhawan, "*Image Representation, Analysis and Classification-Medical Image Analysis"* 2nd Edition, Wiley Online Library Document, 2010.

[38] Dr. M.Seetha," *Accuracy Assessment of Object Oriented and Knowledge Base Image Classification using P-Trees*", 978-1-4244-5586-7/10@2010 IEEE

[39] Nitin Bhatia and Vandana, "*Survey of Nearest Neighbor Techniques*," IJCSIS) Int. J. Comput. Sci. Inf. Secur., vol. 8, no. 2, pp. 302–305, 2010.

[40] Ping Fu,Mianshu Chen,Yuan Sun,Chen Shen, "*Image Semantic Classification Based on Integration of Feature Subspaces*",2010 International Conference on Communications and Mobile Computing",IEEE,DOI 10.1109/CMC.2010.19,2010.

[41] Shuang Li, Zhilin Li,*" Effects of Image Fusion Algorithms on Classification Accuracy*", 18th International Conference on Geoinformatics, 2010

[42] Xianjun Zhang," *A Novel Image Classification Method Based on Manifold Learning and Gaussian Mixture Model*", 978-1-4244-5555-3/10©2010 IEEE

[43] Rama Rao Nidamanuri," *Normalized Spectral Similarity Score (NS3) as an Efficient Spectral Library Searching Method for Hyperspectral Image Classification"*, IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing 1939-1404, 2010

[44] Zhi Zeng," *Similarity-Based Image Classification Via Kernelized Sparse Representation*", Proceedings of 2010 IEEE 17th International Conference on Image Processing 978-1-4244-7994-8/10©2010 IEEE

[45] Wei-Yin Loh, "*Classification and regression trees*", John Wiley & Sons, Inc. Volume 1, January/February 2011

[46] S.Manthira Moorthi," *Kernel based learning approach for satellite image classification using support vector machin*e", 978-1-4244-9477-4/11, 2011 IEEE

[47] Liu Jinmei," *A Refined Quadtree-based Automatic Classification Method for Remote Sensing Image*", 2011 International Conference on Computer Science and Network Technology 978-1-4577-1587-7/11©201l IEEE

[48] Piyush Rai "*Supervised Learning: K-Nearest Neighbors and Decision Trees*", August 25, 2011

[49] Yuguang Huang, Lei Li , "*NAIVE BAYES CLASSIFICATION ALGORITHM BASED ON SMALL SAMPLE SET*", Proceedings of IEEE CCIS 2011.

[50] A. L. Choodarathnakara, Dr. T. Ashok Kumar, Dr. Shivaprakash Koliwad, Dr. C. G. Patil, "*A Challenge in Remote Sensing Data Classification for improving performance*", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) , November 2012

[51] Balasubramanian Subbiah and Seldev Christopher. C, *"Image Classification through integrated K- Means Algorithm*", International Journal of Computer Science Issues, Vol. 9, Issue 2, No 2, March 2012

[52] Harith Al-Sahaf," *Extracting Image Features for Classification By Two-Tier Genetic Programming*", WCCI 2012 IEEE World Congress on Computational Intelligence.

[53] Dr. P.K. Srimani and Smt. Nanditha Prasad, "*Decision Tree Classification Model For Land Use And Land Cover Mapping-A Case Study*", International Journal of Current Research Vol. 4, Issue, 05, pp.177-181, May, 2012

[54] Ruihu Wang, "*AdaBoost for Feature Selection, Classification and Its Relation with SVM-, A Review*", International Conference on Solid State Devices and Materials Science, 2012.

[55] Zhiyong Xiong," *An Algorithm of Image Classification Based on BP Neural Network*", 978-1-4673-0089-6/12 ©2012 IEEE

[56] Yang Shao, Ross S. Lunetta, " *Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points*", ISPRS Journal of Photogrammetry and Remote Sensing 70 (2012) 78–87.

[57] Zilei Wang," *Linear Distance Coding for Image Classification*", IEEE Transactions On Image Processing 1057–7149© 2012 IEEE.

[58] Bo Yu , Li Whang, Zheng Niu, "*A novel algorithm in buildings/shadow detection based on Harris detector*" Science Direct, 2013.

[59] CAO Ying, MIAO Qi-Guang, LIU Jia-Chen, GAO Lin, "*Advance and Prospects of AdaBoost Algorithm*", ACTA AUTOMATICA SINICA, vol 6 June 2013.

[60] M. Merzougui, M. Nasri, B. Bouali, "*ISODATA classification with parameters estimated by evolutionary approach*", 2013 IEEE.

[61] Pooja Kamavisdar, Sonam Saluja, Sonu Agrawal, "*A Survey on Image Classification Approaches and Techniques*", International Journal of Advanced Research in Computer and Communication EngineeringVol. 2, Issue 1, January 2013

[62] Xiukuan Zhaoa, Baiqi Ninga, Libo Liua, Gangbing Songc, "*A prediction model of short-term ionospheric foF2 based on AdaBoost*", Advances in Space Research (2013)

[63] Abass Olaode, Golshah Naghdy & Catherine Todd, *"Unsupervised Classification of Images: A Review*", International Journal of Image Processing (IJIP), Volume (8), Issue (5) : 2014

[64] Haval M. SIDQI and Jamal F. KAKBRA, "*Image Classification Using K-mean Algorithm*", IJETTCS, Volume 3, Issue 1, January – February 2014.

[65] Claudio Persello, Lorenzo Bruzzone*, "Active and Semisupervised Learning for the Classification of Remote Sensing Images*", Ieee Transactions On Geoscience And Remote Sensing, Vol. 52, No. 11, November 2014.

[66] Vrushali Y Kulkarni, Pradeep K Sinha, "*Effective Learning and Classification using Random Forest Algorithm*", International Journal of Engineering and Innovative Technology (IJEIT) Volume 3, Issue 11, May 2014.

[67] Qian Wang, Qingli Li, Hongying Liu, Yiting Wang, Jianzhong Zhu, "*An Improved ISODATA Algorithm for Hyperspectral Image Classification* ", 7th International Congress on Image and Signal Processing, 2014.

[68] Bo Sun, Haiyan Chen,Jiandong Wang, "*An empirical margin explanation for the effectiveness of DECORATEensemble learning algorithm", Knowledge-Based Systems"*, 29 January 2015

[69] Chuen-Horng Lin, Chun-Chieh Chen, Hsin-Lun Lee, Jan-Ray Liao, "*Fast K-means algorithm based on a level histogram for image retrieval*", Expert Systems with Applications 41 (2014) 3276–3283

[70] Sahar A. El_Rahman, "*Hyperspectral Imaging Classification Using ISODATA Algorithm: Big Data Challenge*", 2015 Fifth International Conference on e-Learning.

[71] Thamilselvan P, Dr. J. G. R. Sathiaseelan, "*Image Classification using Hybrid Data Mining Algorithms – A Review"*, IEEE Sponsored 2nd International Conference on Innovations in Information Embedded and Communication Systems ICIIECS'15

[72] Junxia Li, Deepu Rajan, Jian Yang, "*Local Feature Embedding For Supervised Image Classification*", IEEE, 2015

[73] Delin Chu, Li-Zhi Liao, Michael Kwok-Po Ng, and Xiaoyan Wang, "*IncrementalLinear Discriminant Analysis:A Fast Algorithm and Comparison",*,IEEE Transactions On Neural Networks And Learning Systems, 2015

[74] Hugo Jair Escalante, Eduardo F. Morales, L.Enrique Sucar , *"A Naïve ve Bayes baseline for early gesture recognition",* Pattern Recognition Letters, January 2016.

[75] Avinash Kak, "*AdaBoost for Learning Binary and Multiclass Discriminations"*, Purdue University, December 12, 2016.

[76] Sourav De, Siddhartha Bhattacharyya, Susanta Chakraborty, Paramartha Dutta "*Image Segmentation: A Review -Hybrid Soft Computing for Multilevel Image and Data Segmentation",* Computational Intelligence Methods and Applications", Springer International Publishing AG 2016.

[77] Gerard Biau, Erwan Scornet, *"A Random Forest Guided Tour*", An Official Journal of the Spanish Society of Statistics and Operations Research, 2016

[78] Shuprajhaa T, Subasree S, Vaitheeshwari M, Sivakumar S *"A review on Image Processing techniques using Pattern matching in LabVIEW"* International Journal of Advanced Engineering Research and Applications, Volume – 1, Issue – 11, March – 2016

[79] Harith Al-Sahaf,Mengjie Zhang,Mark Johnston,"*Binary Image Classification: A Genetic Programming Approach to the Problem of Limited Training Instances*",Massachusetts Institute of Technology,2016.

[80] Marco Capóa, Aritz Pérez, Jose A. Lozano, "An *efficient approximation to the K-means clustering for massive data*", Knowledge-Based Systems 000 (2016) 1–14

[81] Vishal .T .V, Srinidhi .S,Srividhya .S,Sri Vishnu Kumar .K,Swathika R,"*A Survey and Comparison of Artificial Intelligence Techniques for Image Classification and Their Applications*",International Journal of Science and Research (IJSR) ,Volume 5 Issue 4, April 2016.

[82] Han Liu, Alexander Gegov, "*Rule Based Systems and Networks: Deterministic and Fuzzy Approaches*", IEEE 8th International Conference on Intelligent Systems, 2016

[83] Zhijiang Lou, Jianyong Tuo, Youqing Wang "*Two-step Principal Component Analysis for Dynamic Processes"*, AdCONIP May 28-31, 2017. Taipei, Taiwan

[84] Panos P. Markopoulos, "*Linear Discriminant Analysis With Few Training Data*" IEEE 2017

[85] Artem A. Maksuto, Anastasiya V. Simonenko, Igor S. Shmakov, "*Classifiers Based on Bayesian Neural Networks*", IEEE 2017.

[86] Glen Mitchell, Sanjay Bahadoorsingh, Neil Ramsamooj and Chandrabhan Sharma, "*A Comparison of Artificial Neural Networks and Support Vector Machines for Short-term Load Forecasting using Various Load Types"*, IEEE June 2017.

[87] Joaquin Abellan, Carlos J. Mantas, Javier G. Castellano, "*A Random Forest approach using imprecise probabilities"*, Knowledge-Based Systems, July 2017

[88] Liliya Demidova, Yulia Sokolova, "*A Novel SVM-kNN Technique for Data Classification*", 6th MEDITERRANEAN CONFERENCE ON EMBEDDED COMPUTING, 2017.

[89] Lior Rokach, Oded Maimon, "DECISION TREES", DATA MINING AND KNOWLEDGE DISCOVERY HANDBOOK"

International Journal of Engineering Science Invention (IJESI) is UGC approved Journal with Sl. No. 3822, Journal no. 43302.

Rajni Mehta "An Exploration of Classification Approaches For Real Life Image Sets." International Journal of Engineering Science Invention(IJESI), vol. 6, no. 12, 2017, pp. 55-70.