2

REVIEW OF LITERATURE

The River ice is being classified both beneficially and detrimentally, as they form a significant portion of the annual water cycle. They can cause flooding, restrict the hydro-power production, block water intakes and outfalls, expose toxins, damage the fish habitat, threaten the structures such as bridges, culverts and disrupt the River navigation. Many works have been carried out to differentiate the River ice types into different classes that help in classification accuracy, and solve the problems as they could cause.

According to the texture, River ice was classified as large and difficult. Real River ice samples collected from Stuttgart, Germany were classified, using digital image processing techniques for image de-noising, texture segmentation, feature extraction, feature selection and finally classification. Several works have been proposed to perform the River ice classification from Synthetic Aperture Radar (SAR) images (Frank Weber et al., 2003; Yves Gauthier et al., 2006; Mermoz et al., 2009). Among these methods, to improve the classification accuracy to classify different ice types, feature selection played a major role.

To determine which features should be employed for the best classification result is a key feature in selecting statistical feature selection pattern (Sebastián Maldonado et al., 2011). Feature selection is a technique used to identify a subset of features, which perform best under a given classification system (Jain and Zongker, 1997). Feature selection algorithm reduces the cost of running as well as the feature space by providing a better classification model due to the statically favored feature space that better fits the pattern recognition problem (Jain and Chandrasekaran, 1982).

Sequential Forward Selection (SFS) algorithm (Inza et al., 2000) identified a method which starts with an empty feature set and iteratively evaluates, adds features in a forward manner. If a feature has been identified along with the combination of feature space being evaluated, it can be assigned permanently as a member of the selected set, and also helps in removing errors in possible features to be selected in the next phase of feature space search. This process is being continued, as, after removing all unused feature, the result still gives an improvement to the classification accuracy. The River ice classes were classified in the form of a binary decision tree for this analysis and were applied three times separator to separate. However, this approach does not consider spatial variability features, noises in the images and the patterns within regions of an image. According to the survey, classification is performed without feature selection as image was collected from the original feature set.

This chapter presents a detailed analysis of some of the image denoising methods, image texture segmentation methods, feature extraction methods, feature selection techniques and classification methods for River ice images.

2.1. Feature Subset Selection Methods

Feature selection aims at identifying features in the data-set as important, and discard any other feature irrelevant and redundant to the classification information. Since feature selection methods help in reducing the dimensionality of the data, and classification algorithms can be operated faster and more effectively, after performing feature selection. In some cases, as a result of feature selection there is an improvement in the performance. The reason is that the target concept has more compact and easily interpreted representation.

The objectives of feature selection are manifold, the most important ones being:

a) To reduce dimensionality and to remove irrelevant and redundant data to improve model performance, i.e., prediction performance in the case of supervised classification and better cluster detection in the case of clustering,

b) To provide faster and more cost-effective models, and

c) To gain a deeper insight into the underlying processes that generated the data.

Feature Selection Algorithm (FSA) is a computational solution being motivated to certain definition of relevance. Furthermore, the relevance of a feature has several definitions depending on the objective, and an irrelevant feature is not useful but all relevant features are necessarily helpful for training. Feature selection methods and their representation of the methods in the literature are presented in Figure 2.1.

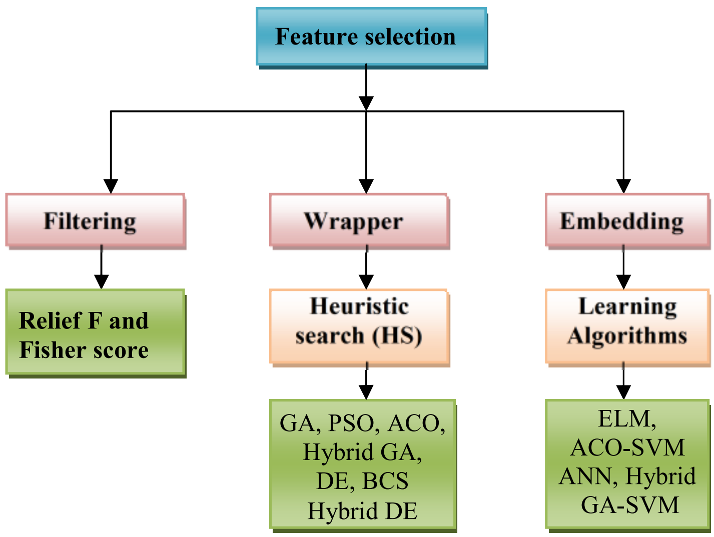

Figure 2.1: Feature Selection Approaches

2.1.1. Feature Selection with Filtering Methods

Filter techniques evaluate the relevance features by looking only at the intrinsic properties of the data. Feature relevance score is calculated and low scoring features are removed. Afterwards, these subsets of features are presented as input to the classification algorithm. Filter technique will easily scale to very high-dimensional datasets, computationally simple and fast, and independent of the classification algorithm, which favor the advantageous over other algorithms. As a result, feature selection techniques tend to be performed only once and then different classifiers can be evaluated. The ignorance of interaction with the classifier and that most proposed techniques express one variable as they form the most common disadvantage.

A. Relief F and Fisher Score

Relief F and Fisher Score proposed by Yuxuan Sun et al., (2011) is mainly implemented for supervising filter approach algorithm using the feature weights. They have considered 'n' instances that are randomly selected from the original features.

The simplest filter algorithm for feature selection proposed by Jian-Bo Yang et al., (2009) is Fisher score. In these criteria, features were selected with the similar values in the same class and the dissimilar values in different classes.

The main disadvantage is that they ignore the interaction with the classifiers. This means that each feature is considered separately, ignoring feature dependencies, leading to worse classification performance when compared to other types of feature selection techniques. To overcome the problem of ignoring feature dependencies, a number of multivariate filter techniques were introduced, aimed at the incorporation of feature dependencies to some degree.

To estimate a relevance index for each feature to measure how relevant a feature is to the target filters were being set. Then filter technique will rank features by their relevance indices and perform search according to the ranks or based on some statistical criterion, like significance level. The most promising characteristic of filters is that the relevance index is calculated, based solely on a single feature, without considering the values of other features. Such implementation implies that filters assume orthogonally between features, which usually is not true in practice. To overcome these problems, heuristic methods have proposed with feature subset selection.

2.1.2. Feature Selection with Wrapper Methods

Model construction with feature was incorporated in Wrapper methods. For Wrapper methods, different feature subsets are selected, a predictive model is constructed for each feature subset and the model which produced from the feature subset with highest predictive performance is selected. Wrapper methods have typically been used for small datasets with a small number of features. Wrapper methods are not suitable for large datasets as encountered in data mining.

A search procedure in the space of possible feature subsets are defined and various subsets of features are being generated and evaluated such that Wrapper methods embedded with model hypothesis search within the feature subset search. Training and testing a specific classification model, rendering this approach tailored to a specific classification algorithm evaluates a specific subset of features. A search algorithm is then “wrapped” around the classification model to search the space of all feature subsets. The space of feature subsets grows exponentially with the number of features, heuristic search methods used to guide the search for an optimal subset. These search methods can be divided into two classes: deterministic and randomized search algorithms.

In order to identify a satisfactory solution, a Heuristic search, the process of intelligently narrowing the search through a potentially very large search space of solutions have been followed. At every decision point in the search space, a heuristic search procedure employs a merit (heuristic) measure to determine the best path to expand in the search space. For the problem of feature subset search, the space of all possible problems is the set of all possible combinations (the power set) of the features in the set of candidate features. The candidate features are those features that have been pre-selected through a process of ranking.

Feature selection and reduction of pattern dimensionality, forms an important step in pattern recognition systems. Making use of the population-based optimization algorithms like Genetic Algorithm (GA) based method, Ant Colony Optimization (ACO)-based method and Particle Swarm Optimization (PSO)-based method for feature selection.

A. Genetic Algorithm (GA)

D.Goldberg (1989) suggested Genetic Algorithm which highlights selection of features with important gender information and improves in the classification performance. GA belongs to the class of randomized heuristic search techniques with attractive approach in feature subset selection. Their use in the area of computer vision is limited, in spite of being used in various pattern recognition applications and updating the local result also does not produce best results.

B. Ant Colony Optimization (ACO)

A novel feature selection method based on Ant Colony Optimization (ACO) was presented by Hamidreza Rashidy Kanan et al., (2007). ACO algorithm was motivated by social behavior of ant’s in their hunt for the shortest pathways to food sources. Most familiar approaches for ACO-Based feature selection utilize the previous information of features. On the other hand, the classifier performance and the extent of chosen feature vector are implemented as heuristic information for ACO. This method can be implemented, without any difficulty and as a consequence of using one simple classifier, its computational complexity is extremely low. ACO is practically hard to implement for larger samples, since updating of the phenomena phase requires more step.

C. Particle Swarm Optimization (PSO)

Bing Xue et al., (2013), proposed a study based on multi-objective particle swarm optimization (PSO) which helps in feature selection. Two PSO-based multi-objective feature selection methods are employed to generate a Pareto front of non-dominated solutions. Introduction of the idea of non-dominated sorting technique into PSO is to address feature selection problems by the first algorithm. The concepts of crowding, mutation and dominance to PSO to seek out the Pareto front solutions were applied by the second algorithm. The two multi-objective approaches are evaluated against two traditional feature selection methods, a single objective feature selection method, two-stage feature selection method and three most familiar evolutionary multi-objective algorithm on 12 benchmark data sets are tested. The major disadvantage of the proposed system is less local search and it does not achieve the best result.

D. Hybrid Genetic Algorithm

To find a subset of features relevant to the classification task, a hybrid genetic algorithm by Huang J et al., (2007) was adopted. Two stages of optimization are involved, in which the inner and outer optimizations associate with each other and accomplish maximum global predictive accuracy in addition to the maximum local search efficiency.

Hybrid GA for feature selection has been formulated by integrating the potential global search capability of GA with certain well-organized local search heuristic approaches (Huerta, E. B., et al., 2000; Shaohua Wu et al., 2013; Deekshatulu, B.L. and Chandra, P. 2013). A new immune clonal GA based on immune clonal algorithm, called ICGFSA, designed to solve the feature selection problem was suggested and discussed. The ICGFSA approach has more searching and operation capabilities because of the clonal selection theory that an antibody has the potential to clone certain related antibodies in the solution space. Each antibody in the search space specifies a subset of the possible features. But in clone selection method, they have some of the disadvantages like poor convergence properties and difficulties in reaching high-quality solutions in reasonable time.

E. Binary Cuckoo Search

The feature selection with the most discriminative set of features by enhancing the recognition rates and to make feature extraction faster was pursued in the last years. A new feature selection technique called Binary Cuckoo Search (BCS), based on the behavior of cuckoo birds, was proposed by Rodrigues D, (2013). The experimentations were done in the perspective of theft discovery in power distribution systems using two datasets acquired from a Brazilian electrical power company and have revealed the robustness of this approach in opposition to several others nature-inspired optimization approaches. The permutation functions applied on nests matrix will affect the performance of the system.

F. Differential Evolution (DE)

Utpal Kumar Sikdar et al., (2014) developed Differential Evolution (DE) based feature selection technique for anaphora resolution in a resource-poor language, namely Bengali. They discussed the issues of adapting a state-of-the-art English anaphora resolution system for a resource-poor language like Bengali. The quality of high accurate detector and the use of appropriate features for anaphora resolution system, help in evaluating the performance of any anaphoric resolver.

The feature selection process is carried out in datasets comprising a huge amount of features. Dimensionality reduction of the original feature set indicates the complication of choosing appropriate features, which produce the most predictive outcome. Rough set theory is a most promising methods employed for feature selection approach, which is still not able to find optimal subsets. On the other hand, it can be extended to be more optimal by means of different optimization approaches. Feature selection technique using the Rough set theory hybrid with Differential Evolution (DE) is formulated by Suresh Chandra Satapathy et al., (2012). In DE, there exist many trial vectors generation strategies, out of which, a few may be suitable for solving a specific complication application, as the proposed approach is experimentally evaluated against other hybrid Rough Set methods such as GA and PSO.

G. Hybrid Differential Evolution (DE)

To overcome the curse of dimensionality problem is one of the primary inspirations for the process of feature selection. Rami N. Khushaba et al., (2011) presented a novel feature selection method utilizing the integration of DE optimization technique and a formulated repair mechanism having feature distribution measures. This approach, termed as DEFS, makes use of the DE float number optimizer in the combinatorial optimization complication of feature selection. With the aim of making the solutions produced by the float-optimizer appropriate for feature selection, a roulette wheel structure is built and supplied with the probabilities of features distribution. DEFS approach is employed to select optimal subsets of features from original datasets with unreliable dimensionality. In the distribution measures the value are measured based on setting a threshold value in the feature selection approach, setting a threshold value will reduce the classification accuracy results.

Khushaba et al., (2008) formulated a feature selection approach depending on DE optimization scheme. The new algorithm, called DEFS, modified real-valued optimizer to deal with the complication of feature selection. DEFS greatly lessens the computational costs and simultaneously proving to provide commanding performance. The DEFS approach is executed to a Brain-Computer-Interface (BCI) application and compared against other dimensionality reduction methods. The results point out the importance of this algorithm in terms of solutions optimality, computational cost and memory requirement.

2.1.3. Feature Selection with Embedding Methods

In Embedded techniques, an optimal subset of features is built into the classifier construction and can be seen as a search in the combined space of feature subsets and hypotheses. Just like Wrapper methods, Embedded methods are specific for a learning algorithm. Embedded methods have the advantage of being less computationally intensive than Wrapper methods. In Wrapper approach, the interaction with the classification model is included. The Embedded method (for instance Guyon and Elisseeff, 2003) is similar to the Wrapper approach in the way that the features are specifically selected for a certain inducer, but it selects the features in the process of learning.

A. Hybrid Genetic Algorithm (GA) approach combined with Support Vector Machines (SVM)

Edmundo Bonilla et al., (2005), proposed a Hybrid Genetic Algorithm (GA) approach combining the Support Vector Machines (SVM) for the classification of high dimensional Microarray data and in association with a fuzzy logic based pre-filtering technique. A SVM classifier evaluates and GA optimization evolve gene subsets. Using archive records of "good" gene subsets, a frequency based technique helps to identify the most informative genes.

B. Extreme Learning Machines (ELM)

M. Termenon et al., (2013) presented a Computer Aided Diagnosis and image biomarker identification system for cocaine dependence selecting appropriate areas from a collection of brain structural Magnetic Resonance Images (MRIs). A region selection phase discovers the most appropriate watershed areas. M. Termenon et al., (2013) formulated two different schemes to differentiate region relevance:

(a) Embedded procedure using extreme learning machines (ELM), and

(b) Execute correlation distribution percentiles to choose most discriminant areas.

Feature selection path shows best feature subset having the context to the number of chosen features and the generalization error. In every subset size, the sparsity-error trade-off curve confirms the matching generalization errors. To choose appropriate feature subsets and obtain functional domain information was being helped by these graphical tools. In order to obtain these tools, an extreme learning machine proposed by Frenay Benoit et al., (2013) was used here. They are fast to train and estimate their generalization error by using the PRESS statistics. A new approach is commenced, which inserts an additional layer to typical ELMs to optimize the subset of selected features is introduced.

C. Ant Colony Optimization (ACO) and Support Vector Machines (SVM)

Cheng-Lung Huang et al., (2009), presented a novel hybrid ACO-based classifier model that incorporates Ant Colony Optimization (ACO) and Support Vector Machines (SVM) to improve classification accuracy with a little and proper feature subset. Concurrently, for the purpose of optimizing the feature subset and the SVM kernel constraints, the feature weight and the pheromones are passed down to find out the transition probability, the weight vector and the classification accuracy of the feature offered by the SVM classifier are taken into account to update the pheromone. The major disadvantage of SVM classifier is that it is applied only for two classes.

One of the significant research problems in multivariate analysis is the selection of a subset of input variables which can predict the desired output with an acceptable level of accuracy. The goal of attaining the elimination of the variables that produce noise or strictly correlated with other already selected variables was achieved. Feature subset selection (selection of the input variables) is essential in parallel analysis and in the field of classification and modeling.

D. Ant Colony Optimization (ACO) and Artificial Neural Networks (ANNs)

Based on Ant Colony Optimization (ACO) and Artificial Neural Networks (ANNs), Rahul Karthik Sivagaminathan et al., (2007) presented a hybrid method to report feature selection. In this method, ANNs is employed as a classification function rather than using the nearest neighborhood algorithm for evaluating the ‘‘goodness’’ of the subset developed at each stage. The proposed hybrid model yields promising results by demonstrating data sets from the medical diagnosis domain. In order to improve feature selection results, the local pheromone update in the ACO becomes one of the most important strength parameters. However, quantification of the impact of exploitation probability factor and other important parameters need to be studied further.

For identification of key feature genes in a complex biological phenotype, Li et al., (2005) proposed an efficient hybridization genetic algorithm to exploit fully their respective. The large B cell lymphoma can be diffused by applying this technique to the microarray data. The behaviors and properties for mining the high-dimension data of genome-wide gene expression profiles can also be demonstrated using this technique.

Among all three feature selection methods, the Wrapper and Embedded methods are the most popular as the Wrapper methods have the main advantage of considering the interaction between feature subset search and model selection, thereby to consider feature dependencies. The major risk of over-fitting is the common drawback of the filter technique out of these techniques. These techniques are also computationally intensive, if there exist a high computational cost for building the classifier model.

Similar to Wrapper approach, Embedded approach (Guyon and Elisseeff, 2003) selects the features specifically for a certain inducer, and the only difference is that it has been selected in the process of learning. In view of this, Wrapper methods are being combined with learning algorithm for improving the feature selection results.

2.2. Impact of Pre-processing in Feature Selection

The quality of the image can be improved by pre-processing the images which includes the common steps such as removing low-frequency background noise, normalizing the intensity of the individual images and also removing reflections and masking. This work focuses on the improvement of the feature selection and classification results by removing the noise from images. The impact of pre-processing methods in feature selection methods are analyzed here and are illustrated in Figure 2.2.

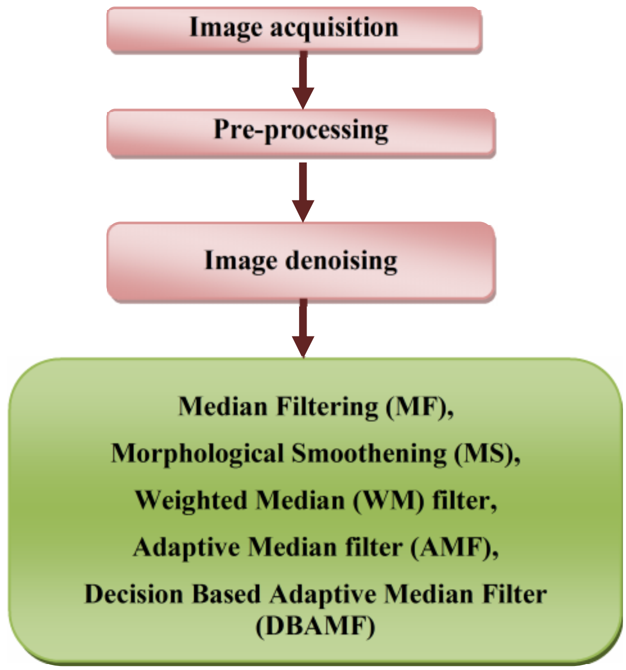

Figure 2.2: Denoising Filters used in Pre-Processing

A. Median Filter (MF)

In the pre-processing step, Nivetha et al., (2014) developed an image enhancement technique for applying some of the pre-processing techniques like Median filtering, Histogram equalization, noise removal and so on. The MATLAB software has been used for image pre-processing. The selective removal of the redundancy in the scanned images without affecting the details that are required in the diagnostic process is the main aim of the image pre-processing.

With the help of SVM classifier the extracted images are classified, based on the stages of disease, so it will be useful for the physician to provide therapeutical suggestions. The higher rate of early detection of lung cancer has been possible by using this proposed system. The image features are extracted and particular features are selected, using GLCM method and genetic algorithm respectively. The evaluations of other feature extraction and feature selection methods were also investigated for the classification of lung images. The widespread screening by CT or MRI is impractical, thus implying the most common procedure for diagnosis of lung disease is chest radiology. The common lung diseases include asthma, COPD and infections like influenza, pneumonia, tuberculosis, lung cancer and many other breathing problems. The image pre-processing step is thus necessary to improve the quality of the images and also to make the feature extraction phase more reliable.

Computer Aided Diagnosis (CAD), using image processing, have been studied by many researchers for improving the performance of the breast cancer screening. The breast cancer detection, using mammogram images, has been reported by K.Thangavel et al., (2012). By using median filtering, pre-processing is initially carried out, and features are extracted followed by identification of abnormality by means of classification. In case of employing all extracted features, misidentification exists, and hence, feature selection procedure is required. This work proposes Fuzzy- Rough set feature selection with π membership function. The abnormalities are classified, using the selected features, with the help of Ant-Miner and Weka tools.

The noise pixel in the images can be detected using a novel median filtering method, as presented by X L Ma, (2009). However, the median filter cannot distinguish between noise and fine detail and hence removes both, as the difference cannot be sorted out.

The filtering is thus carried out based on size, in which anything relatively smaller in size compared to the neighborhood will have minimal effect on the value of the median, and will be filtered out. Most of the denoising solutions reported so far have focused largely on noise removal by ignoring the edge details. Therefore, other versions of Median filters have been used in this study to overcome the above drawbacks.

B. Median Filter and Morphological Smoothening (MS)

The pre-processing technique for noise removal from the lung CT images was discussed by Bhuvaneswari et al., (2014). In the data preprocessing step, the input image is considered for effective feature extraction techniques, by applying the median filter and morphological smoothening. The automatic classification of lung images into Pleural effusion, Emphysema, Bronchitis and normal lung scan has been studied. Features extraction can be carried out based on Gabor filter (H.B. Kekre et al., 2010, Walsh hadamard transform (Alina Banerjee et al., 2013) whereas selection can be through Correlation based Feature Selection (CFS) (Mital Doshi et al., 2014), and Principal Component Analysis (PCA) (Ronald A. Holser, 2012). Then, the classification is done using K- Nearest Neighbour (KNN) (M. Akhil Jabbar et al., 2013). The naive bayes, decision tree, J48, MLP-NN are the widely used classifiers to train, test and classify the features among which the MLP-NN classifier offers performance of about 81.4% compared to other classifiers. This works assists in the detection of the lung diseases using image processing technique.

C. Weighted Median Filter

A novel image-denoising filter based on the standard median (SM) filter is proposed by Chin-Chen Chang et al., (2008) in which a threshold and the standard median is used for noise detection. The original pixel value, which is closer to or the same as the standard median, will also, be changed to a newer value using weighted median (WM) filter. The experimental results showed that, Weighted Median (WM) filter performed better that the Standard Median (SM) filter and the Tri-State Median (TSM) filter.

D. Adaptive Median Filter

The image adaptive median filtering algorithm has been described by Yang Heng-fu, et al., (2009) in which luminance, edge and texture characteristics of the host image were exploited. The possible noisy pixels were determined based on the noticeable difference using Human Visual System (HVS) masking characteristics. The sizes of the filter window size were adjusted adaptively by the noise density, and noisy pixels were removed by improved median filtering algorithm. The image details were preserved as well as the salt and pepper noise was removed using this method.

The adaptive median filter was also used by Lukac et al., (2003) in which local entropy of the samples is developed inside the filtering window. As the proposed method is fully adaptive, no optimization is required, and moreover, the main disadvantages based on the local contrast probability in the work of Beghadi and Khellaf (year) has been eliminated. The excellent performance of the proposed method is due to the successful analysis of input samples, and the local entropy concept enables efficient differentiation between outliers and desired edge samples.

E. Decision Based Adaptive (DBA) Median Filter

The detection of salt and pepper noise in the image has been proposed in Decision Based Algorithm (DBA). The pixel values are checked against the maximum and minimum values in the selected window selected for detecting the noisy and noise-free pixels in the image. The maximum and minimum value of the impulse noise exists in the dynamic range (0, 255). The noise-free pixel is identified to have a value between the minimum and maximum values in the currently processed window and no modification will be done to that pixel. If the pixel value does not lie within the range, it is identified as a noisy pixel, and hence will be replaced by either the median pixel value or by the mean of the neighboring processed pixels in case of noisy median.

Though it shows better performance at low noise rates, the effective suppression of impulse noise and preservation failed in this method. To overcome such problems and to consider noise-free pixels for calculation, (Anisha Bhatia, 2013) proposed a Decision Based Adaptive Median filter to calculate the median value from the window. The calculation of the median starts from the rest of the pixel values, by omitting the pixel values of 0 or 255. Thus, the use of adaptive windowing approach is this algorithm helps, in expanding the size of its filtering window, based on the local noise density for facilitating the filtering high density of salt-and-pepper noise.

Hung-Hsu Tsai et al., (2012) designed a Median-type filter with a 2-level impulse noise detector called DPSM filter, for image enhancement, using a system that includes Decision tree, Particle swarm optimization and Support vector Machine regression. First, a pixel is determined to be contaminated by impulse noises or not by employing a varying 2-level hybrid Impulse Noise Detector (IND). The construction of 2-level IND is by a Decision Tree (DT), built based on the combinations of 10 impulse noise detectors. The DT is also optimized by exploiting the particle swarm optimization (PSO) algorithm. Further, the corrupted pixels were restored in the DPSM filter, using Median-Type filter with Support Vector Regression (MTSVR).

2.3. Impact of Image Segmentation Methods without Feature Selection

The method of grouping a set of objects, in which the objects in the same group (called a cluster) are more or less similar to each other, than to those in other groups (clusters), is termed as cluster analysis. In statistical data analysis, clustering is an important task used in many fields such as River ice image classification, segmentation of the image into several regions, computer vision, remote sensing and stock market analysis. In case of image segmentation, it is defined as the process of identifying groups of similar image primitive. The methods followed so far, to perform image segmentation without feature selection impact, are illustrated in Figure 2.3.

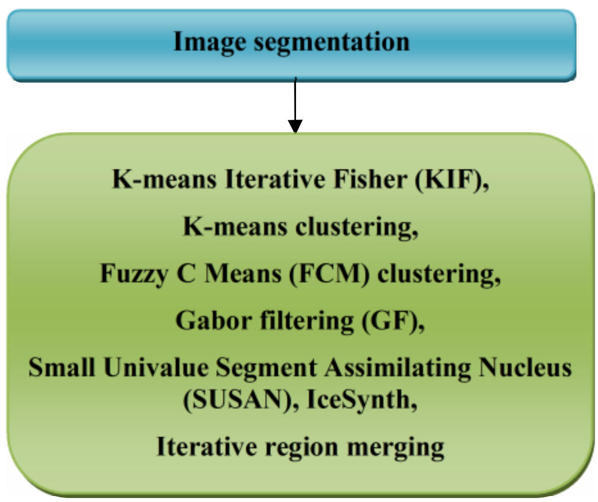

Figure 2.3 Image Segmentation Methods

A. K-Means Iterative Fisher (KIF)

The problem of image texture segmentation can be solved by a robust and unsupervised clustering algorithm called K-Means Iterative Fisher (KIF) algorithm as proposed by Clausi (2002). KIF algorithm involves two steps in which K-Means is applied initially, and then the parameters required for a Fisher Linear Discriminant (FLD) are estimated, using the K-Means class assignments. The implementation of this unsupervised methodology proved to be better than other published texture segmentation results that employ a wide variety of test imagery. However, this cannot be applied to color image samples which remain as the only drawback.

B. K-Means Clustering

A novel image segmentation method was presented by Anil Z Chitade et al., (2010) which is based on color features with unsupervised K-Means clustering algorithm. This step requires no training data, and the whole work is partitioned into two stages. In the first stage, the color separation of satellite image was enhanced using de-correlation stretching. In the second stage, K-Means clustering algorithm was used for grouping the regions into a set of five classes. Based on this two-step process, the computational cost can be reduced by avoiding the process of feature calculation for each pixel in the image. Even though the color is not commonly employed for image segmentation purposes, it provides a high discriminative power of areas present in the image.

In most of the image analysis or computer vision applications, image segmentation is an important task. Zainab M. Hussain, (2013) presented a texture image segmentation method using Gabor Filter, Anisotropic Filter and K-Means clustering algorithm. The texture in the image was analyzed using the Gabor filter as a multi-channels filter. The anisotropic diffusion filter helped in the extraction and enhancement of the obtained texture features. Then the K-Means algorithm helps in clustering pixels into a number of clusters that represent the texture regions. The evaluation of segmentation quality in this method was done using Ultimate Measurement Accuracy (UMA) metric.

C. Fuzzy C Means (FCM) Clustering

A satellite image segmentation technique, using the features of M-band Fuzzy C-Means features, was presented by Manimala Singha et al., (2011). The basic mechanism behind the image features is understood, by providing the variations in the reflectivity of surface materials, across different spectral bands in remotely-sensed multispectral imagery. In case of remote sensing, Fuzzy methods have received growing interest, owing to their significance in situations having inherent Fuzzy geographical phenomena.

In the proposed approach, multi-spectral image has been enhanced initially, followed by applying clustering technique using L a*b* color space and the vectors as inputs for the K-Means or FCM clustering methods. In case of segmented image, the regions are distinct from each other with respect to color and texture characteristics, and hence application of the number of the C defined by user is impractical to all the image samples.

A Fuzzy Kohonen Clustering Network (FKCN) clustering was proposed by Denzil Vinod Fernandes et al., (2011) which seem to be closely related to the Fuzzy C-Means (FCM) algorithms. The several drawbacks of KCN can be overcome, by integrating FCM with Kohonen Clustering Network (KCN), which is named as FKCN. The principles of fuzzy membership values are combined for learning rates and those rates subsequently help in updating the cluster center. The formulation of the present algorithm is based on the modification of the objective function of the standard FCM algorithm. This algorithm is known to be very efficient, as the time taken for training the neuron has been reduced. However, the improvised version of this algorithm will help in implementing for applications such as brain mapping, retinal image segmentation, etc.

D. SUSAN Algorithm

The principle and availability of SUSAN algorithm for corner detecting has been analyzed by (HENG,Y. et al., 2011; Li He-lin et al., 2012). This formed the basis for adopting the improved SUSAN algorithm for extracting the corner features of the Yellow River Model images. The self-adaptive threshold method of gray level difference has been utilized in this method, followed by which the method of setting its upper and lower limits of geometric threshold was adopted, to realize the automatic identification and detection for the Yellow River corner. Experimental results proved that this method of corner detection is highly suitable for de-noising. The boundary of River images was found to be clear, true and meticulous with accurate location. Although the detection of stereo matching and 3D reconstruction is enabled using this method, matching of different River image samples with 3D construction, however, remains a problem.

E. IceSynth

In general, tasks such as intractability of getting ground-truth segmentation data from Polar areas, the assessment of automatic SAR sea-ice image segmentation approaches are restricted to tests, and simple synthetic tests related to valid SAR imagery in accordance with the pseudo-ground truth data and basic shape primitives respectively. The evaluation of automatic segmentation approaches, in a systematic and consistent way, by means of using realistic circumstances which is always difficult. Hence, (Alexander Wong et al., 2009) introduced image synthesis system named IceSynth to overcome this limitation. It is capable of generating various synthetic sea-ice images that represent real SAR sea-ice imagery. In case of IceSynth, SAR sea-ice textures, are synthesized by means of stochastic sampling for each ice type, are estimated in accordance with non-parametric and local conditional texture probability distribution function. A stochastic sampling scheme, in accordance with the non-parametric local class probability distribution estimates, helps in generating large-scale sea-ice structures of several ice categories based on ice classification priors derived from real SAR sea-ice imagery, and also, some of the extreme types of River ice types based on the quantities becomes complex are calculated.

F. Iterative Region Merging

The method of segmentation of SAR sea ice images using a hybrid method was presented by Yue, B. and Clausi, D. A. (2001); Xuezhi Yang et al., (2007), constituting a preliminary watershed segmentation, subsequently a region merging. Iterative bilateral filtering mitigates speckle noise and suppresses irrelevant image details that will further significantly alleviate increased segmentation of watersheds. The precise location of object boundaries is possible, using the watershed algorithm as the edges are preserved extremely well by bilateral filtering. The final segmentation is done by applying an iterative region integrating on the watershed areas, with the consideration of local boundary potentials and local statistics. The effectiveness of this approach has been revealed on the segmentation of SAR sea ice images. However, some other noise such as the Gaussian, Salt and Pepper cannot be removed by these segmentation methods.

2.4. Impact of Feature Extraction Methods

The accurate description of large set of data requires amount of resources which is possible, by the method of feature extraction. During the analysis of complex data, the major problems seemed to be arising from the number of variables involved. A large amount of memory and computation power is generally required for analyzing large number of variables. Otherwise, a classification algorithm that overfishs the training sample may be required for new samples. The methods of constructing combinations of the variables, to overcome these problems, are generally termed as feature extraction in which data is represented with sufficient accuracy. Feature extraction methods that are performed with and without feature subset selection techniques are presented in Figure 2.4.

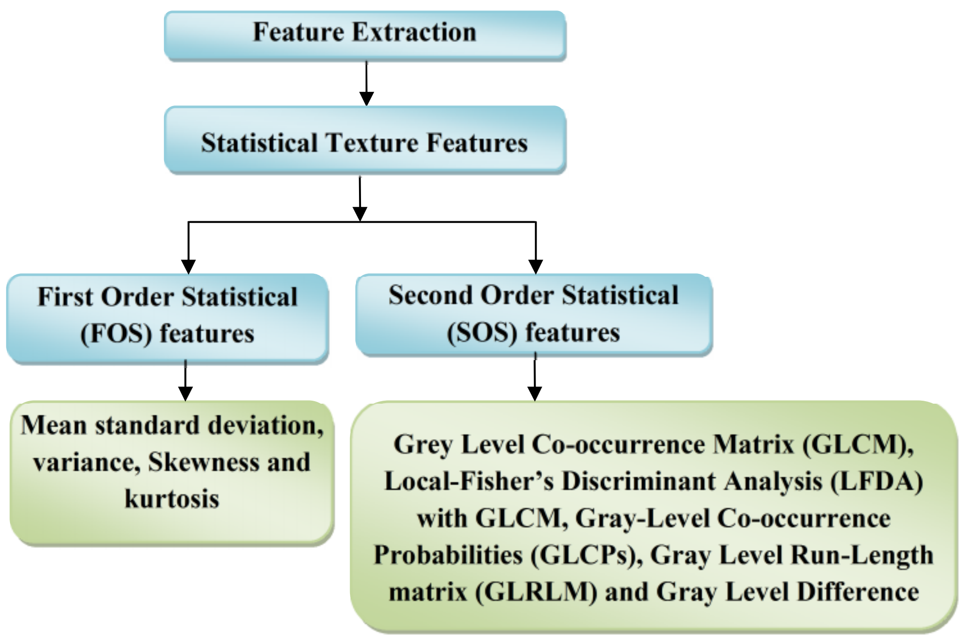

Figure 2.4 Feature Extraction Methods

2.4.1. Impact of Feature Extraction Methods in Feature Selection

A. Grey Level Co-occurrence Matrix (GLCM)

The texture was described as a measure of surface roughness by Naseri et al., (2012), if the intensity values of an image were considered as elevations. From the preprocessed image, four first order features, namely, Mean, Standard deviation, Entropy and Variance are extracted. Then, the second order textural features are extracted using Gray-Level Co-occurrence Matrix (GLCM) which represents a set of features to reduce the misclassification of glaucomatous images. The method of different combinations of pixel brightness values in an image can be depicted by GLCM. Second order textural features are extracted for different quantization levels, namely, 8, 16, 32, 64, 128 and 256 occurring in four orientations such as 0^0, 45^0, 90^0 and 135^0 for various distances. The method of Sequential Forward Floating Selection (SFFS) was employed to select the extracted features. The selected features are then fed to Back Propagation Network (BPN) in order to classify features as normal and abnormal images. The results showed that the proposed computer aided diagnostic system achieved 96% Sensitivity, 94% Specificity and 95% accuracy and hence were suitable for screening purposes. This study has reported that the analysis of gray levels has found significance in the classification of glaucoma.

The monitoring of River ice based on the contextual analysis approach, using RADARSAT data, was presented by Gauthier et al., (2003). In this method, it is assumed that the radar backscattering is influenced by the structure and composition of the ice cover. The optimization of the River ice characterisation from radar data is done by using contextual information. The establishment of context of the River channel and environment is performed, and homogeneous reaches are determined through a GIS, and are determined. The texture analysis helps in the establishment of the spectral context, by improving unsupervised classification, using a single texture parameter. The use of contrast measures, orderliness measures and descriptive statistics based on the discrimination Grey Level Co-occurrence Matrix (GLCM) between different ice conditions are measured.

B. Genetic Algorithm (GA) based feature selection and Local-Fisher's Discriminant Analysis (LFDA) with GLCM

The drawback in Hyper-Spectral Imagery (HSI) image analysis has been effectively solved by means of a robust dimensionality reduction approach, using spectral and spatial features. GA-LFDA is a novel dimensionality reduction algorithm proposed by Minshan Cui et al., (2013). In this method, both feature selection and projection based on Genetic Algorithm (GA) and Local-Fisher's Discriminant Analysis (LFDA) respectively, are performed in a raw spectral-spatial feature space for achieving superior dimensionality reduction. The spatial content of the HSI has been fully exploited by this approach with the help of Grey Level Co-occurrence Matrix (GLCM) extracted textural features from each spectral channel.

C. Gray Level Co-occurrence Probabilities (GLCPs)

In feature extraction methods image, the main aspect of the computer-assisted discrimination of SAR sea ice imagery lies in texture interpretation. The widely used approach to solve this problem is Co-occurrence probabilities. However, there are also other texture feature extraction methods, but their ability of SAR sea ice imagery interpretation has not been fully studied.

Gray Level Co-occurrence Probabilities (GLCPs) were presented by Clausi D. A. (2004). However, only limited reports are available, based on the comparison of different texture methods, especially related to segmenting remotely sensed imagery. The investigation of the significance of window size in texture feature consistency and separability of handling multiple textures within a window is performed. Moreover, required testing has been performed on synthetic (MRF generated) samples. It is due to the fact that, compared to MRFs, GLCPs are highly sensitive to texture boundary confusion with respect to segmentation procedures.

D. Gray Level Run-Length Matrix (GLRLM)

In another work of Chang et al., (2008), a seven feature extraction technique based on gray level run-length matrix, laws texture energy measures, neighboring gray level dependence matrix, wavelet features and Fourier feature were adopted based on local Fourier co-efficients. Using GLRLM, 78 textural features were extracted, among which six were selected and classified, using five binary SVM. The proposed method helped in successfully identifying six kinds of thyroid nodules. However, time consumption remained as a drawback.

E. Gray Level Difference Matrix (GLDM)

A Gray Level Difference Matrix (GLDM) method was proposed by Priyanjana Sharma et al., (2013) for feature extraction. The module of feature extraction constitutes 89 features from five textures (features) analysis methods for each, and every ROI were calculated. For given CT image, 6 features from First Order Statistical (FOS), 48 features from Spatial Gray-Level Dependence Matrix (SGLDM), 20 features from Gray-Level Difference Matrix (GLDM), 12 features from Laws Texture Energy Measures (TEM) and 3 features from Fractal Dimension Measurements (FDM) were extracted. The diagnostic Sensitivity and Specificity will be improved and at the same time interpretation time will be decreased by means of different feature extraction methods. However, various issues regarding the application of CAD systems into clinical practice still need to be investigated.

2.5. Importance of Feature Selection Approaches in Classification

The process in which ideas and objects are recognized, differentiated and understood is defined as classification, or otherwise called as categorization. The impact of the classification method with and without feature selection methods are investigated in this work. The representation of existing methods of classification with and without feature selection methods is diagrammatically represented, as shown in Figure 2.5.

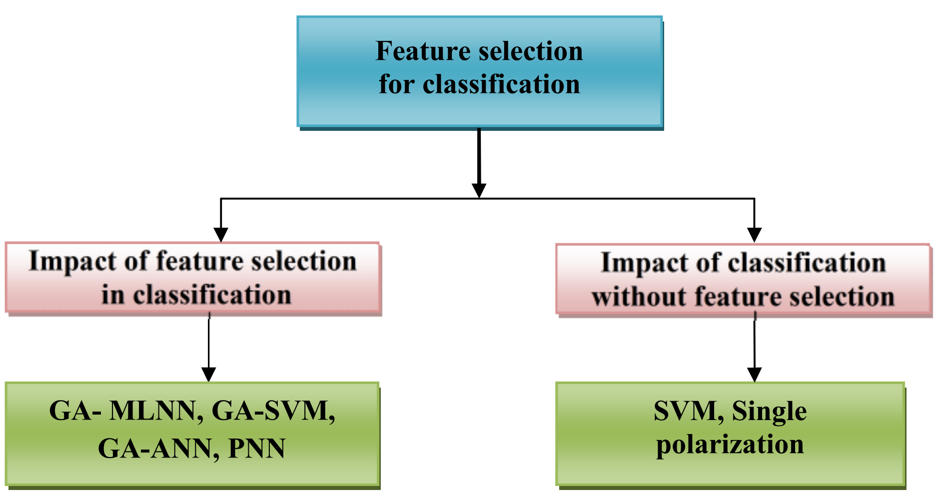

Figure 2.5 Classification Methods

2.5.1. Impact of Feature Selection Approaches in Classification

A. Genetic Algorithm with Support Vector Machine (GA-SVM)

Genetic algorithm has been confirmed to be an extremely constructive solution in a great variety of optimum search problems. Recently, Huang and Wang postulated a genetic algorithm to optimize these parameters and input feature subset of SVM without any loss of accuracy in classification problems (C. L. Huang et al., 2006), timing of events is not explicitly modeled; ordering of events is preset by the analyst.

The contrast between ice and wet snow is remarkable. The ice from the ground on SAR images were difficult to distinguish due to similar backscatter characteristics. In this study, (Lei Huang et al., 2011), revealed the distinction by target decomposition. Support Vector Machines (SVMs) are performed to classify the glacier areas using the selected features. The glacial areas are classified into six parts: wet snow, ice, River outwash, soil land, rocky land and others.

B. Differential Evolution (DE) with Support Vector Machines (DE-SVM)

Ekbal, U. K. S. A. and Saha, S. (2012); Utpal Kumar Sikdar et al., (2014), suggested a DE based two-stage evolutionary scheme for Named Entity Recognition (NER). The initial stage takes care of the complication of related feature selection for NER inside the frameworks of two common machine learning methods, namely, Conditional Random Field (CRF) and Support Vector Machine (SVM). The solutions of the ultimate best population offered different diverse set of classifiers; few are efficient in terms of recall while few are efficient in terms of Precision. In the second stage, a new scheme for classifier ensemble for integrating these classifiers was formulated. This scheme is extremely common and can be executed for any classification problem.

C. Probabilistic Neural Network (PNN) with Feature Selection

N. Azzri and A. Mohamed (2009) postulated a 39-bus system by means of the Probabilistic Neural Network (PNN) together with enhanced feature selection and extraction methods. The inspected power systems are classified into smaller relaying the coherency of the regions when subjected to interruptions. This is to diminish the quantity of data sets derived for the relevant regions. Transient stability of the power system was decided in accordance with the generator relative rotor angles derived from time domain simulations and performed by considering three phase faults at various loading circumstances. The data obtained from the time domain simulations were employed as inputs to the PNN classifier. An improved feature selection and extraction techniques integrate to diminish the input features to the PNN, which is employed as a classifier to decide whether the power system is stable or unstable. It can be noted that the PNN with improved feature selection and extraction techniques lessens the time taken to train the PNN, without influencing the accuracy of the classification results.

A Probabilistic Neural Network with image and data processing approaches was utilized to execute an automated brain tumor classification, was postulated by Othman et al., (2011). The conventional method for medical resonance brain image classification and tumors recognition is through human examination. Operator-assisted classification techniques are unworkable for huge amounts of data and are also non-reproducible. Medical Resonance images include a noise generated by operator performance which can guide to severe inaccuracies in classification.

2.5.2. Performance of Classification Methods without Feature Selection Approaches

A. Support Vector Machine (SVM) and Probabilistic Neural Network (PNN)

Saiti et al., (2009), have proposed SVM and PNN for classification. Selection of good subsets of features was done by using Genetic Algorithms (GAs) for improving the diagnosis rate. The classification accuracies obtained by this method is better, but SVM has performed better than PNN.

B. Single-Polarization based Classification

Mermoz et al., (2009) proposed a polarimetric SAR airborne image of the Saint-François River, Quebec, for investigation. Rule-based hierarchical classification is compared with a Wishart classification. A single-polarization-based classification is used to show the limits of this approach in discriminating water from ice. The hierarchical classification more accurately separates areas of ice from areas of open water. Detection of the thermal ice class is not highly accurate. Thermal and frazil ice classification is performed better in hierarchical classification than in Wishart classification. Rules generated during classification task, if the number of image samples become high it becomes complex to generate more rules. So a specific pruning is required to remove irrelevant rules in the classification task.

2.6. Comparative Study

During the recent years, (Yves Gauthier et al., 2006) formulated the potential of RADARSAT-1 data (C-band, HH polarization) for the purpose of classifying River ice categories and to determine ice cover features for harmless and efficient dam processes which British Columbia Hydro and Power Authority (BC Hydro) has been looking for. The preliminary results of this project have performed a combination of classification schema, to characterize the SAR data, providing helpful information regarding the ice front position and the ice states upstream and downstream of this point. It clearly reveals that, an integration of texture and backscattering images improvise the discrimination of ice cover categories, at some stage in freeze-up. Two preprocessing methods were used for the preparation of the original backscattering images. Firstly, images were processed according to a filtering approach. A filter named Kuan was applied to reduce speckle noise. To achieve further smoothing, the resulting images were again filtered with median filter. The original 16 bit power images are used by the second approach to create a texture image. A texture image extracts information from the spatial distribution pattern of the backscatter values. Efficiency of both methods as input to ice type classification is assessed, through the Fuzzy K-Means unsupervised algorithm, which segments the image, according to an iterative measurement of the distance to class mean and a predetermined number of desired clusters. The classification of ice cover types is possible with relatively high degree of confidence.

Frank Weber et al., (2003) found semi-automated classification of River ice categories on the Peace River using RADARSAT-1 SAR imagery. The investigation of the images was carried out using (*i*) Visually, and (*ii*) an unsupervised Fuzzy K-Means classification approach. The unsupervised classification partition the data into seven groups, which indicate the foremost ice cover category, observed on the Peace River. In order to simplify the process in an operational atmosphere, the unsupervised ice classification was semi-automated. The ice classification normally corresponds well with air-truth data and backscatter signatures from SAR images, as they were generated by spatial distribution of ice cover types. The position of the boundaries among the ice categories appears to be precise. The ice maps can be employed for ice monitoring, confirmation of ice models or decision-making functions.

Leen-Kiat Soh and Costas Tsatsoulis (1999) postulated a groundwork investigation for mapping sea ice patterns (texture) with 100-m ERS-1 SAR imagery. The authors employed Gray-Level Co-occurrence Matrices (GLCM), for the purpose of quantitatively evaluating textural constraints and representations, to determine the parameter values and representations. These values and representations are finest for plotting sea ice texture. They conducted experiments by investigating the consequences textural descriptors like entropy have in the representation of several sea ice textures for the quantization points of the image and the displacement and orientation values of the GLCM. A comprehensive gray-level demonstration of the image is not essential for the purpose of texture mapping, an eight-level quantization representation is undesirable for textural representation, and the displacement feature in texture measurements is more significant than orientation revealed by these studies. In addition, they formulated three GLCM implementations and assessed them by a supervised Bayesian classifier on sea ice textural circumstances. This experiment gave the result as the most excellent GLCM implementation in indicating sea ice texture was the one that exploits a variety of displacement values in order that both micro textures and macro textures of sea ice can be effectively derived. The definition of quantization, orientation and displacement values that are the finest for SAR sea ice texture analysis by means of GLCM were found experimentally.

An application of GLCM to texture based similarity evaluation of rock images was presented by Mari Partio et al., (2002). Retrieval results were evaluated and provided for two databases. One includes the complete images and the other with blocks acquired by dividing the original images. By computing distance among the feature vector of the query image and supplementary feature vectors, the database retrieval outcomes were obtained. Performance of the co-occurrence matrices was assessed with Gabor wavelet features. Co-occurrence matrices executed well for the particular rock image dataset. This similarity assessment application could lessen the cost of geological examinations, by providing enhanced accuracy in automatic rock sample selection.

S.Jayalakshmi and M.Sundaresan (2014) carried out a research work exploring the Iris Segmentation process with Fuzzy C Means approach and K-Means clustering apporach. Image acquisition, filtering, inner boundary localization, outer boundary localization and omission of eyelids and eyelashes were incorporated in this segmentation approach. Here, the segmentation process with the help of the images from CASIA iris dataset image existing online. The entire approaches are executed individually and the results are noted. The noted results of segmentation by means of FCM provide better accuracy rate of 98.20% and low error rate.

Microwave applications use micro strip antennas owing to their several benefits like small size and weight, compatibility and conformability with microwave Integrated Circuits (ICs), etc. There are many benefits like accessibility of a huge number of adaptable parameters in the form of aperture length, width and stub parameters, since; they have aperture coupled micro strip line feeding technique.

Design of aperture coupled micro strip antennas, by means of a newly formulated fast optimization algorithm called Differential Evolution was proposed by A. Deb et al., (2011). The optimization algorithm decides the dimensions in addition to the transmission line feed positions which suit ensembles optimum matching over a particular range of frequencies and substrate. Using Differential Evolution (DE) in the optimal design of aperture coupled micro strip antennas, fitness function is assessed by means of, method of moments approach executed through IE3D. In addition to the optimal design of the antenna possessing the traditional optimization approaches like real coded GA and PSO were experimented and their results and performances are compared with the DE technique.

2.7. Findings of the Literature Study

Table 2.1: Merits and Demerits of Feature Selection Methods

Feature Selection Methods	Merits	Demerits
Relief F and Fisher score	• Combined different measures of coding potential prediction, and then to retain only the rel evant ones. • When all model types involved in the present problem have been definite, certain features still include model type which is not referred to in the current issue.	• Filter methods are that they ignore the interaction with the classifiers. • This means that each feature is considered separately, thereby ignoring feature dependencies, which may lead to worse classification performance.
Genetic Algorithm (GA)	• Best feature selection results when compared to other conventional feature selection. • Local features selection achieves best results.	• Since generation the entire feasible permutations of holes and accessible blocks is difficult, they are bound to have over-fitting with sample data using the mentioned simple weighted approach. • In contrast, when the algorithm is not trained by means of long sufficient generations, subsequently the usability would be less.
Ant Colony Optimization (ACO)	• Limited ability to sense local environment. • Act concurrently and independently. • High quality solutions emerge via global co-operation.	• Convergence is guaranteed but it takes time. • Uncertain Coding is not straightforward.

Feature Selection Methods	Merits	Demerits
	• Positive Feedback leads to rapid discovery of good solutions.	
Particle Swarm Optimization (PSO)	• Its simplicity and easy implementation. • Algorithm can be used widely in the fields such as function optimization. • PSO adopts the real number code, and it is decided directly by the solution. The number of the dimensions is equal to the constancy of the solution.	• The major disadvantage of the proposed system is less local search does not achieve best result. • The method easily suffers from the partial optimism, which causes the less exact at the regulation of its speed and the direction. • The method cannot work out the problems of scattering and optimization. • The method cannot work out the problems of non-co-ordinate system, such as the solution to the energy field and the moving rules of the particles in the energy field.
Hybrid Genetic Algorithm	• The inner and outer optimizations co-operate with each other. • It achieves the high global predictive accuracy as well as the high local search efficiency.	• Poor convergence properties. • It has difficulty in reaching high-quality solutions in reasonable time.
Binary Cuckoo Search (BCS)	• It performs faster than the other optimizations.	• The permutation functions applied on nests matrix affect the performance of the system.

Feature Selection Methods	Merits	Demerits
Differential Evolution (DE)	• Simple yet and powerful population method for all applications. • Direct search algorithm for locally and globally optimizing functions with real value parameters.	• In DE, there exist many trial vectors generation strategies, out of which a few may be suitable for solving a particular problem only. • If the number of data samples becomes large it becomes hard to implement. • The weight values calculation for each feature values becomes assumption. • Population matrix will contain either the trial vector.
Hybrid Differential Evolution	• Unique solution for the output unit parameters is calculated. • Learning mode is suitable for the online applications.	• Setting threshold value will not exactly remove irrelevant features. • The classification results of the method will reduce the classification results.
Genetic Algorithm-Support Vector Machine (GA-SVM)	• Optimize the SVM classification parameters, the prediction accuracy.	• Classification accuracy is only improved. • More time to complete the classification.
Extreme Learning Machines (ELM)	• Completes the whole process without delay and produces a unique optimal solution without the requirement of iterations. • Accordingly, it has the benefits of simple parameter selection and quick learning speed.	• Current reported ELM and its superior version are only depends on the empirical risk minimization principle, which possibly will experience from over-fitting.

<table>
<tr><th>Feature Selection Methods</th><th>Merits</th><th>Demerits</th></tr>
<tr><td>Ant Colony Optimization (ACO) and Support Vector Machines (SVM)</td><td>Optimize the SVM classification parameters, the prediction accuracy.
SVM with local search achieves best results.</td><td>Theoretical analysis is difficult.
Time to convergence is uncertain (but convergence is guaranteed).</td></tr>
<tr><td>Ant Colony Optimization (ACO) and Artificial Neural Networks (ANNs)</td><td>A neural network can carry out operations that a linear program cannot.
If an element of the NN fails working, it can prolong without any setback by their parallel character.
A NN learns and does not require to be reprogrammed.
It can be realized in any application and without any complication.</td><td>The neural network needs training to operate.
The structural design of a NN is different from the structural design of microprocessors, as a result needs to be emulated.
Necessitates high processing time for huge NNs.</td></tr>
<tr><th>Proposed Methods</th><th colspan="2">Advantages</th></tr>
<tr><td>ELM-DEFS</td><td colspan="2">ELM-DEFS enables in selecting the best feature set with higher accuracy.
The error rate of the selected features results are calculated based on the ELM.
ELM for optimizing the weights in DEFS features selection method.
Improved classification accuracy for River ice types.</td></tr>
<tr><td>SELM-DEFS</td><td colspan="2">The algorithm uses the ideas of ELM for sequential learning which has been shown to be extremely fast with generalization.
For selecting the number of hidden nodes, no other control parameters have to be manually chosen. Thus faster feature selection process is obtained and the classification accuracy for River ice types is improved.</td></tr>
</table>

Table 2.2: Merits and Demerits of Pre-Processing Methods

Pre-Processing Methods	Merits	Demerits
Median Filter	• It is extremely easy to recognize and trouble-free to compute. In certain scenarios, it is derived just by examination. • Median lies at the center division of the series and consequently it is not influenced by the extreme values.	• In uncomplicated series, the item values have to be organized. When the series includes large number of items, subsequently the process becomes tedious. • It is a less representative average since it is not based on the entire items in the series.
Median Filter and Morphological Smoothening	• It reserves the edges without any loss. • Image quality is improved.	• It removes only corrupted pixel by the median value or by its neighboring pixel value.
Image Enhancement	• Enhanced results with exact identification of the contrast features. • Improved image quality.	• Noises in the images are not removed. • Quality of the image only improved.
Decision Based Adaptive (DBA) Median Filter	• Solves the problem of the median filtering. • Decision of the median filtering is solved based on the decision tree algorithm.	• Preserve the edges and fine details of the original image. • Fine details of the image edges are not removed.
Adaptive Median Filter	• Median filter and its variants usually perform impulse noise removal effectively. • Remove impulse noise exactly and smoothing of other noise. • Reduce distortion, like excessive thinning or thickening of object boundaries.	• It removes only salt and pepper noises in the images. • Edge details are not considered while removing noise in the images.

Proposed Filter	Advantages
DBA-PSO	• It performs efficient noise removal for Salt and Pepper noise. • Removal of noise without destroying the image information. In this scenario, there is a need for adapting the filter weights to the direction of edges accordingly using the PSO. • Fine tunes the details of the original image.

Table 2.3: Merits and Demerits of Segmentation Methods

Segmentation Methods	Merits	Demerits
K-Means Iterative Fisher (KIF)	• It requires no a priori knowledge of the number of classes, is a non-parametric solution. • It is computationally efficient compared to other methods used for clustering.	• The number of cluster selection and centroid selection becomes random. • Reduces the clustering or segmentation results.
K-Means Clustering	• If variables are huge, then K-Means most of the times computationally faster. • Low complexity is O(nkt), where t = #iterations.	• Necessity of specifying K. • Sensitive to noise and outlier data points. • Clusters are sensitive to initial assignment of centroid. • K-Means is not a deterministic algorithm. • Clusters can be inconsistent from one run to another.

Segmentation Methods	Merits	Demerits
Fuzzy C Means Clustering	• Allows a data point to be in multiple clusters. • A more natural representation of the behavior of genes. • Genes usually are involved in multiple functions.	• Need to define C, the number of clusters. • Need to determine membership cutoff value. • Clusters are sensitive to initial assignment of centroid. • FCM is not a deterministic algorithm. • Without consideration of the feature selection segmentation is performed.
Gabor Filter	• To identify different textured and non-textured regions in an image. • To classify/segment different texture regions in an image. • To extract boundaries between major texture regions.	• In Gabor filtering methods some of the texture is ignored. • Thus traditional Gabor filters are less suitable for this purpose.
Fuzzy-Kohonen Algorithm	• Centroid values are automatically updated, and objective function of the FCM is modified. • Improved segmentation results than the k means clustering.	• It does not extract various features of the image by using the various parameters such as area, mean, standard deviation, entropy, energy, etc. • Noises in the images are not removed important features are not selected.

Segmentation Methods	Merits	Demerits
SUSAN Algorithm	• This method of corner detection is good for de-noising. • Its boundary is clear, true and meticulous and the location is accurate.	• Feature selection and feature extraction are not performed. • Segmentation results get reduced and affect the classification performance.
IceSynth	• Assess automatic segmentation approaches in a systematic and consistent manner using realistic scenarios.	• Various ice categories based on ice classification priors obtained from real SAR sea-ice imagery, to calculate some extreme quantiles becomes complex. • Noises in the images samples are not removed.
Iterative Region Merging	• The efficiency of the proposed method has been demonstrated on the segmentation of SAR sea ice images.	• Some other noises such as Gaussian, salt and pepper, etc. are not removed in these segmentation methods. • Important feature in the River ice images are not selected.
Proposed Method	**Advantages**	
Fuzzy C Means Clustering	• Allows a data point to be in multiple clusters. • Fast, robust and easier to understand for image segmentation. • Gives best segmentation result for River ice images with distinct or well separated textures from each other.	

Table 2.4: Merits and Demerits of Feature Extraction Methods

Feature Extraction Methods	Merits	Demerits
Grey Level Co-occurrence Matrix (GLCM)	• It represents a set of features to reduce the misclassification of glaucomatous images. • It depicts how different combinations of pixel brightness values occur in an image.	• They require a lot of computation (many matrices to be computed). • Features are not invariant to rotation or scale changes in the texture, since the features are efficiently selected.
Gray-Level Co-occurrence Probabilities (GLCPs)	• Probabilities values are calculated to features to measure the feature extraction efficiently.	• Probabilities are calculated with less number of considerations of the features. • Gray level differences are not considered.
Gray Level Run Length Matrix (GLRLM)	• Shot-run emphasis (SRE), Long-run emphasis (LRE), Gray level Non-uniformity (GLN) and Run Percentage (RP) features are mainly focused.	• Sharpness (SHP), Second Moment of distribution of gray difference is not measured in the feature extraction step. • Larger level of the gray level differences is not focused.
Gray Level Difference Matrix (GLDM)	• This technique is usually used for extracting statistical texture features.	• It mainly focuses on statistical feature only. Other types of features are not considered.

Feature Extraction Methods	Merits	Demerits
Proposed Method	**Advantages**	
Grey Level Co-occurrence Matrix (GLCM)	• GLCM based texture image was considered to be the most suitable for a pixel-based classification. • GLCM mean texture differs from a standard average filter, because the pixel value is weighted by its frequency of occurrence in combination with a certain neighbour pixel value. • Texture is calculated over the entire image, thereby covering a wide range of patterns.	

Table 2.5: Impact of Feature Selection Approach in Classification Performance

Feature Selection based Classification Methods	Merits	Demerits
Genetic Algorithm with Multilayered Neural Network (GA-MLNN)	• The selection of such a subset will reduce the dimensionality. • Improved classification results in the MLNN.	• In simple series, the local selection of the features is not exact. • It reduces the classification accuracy.
Genetic Algorithm with Support Vector Machine (GA-SVM)	• Genetic algorithm has been proven to be very effective solution in a great variety of approximately optimum search problems.	• More time complexity to perform the classification. • Less classification results in terms of the time, noises in the images samples are not removed. • Timing of events is not explicitly modeled; ordering of events is preset by the analyst.

Feature Selection based Classification Methods	Merits	Demerits
Differential Evolution with Support Vector Machines (DE-SVM)	• Some are effective with respect to recall whereas some are effective with respect to Precision. • Approach is very general and can be applied for any classification problem.	• Feature extraction steps are not focused. • The weighted matrix in the DE is not exactly measured.
Support Vector Machine (SVM) and Probabilistic Neural Network (PNN)	• The classification accuracies obtained by this method is better, but SVM has performed better than PNN. • Selection of good subsets of features for improving the diagnosis rate.	• Requires a representative training set. • Noise in the images is not removed.
Exclusive Classification Techniques without Feature Selection	**Advantages**	**Major Limitations**
Single-Polarization based Classification	• Limits of this approach in discriminating water from ice.	• Detection of the thermal ice class is not highly accurate. • Classification alone performs individually.
Probabilistic Neural Network (PNN)	• An inherently parallel structure. • Guaranteed to converge to an optimal classifier as the size of the representative training set increases. • No local minima issues.	• Takes more time to perform the classification since all the features are taken into the consideration of the classification. • It also affects the

Feature Selection based Classification Methods	Merits	Demerits
	• Training samples can be added or removed without extensive retraining.	classification. performance of the system. • Large memory requirements. • Slow execution of the network.
Proposed Classifier	**Advantages**	
PNN Classifier	• Selection of optimal subsets of features from feature selection methods provides higher dimensionality reduction, which in turn improves the overall performance of classification model.	

2.8. Summary

This chapter gives the overview of all the existing methods of image preprocessing, image segmentation, feature extraction, feature subset selection and classification. With a view to improve the optimal feature subset selection and classification results, many methods have been proposed using different procedures. The inherent merits and demerits of each one of the existing methods have also been analyzed. As a detailed analysis of the existing techniques has been made systematically, it will substantially help the present research work in extending the performance of feature selection techniques from River ice image samples to improve the classification result for various ice stages. This also enables overcoming the limitations of the existing techniques, and thereby, the accuracy of final River ice type classification can be improved with optimal feature subset selection methods.