Comparative Analysis of Spatial Image Classification Techniques Using Rough-Sets

Vasundhara, D. N.[1]* and Seetha, M.[2]
[1]Department of C.S.E, VNRVJIET, Research Scholar in JNTU Hyderabad, India
E-mail: vasundhara_d@vnrvjiet.in,
[2]Department of C.S.E, G Narayanamma Institute of Technology and Science, Hyderabad, India
E-mail: smaddala2000@yahoo.com
**Corresponding Author*

Abstract

Many traditional spatial image classification techniques which are developed over past years and exists in literature like maximum likelihood, minimum distance, and parallelepiped suffer from problem of misclassification. Today, expert systems along with machine learning methods are getting universality in this area due to effective classification. This paper emphasizes on Artificial Neural Network (ANN) and "Rough Set" based "Artificial Neural Network" (RS-ANN) methods. In RS-ANN technique, Rough set (RS) is used as an attributes selection mathematical tool which excludes the redundant attributes. Further, this reduced dimensionality data set is given to Artificial Neural Network (ANN) classifier correspondingly. This process improves the classification accuracy and performance. The experiments are performed with LISS III spatial images of Hyderabad region collected from NRSC. The comparative analysis is carried out between ANN and RS-ANN techniques by considering efficiency measures like root mean square error and execution time. It is observed that RS-ANN technique outperforms the ANN technique for the spatial images with respect to accuracy and performance.

1. Introduction

Image categorization is concerned with identifying different types of images on the earth's surface, satellite collected image and spatial image and has been used for many purposes, such as urban planning and management, forestry, environ- mental monitoring, agriculture, etc., (Chang, 2007). Methodologies for classification of images are categorized into two and they are "supervised learning and unsupervised learning" strategies (Scholkopf and Smola, 2002 and Lin and Wang 2002). Supervised classification method takes an object, it is typically a vector as an input and outputs a desired value (Abonyi and Szeifert, 2003, Osuna et al., 1997 and Joachims, 1998). In unsupervised classification method, a large number of unknown pixels can be examined. It divides all pixels in form of number of different classes inside in image. The need of input in the unsupervised classification method is the number of classes.

In the image classification, the images consist of allotting a distinct class membership to individual pixel existing in the image, in this instance based process, particularly on its spectral signature (Li et al., 2014).

Many machine learning techniques are adapted to the classification task to gain the improved classification performance including accuracy (John et al., 1994 and Kryszkiewicz, and Rybinski, 1996). Rough set theory has been emerged as an innovative strategy to manage the ambiguity that is being used for the purpose of identification of data dependencies, significance factor of feature characteristics, design patterns in representation of data, attribute space dimensionality compression, and further the categorization of images (Pawlak, 1982, Petrosino and Salvi, 2006 and Cortes and Vapnik, 1995).

1.1 Related Work

There exists many techniques in literature for image classification procedure, which have been proposed in past years and shown in Table 1. Rough set phenomenon can be applied for dimensionality reduction in various domains (Pawlak, 1996). Especially its applications are so wide in machine learning areas and knowledge based decision support systems to improve the accuracy (Pawlak, 1997). Hence it warrants an investigation in this study.

Table 1: Image classification Techniques in the literature

Year and Authors	Method Used	Advantages	Limitations
Quinlan, 1996	Decision tree method	Data classification without much calculation while dealing with noisy or incomplete data.	Performance is weaker in domains with a preponderance of continuous attributes than for learning tasks that have mainly discrete attributes.
Abonyi and Szeifert, 2003	Supervised fuzzy classifier.	Individual treatment to all the pixels involved	Non-usage of training data for profile generation
Boser et al., 1992	Optimal boundary classifier	Maximizes the margin between the training patterns and the decision	Lack of transparency of results.
Goodman et al., 2002	Artificial Immune Recognition System (AIRS)	Not necessary to know what the appropriate settings for the classifier are in advance. The most important element of the classifier is self-determined.	More number of user parameters
Joachims,1998 and Osuna et al., 1997	Support Vector Machine	They prove to be very robust, eliminating the need for expensive parameter tuning.	Long training time for large datasets.
Pawlak, 1996, 1997 and1982	Rough Set Theory (RST)	Provides efficient algorithms for finding hidden patterns in data; finds minimal sets of data (data reduction);	Optimal decision rule generation from data.
Ahmed et al., 2011	Artificial Neural Network	ANN does not impose any restrictions on the input variables (like how they should be distributed)	ANN becomes vulnerable if the number of neurons in corresponding hidden layers is not appropriate.
Stanczyk, 2010	Presented the rough set based features analysis for ANN classifier	Feature Reduction	The disadvantages of the ANN and RS are rectified by each other.
Vasundhara and Seetha, 2016	Proposed Rough-Set and Artificial Neural Networks Based mechanism for image classification.	Feature Reduction of a spatial image	The disadvantages of the ANN and RS are rectified by each other for spatial image classification.

1.2 Motivation and Contribution

Image classification has been a fascinating and challenging task, which has attracted countless researchers over past years. ANN strategy is proved to be a remarkable classification strategy (Quinlan, 1996). The topological behavior of the structure and its learning procedure are the key factors in describing the performance efficiency of a particular NN. But the determination of the behavior of the structure of ANN cannot be achieved by following a single exact rule (Ahmed et al., 2011). Consequently, ANN becomes vulnerable if the number of neurons in corresponding hidden layers is not appropriate. Rough set theory, evolved in early 1980's, is kind of recent mathematical tool which is very important for the acquisition of knowledge factor and categorical assignment (Stanczyk, 2010 and Vasundhara and Seetha, 2016). This can be achieved through machine learning, inductive reasoning, systems with decision support, etc. (Goodman et al., 2002 and Boser et al., 1992). The contributions of the paper are as follows:

- Firstly, overview of ANN is presented.
- The overview of the proposed RS-ANN classification mechanism along with the experiment results are presented.
- The detailed performance and comparative analysis of the experiments with spatial images are presented.

The organization of the paper is as follows: Section 2 overviews the ANN classifier.

Proposed RS-ANN classification model along with its experimental results are discussed in section 3. Section 4 presents the experiment results and comparative analysis for both ANN and RS-ANN. The conclusions of the paper are consolidated in section 5.

2. ANN Classification

In this section a detailed overview or the ANN classification technique is presented.

2.1 Overview

The structure of a neuron: $p_1, p_2, . p_n$are inputs, $w_1, w_2, . w_n$ are weights, b *is* bias, S is the sum of weighted inputs and bias, $g$is transfer function, and 'a' is output. ANN follows back propagation scenario. Assume the sigmoid function $g(x)$ as:

$$g(x) = \frac{1}{1 + e^{-x}}$$

Equation 1

The following steps describe the steps involved in the execution of the algorithm:

1. Weights are initialized to a very low values randomly.

2. The sigmoid function is applied after calculating the sum of the weights getting into the unit and feeding the input vectors $x_0; x_1; \ldots; x_n$ along the network. Further, except the output from the input class all other outputs are set to zero viz., $d_0; d_1; \ldots; d_5$

3. Taking y_j as the actual output for each desired output from the step 2, the error for each unit is calculated using the equation" $\delta_j = y_j(1 - y_j)(d_j - y_j)$".

4. Similarly for each hidden units it is calculated using "$\delta_j = x_j(1 - x_j) \sum \delta_j w_{jk}$"

5. Each of $W_y(t + 1) = W_y(t) + \eta \delta_j x_i$ is added to the delta weights.

6. The steps from 2 to 5 are done repeatedly until the error is achieved below limit and the weights are saved to use as reference to the algorithm for recognition.

3. Proposed RS-ANN Classification

In proposed RS-ANN mechanism, RST yield the rules and further establish the vectors with most essential and critical features used for input. This basic rule in contracted format is then used for the purpose of training the spatial data bases.

3.1 Overview

The image processing model functions as under:

Image feature selection algorithm using Rough Sets:

Step 1: Since image consists of pixels. In the pre-processing phase, basic image information for each pixel (e.g. RGB values, pixel intensity values, gray code level value etc.) will be obtained. These will be called attribute values for each pixel.

Step 2: Draw the pixel, attribute-value information table. A pair$I = \{P, A\}$, in which $P\{P_1, P_2, \ldots, P_n\}$is meant for finite set of each pixel and non-empty set is called a system of information. $A\{A_1, A_2, \ldots, A_n\}$: meant for attribute set which is finite and non-empty. The process in which universal set is classified into modules or equivalence classes based on some attribute values, is called Granularization.

Step 3: In this step, the Discernibility matrix for the above pixel, attribute value information table is drawn. Here the explanation of two rough set theory terms is as below:

Discernibility: When two objects have different attribute values then they can be distinguished based on that attribute value. In RST, this is denoted as term discernibility.

Discernibility matrix: A discernibility matrix of an information table $I = (P, A)$ is a matrix which is symmetric $|P| \times |P|$ in which the entries are given by "$c_{ij} = \{a \in A | a(x_i) \neq a(x_j)\}$" $i, j = 1, .., |P|$ Individual c_{ij} consists of those features, who offers the difference between objects i and j

Step 4: For each row of the discernibility matrix, compute discernibility function as:

$$f_1, f_2, \ldots, f_{|P|} = \wedge\{\vee c_{ij} | 1 \leq j \leq i \leq |P|, c_{ij} \neq \emptyset\}$$

Equation 2

Step 5: Compute resultant discernibility function as $F = \wedge \{f_i, 1 \leq i \leq |P|\}$

Step 6: Attributes shrinking algorithm for computing $(F, \{f_i, 1 \leq i \leq |P|\})$ the algorithm steps are as below:

- While $(F, \{f_i, 1 \leq i \leq |P|\})$ are not in minimized form
- Remove the super-sets by applying Absorption Rule in $(F, \{f_i, 1 \leq i \leq |P|\}) I P)$
- Replace strong shrinkable attributes$(F, \{f_i, 1 \leq i \leq |P|\})$.

Note: strong shrink ability implements where the clause attributes are either together present or missing in all clauses. In that case, those attributes may be substituted by a single temporary attribute.

- $s \Leftarrow$ most frequent attribute: $(F, \{f_i, 1 \leq i \leq |P|\})$
- apply Expansion Rule $(s, (F, \{f_i, 1 \leq i \leq |P|\}))$
- substitute the strong shrinkable classes $(F, \{f_i, 1 \leq i \leq |P|\})$
- *RED*$\Leftarrow$ calculate the reducts $(F, \{f_i, 1 \leq i \leq |P|\})$
- return (*RED*)

Step 7: Set of entire prime implicants in computed discernibility function is obtained. In RST point of view these all are called as Reducts. In general, it is the process of feature selection in rough set theory.

Image classification using ANN: After the feature selection the network is trained by employing "back propagation mechanism" which can be utilized for image classification.

4. Experiment Results and Comparative Analysis

Here, the experimental results and comparative analysis of both ANN and RS-ANN are presented. System Specifications: Our system specifications are as below:

Software Specifications -
OS - Ubuntu 16.04 LTS, 64 bit
R Language Environment version 3.3.2 R Studio IDE 1.0.44
Java version: '1.8.0 111'
NetBeans product version: NetBeans IDE 8.2
Hardware Specifications -
RAM size - 4 GB
Processor – "Intel core i3 4030U CPU @1.90GHz × 4"

4.1 ANN Procedure Experimental Details

Experiment details of ANN procedure are as follows:

Input Dataset: Image classification using ANN is performed on the spatial LISS III images of Hyderabad, Telangana State region, which are acquired from NRSC and shown in Figures 1 to 3.

Output Details: After executing the ANN procedure on the images shown in figures 1 to 3, the output images are shown in Figures 4 to 6.

The accuracy and performance of the classification using ANN is measured with the Mean Squared Error (MSE) and Execution Time (ET) and the results are tabulated in Table 2.

Table 2: Accuracy and Performance using ANN

	MSE (%)	ET(Sec)
Image 1	13.99	80.40
Image 2	14.79	86.14
Image 3	14.34	85.31

4.2 RS-ANN Procedure Experimental Details

Experiment details of RS-ANN procedure are summarized as follows:

Procedure: In this procedure, Java is utilized for RS features selection. It gives the list of selected and extracted features. Further new file is produced with features are extracted. This file is utilized as input to our ANN classifier, finally classification is performed on the same images.

Output Details: After executing the RS-ANN procedure on the images shown in Figures 1 to 3, the output image is shown in Figures 7 to 9. By observing the spatial patterns of Figures 4, 5 and 6 with corresponding Figures 7, 8 and 9 respectively, it can be concluded that the spatial patterns are clearer thus identifying the discernible features, decreasing the mean square error and the execution time in the Figures 7 to 9 which obviously confirms the efficacy of the RS-ANN Technique. The accuracy and performance of the classification using RS-ANN is measured with the Mean Squared Error (MSE) and Execution Time (ET) and the results are tabulated in Table 3.

Table 3: Accuracy and Performance using RS-ANN

	MSE (%)	ET(Sec)
Image 1	12.51	36.80
Image 2	13.23	38.23
Image 3	12.83	37.56

Figure 1 : LISS-III Original Image 1

Figure 2: LISS-III Original Image 2

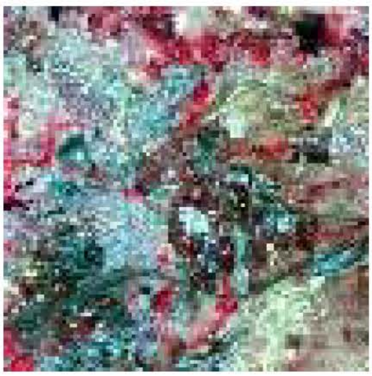

Figure 3 : LISS-III Original Image 3

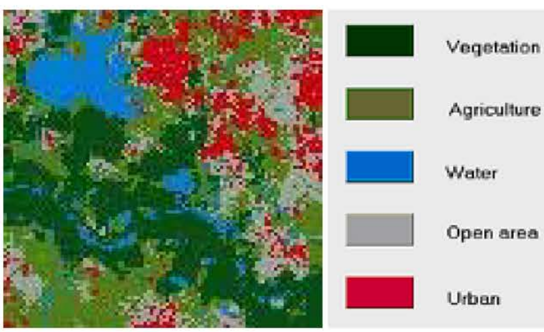

Figure 4: Classified Image of Figure 1 using ANN

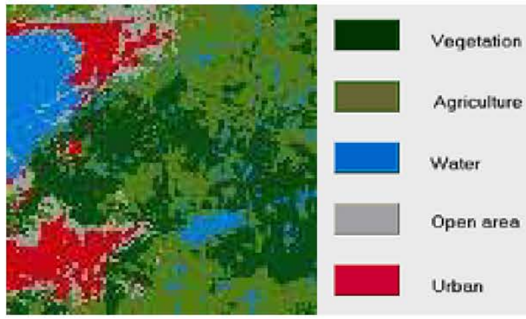

Figure 5: Classified Image of Figure 2 using ANN

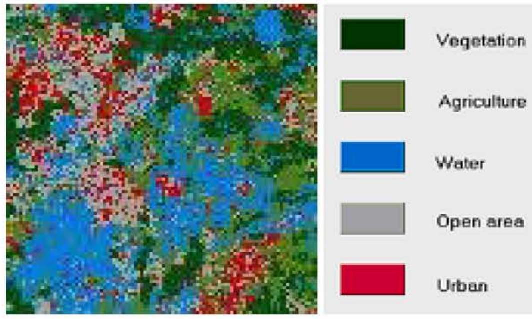

Figure 6: Classified Image of Figure 3 using ANN

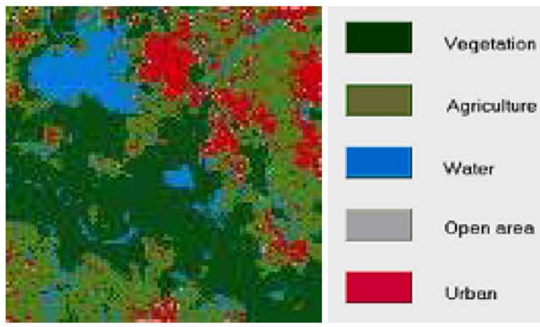

Figure 7: Classified Image of figure 1 using RS-ANN

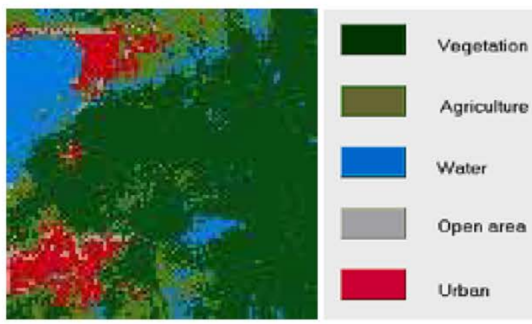

Figure 8: Classified Image of figure 2 using RS-ANN

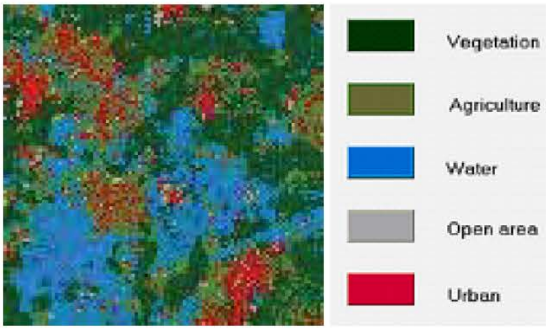

Figure 9: Classified Image of figure 3 using RS-ANN

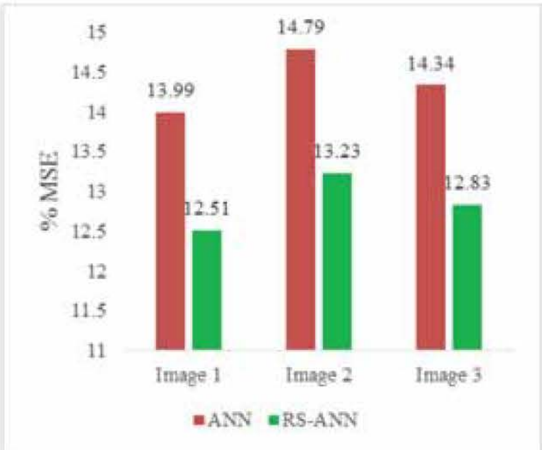

Figure 10: Comparative analysis of MSE% for ANN and RS-ANN techniques for input images

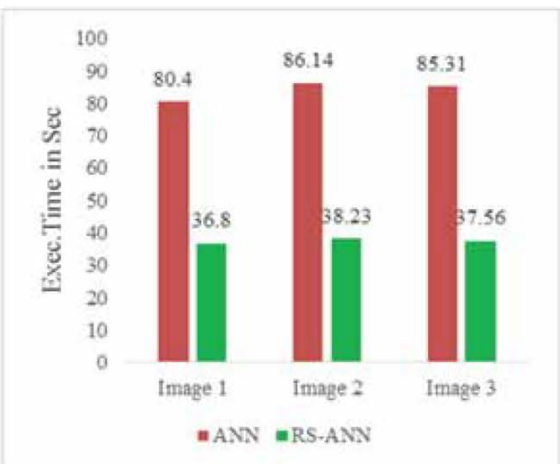

Figure 11: Comparative analysis of Execution Time for ANN and RS-ANN techniques for input images

4.3 Comparative Analysis

The experiment is performed on both - ANN and RS-ANN procedures by utilizing NRSC LISS III images of Hyderabad, Telangana State.

The comparative results are shown in Figures 4 and 5. From the Figures 10 and 11 the increase in the performance with the proposed RS-ANN classification technique is calculated and tabulated in the Table3. From Table 4 it can be concluded that the proposed RS-ANN technique is achieved significant increase in the accuracy viz., around 10.5% for all images and around 55% prominent decrease in the execution time.

Table 4: Increase in the performance using RS-ANN Technique

	ANN		RS-ANN		% Change	
	MSE (%)	ET (Sec)	MSE (%)	ET (Sec)	% Decrease in MSE	% Decrease in ET
Image 1	13.99	80.40	12.51	36.80	10.58	54.23
Image 2	14.79	86.14	13.23	38.23	10.55	55.62
Image 3	14.34	85.31	12.83	37.56	10.53	55.97

Hence the proposed RS-ANN Technique is not only accurate but also faster and can be executed in very less time when compared to ANN. From Table 4 it can be concluded that the proposed RS-ANN technique is achieved significant increase in the accuracy viz., around 10.5% for all images and around 55% prominent decrease in the execution time. Hence the proposed RS-ANN Technique is not only accurate but also faster and can be executed in very less time when compared to ANN.

5. Conclusions

The effectiveness and accuracy of classification task plays a vital role in the analysis of spatial image classification techniques. This paper elucidates the experimental analysis using the Rough Set based Artificial Neural Networks (RS-ANN) method with ANN technique for classification on LISS III images from NRSC. Rough set (RS) theory is an effective mathematical tool for feature selection to eliminate the redundancy. Further, Artificial Neural Networks (ANN) is successfully employed for classification of spatial images to reduce the misclassification. It is observed that the proposed RS-ANN is successfully implemented with 10.50% increase in the accuracy and 55% increase in the execution time. Hence, it is concluded that RS-ANN technique outperforms the ANN technique for the spatial images with respect to accuracy and performance. The work can be extended further for other accuracy measures and to identify the most appropriate features.

References

Abonyi, J. and Szeifert, F, 2003, Supervised Fuzzy Clustering for the Identification of Fuzzy Classifiers. *Pattern Recognition Letters*, Vol., 24(14). 2195-2207.

Ahmed, S. A., Dey, S. and Sarma, K. K. 2011, Image Texture Classification Using Artificial Neural Network (ANN). Emerging Trends and Applications in Computer Science (NCETACS), *2011 2nd National Conference on IEEE*. doi: 10.1109/NCETACS.2011.5751383.

Boser, B. E., Guyon, I. M. and Vapnik, V. N., 1992, A Training Algorithm for Optimal Margin Classifiers. COLT '92 Proceedings of the Fifth Annual Workshop on Computational Learning Theory. 144-152, doi:10.1145/130385.130401

Chang, C. I., 2007, *Hyperspectral Data Exploitation: Theory and Applications*. New York, NY, USA: Wiley. doi:10.1002/0470124628

Cortes, C. and Vapnik, V., 1995, Support-Vector Networks. *Machine Learning*, Vol. 20(3), 273-297.

Goodman, D. E, Boggess, L. C. and Watkins, A. B., 2002, Artificial Immune System Classification of Multiple-Class Problems. Intelligent Engineering Systems through Artificial Neural Networks, Fuzzy Logic, and Evolutionary Programming Complex Systems and Artificial Life, Vol. 12, 179-184.

Joachims, T., 1998, Text Categorization with Support Vector Machines: Learning with Many Relevant. *In Proceedings of the 10th European Conference on Machine Learning*, 137-142.

John, G. H., Kohavi, R. and Pfleger, K., 1994, Irrelevant Features and the Subset Selection Problem. *In Proceedings of the 11th International Conference on Machine Learning*. 121-129, https://doi.org/10.1016/B978-1-55860-335-6.50023-4.

Li, D., Wang, J., Zhao, X., Liu, Y. and Wang, D., 2014, Multiple Kernel-Based Multi-Instance Learning Algorithm for Image Classification. *Journal of Visual Communication and Image Representation*, Vol. 25(5), 1112-1117. https://doi.org/10.1016/j.jvcir.2014.03.011.

Lin, C. F. and Wang, S. D., 2002, Fuzzy Support Vector Machines. *IEEE Transactions on Neural Networks*, Vol. 13(2), 464 - 471. doi: 10.1109/72.991432.

Kryszkiewicz, M. and Rybinski, H., 1996, Attribute Reduction versus Property Reduction. *In Proceedings of the Fourth European Congress on Intelligent Techniques and Soft Computing*, 204-208.

Osuna, E., Freund, R. and Girosi, F, 1997, Training Support Vector Machines: Application to Face Detection. *In Proceedings of Computer Vision and Pattern Recognition, Puerto Rico*, 130-136. doi: 10.1109/CVPR.1997.609310.

Pawlak, Z, 1996, Why Rough Sets. IEEE International Conference on Fuzzy Systems. Vol. 2, 738 – 743.

Pawlak, Z, 1997, Rough Set Approach to Knowledge-Based Decision Support. *European Journal of Operational Research*, Vol. 99(1), 48-57.

Pawlak, Z., 1982, Roughsets. *International Journal of Computer and information Sciences*, Vol. 11(5), 341-356.

Petrosino, A. and Salvi, G, 2006, Rough Fuzzy Set Based Scale Space Transforms and their use in Image Analysis. *International Journal of Approximate Reasoning*, Vol. 41, 212-228. https://doi.org/10.1016/j.ijar.2005.06.015.

Quinlan, J. R., 1996, Improved use of Continuous Attributes in C 4.5. *Journal of Artificial Intelligence Research*, Vol. 4(1), 77-90.

Scholkopf, B. and Smola, A. J, 2002, Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond. The MIT Press.

Stanczyk, U., 2010, Rough Set-Based Analysis of Characteristic Features for ANN Classifier, Hybrid Artificial Intelligence Systems, doi: 10.1007/978-3- 642-13769-3 69, 565-572.

Vasundhara D. N. and Seetha M., 2016, Rough-Set and Artificial Neural Networks Based Image Classification, *2nd International Conference on Contemporary Computing and Informatics* (IC3I), Noida, 2016, 35-39. doi: 10.1109/IC3I.2016.7917931.