



Classification of Synthetic Aperture Radar Images in Urban Areas with the Combination of FCM and SVM

Ahmad Reza Aghajani Fesharaki¹, Mehran Emadi²

Master Student, Dept. of Electrical Engineering, ACECR Institute of Higher Education (Isfahan Branch), Isfahan, Iran¹

Assistant Professor, Dept. of Electrical Engineering, Mobarakeh Branch, Islamic Azad University, Mobarakeh, Iran²

ABSTRACT: A popular way of providing land coverage maps of the Earth's surface is classifying remote sensing images. Synthetic aperture radar data are used more frequently compared to other types of data, since they are independent of sunlight and highly focus on atmosphere conditions. In this study, two sets of data were used for algorithm implementation and results evaluation. The data include simulated and real data. Algorithm is first implemented on simulated data and results are evaluated, next, the real data undergo the same process. Both data sets have undergone different classifiers implementation with regard to features, land coverage maps have been provided, and results evaluated and compared. To classify simulated data, two images and their reference land are used as training data and test data. K-nearest neighbour and support vector machine have been used for this purpose. Ultimately, with regard to their reference land results are evaluated. This method's accuracy for simulated data, not taking GLCM tissue feature into account, was 98.76%, and 99.24% when taking GLCM tissue feature into account.

KEYWORDS: GLCM tissue, tissue feature, vector machine, K classifier

I. INTRODUCTION

Classifying land coverage especially in urban areas is one of the leading synthetic aperture radar and pre-processes in automatic detection of land cover classification of remote sensing images.

Further, mapping, meteorology, identifying targets, and environmental scanning also need satellite images classification. Both supervised and unsupervised classification methods are used for remote sensing images. Providing accurate and up-to-date land coverage information is quite vital in urban management.

SAR can assist urban management in gaining accurate information, as a modern method capable of shooting in different atmosphere conditions and in day and night.

Some related previous studies and methods are:

Qi et al. (2012) adopted 66 features of Radarsat-2 satellite images for land coverage classification in object-oriented analysis, and their results showed improved classification accuracy compared to common methods.

Maghsodi et al. (2012) suggested non-parametric and class-based feature selection methods for non-parametric and class-based feature selection from among 58 primary parameters in forest classification. They showed that feature selection significantly increases classification accuracy compared to Wishart classifier and SVM with complete feature set, but has not been used here because of urban environment complexity.

Maghsodi et al. provided a new method for Hotelling-Lawley as a new test statistic for change detection in multi-perspective polarimetric radar images in 2014. In this method, text classification has been presented for classifying polarimetric radar images with synthetic aperture. Ultimately, this method showed 16%, 11%, and 7% accuracy improvement compared to Wishart, Wishart-Markov, and SVM, respectively.

Maghsodi et al. in 2015 proposed a method for improving classification of urban areas using radar polarimetric data and multi-purpose improvement methods. In the method proposed NSGA-II multi-purpose improvement algorithm is used as the search tool, and two-class support vector machine SVM and ANFIS neuro-fuzzy inference system in the evaluation stage.

In the method proposed in this article, SVM classifier and K-nearest neighbor have been used for classifying land coverage. The accuracy of the method for simulated data, not taking GLCM tissue feature into account, was 98.76%,

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

and 99.24% when taking GLCM tissue feature into account. Classification accuracy for Quebec real data, regardless of tissue feature was 85.12%, and 89.46% with the tissue features considered. The classification accuracy for San Francisco Bay real data, regardless of texture was 91.76 percent, and 93.46 percent were taking tissue feature into account. Results show employing GLCM tissue feature increases the classifier accuracy.

II. METHODOLOGY

Introducing the flowchart proposed:

As mentioned above, two sets of data are used in this study for algorithm implementation and results evaluation. To classify real data three SAR images are used; one as the training data and the other two as test data. Like simulated data, KNN and SVM classifiers, and to train them first and second type features have been used. Since these data lack reference land, an FCM clustering algorithm has been used to cluster training data, and a number of samples have been chosen from each cluster for classifier training. After training the classifier, two images on evaluation are given to the classifier and classified. In the end, results of this classification will be compared to those of its FCM classifier and estimate classifier accuracy on real SAR data. It is vital to note that since FCM results have been used instead of reference land, in the evaluation phase the same results should be used to estimate classifier accuracy. Methodology flowchart is shown in Fig. 1.

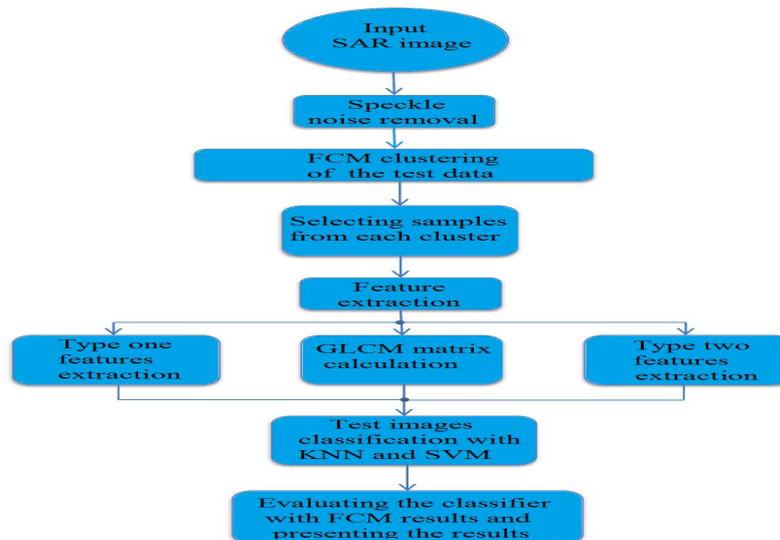


Figure 1: methodology flowchart

Introducing Fuzzy C-means(FCM) algorithm

FCM is an accepted method in image classification due to its great features for recognizing clusters and data security, missing in hard classification methods like k-means. FCM attaches pixels to classes using fuzzy membership function. Assume $X=(x_1, x_2, \dots, x_n)$ is an image with n pixels that want to be clustered in C clusters. Function J is the objective function defined for FCM.

$$J = \sum_{j=1}^n \sum_{i=1}^c u_{ij}^m \|x_j - \mu_i\|^2$$

In which u_{ij} is the membership level of pixel x_i from the i th cluster and m is the fuzzifier determining the fuzziness in each part, located in $[1 - \infty)$. Membership function and category centers in each level are updated through the following:



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

$$2: u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_j - \mu_i\|}{\|x_j - \mu_k\|} \right)^{\frac{1}{(m-1)}}}$$

$$3: \mu_i = \frac{\sum_{j=1}^N u_{ij}^m x_j}{\sum_{j=1}^N u_{ij}^m}$$

FCM is sensitive to noise and other image artifacts. Further, if classes are assigned to each pixel based on the light intensity distribution, the algorithm will be extremely sensitive to light intensity changes. To solve this, some pre-processes are done to make the image smoother.

Image classification:

Here, through extracting proper statistic features and using proper classifiers, the area's images are classified and the urban areas detected. In this project SVM classifier and k-nearest neighbor are used. Features used to train these classifications are divided into two groups of type one and type two features.

Feature extraction:

Type-one features:

Type one features show the level of similarity of a gray degree. Type one statistical features can be extracted using brightness intensity histogram. These features are gained through some mathematical calculations on the image. Type one features provide information about manner of gray levels distribution, but not the status of gray levels to each other. Mean, variance and standard deviation are used type one features.

Type-two features:

Type-two features are expressed as texture features. Texture features provide important information about the structural status of an image parts and their relationship with other parts. Image texture and how to use its features are discussed further.

Classifier:

SVM classifier:

This classifier is a method for detection and identification of data based on cloud plate boundary, so that the cloud has the greatest distance from the adjacent data. Assuming linear separability of the clusters, the cloud acquires a page with maximum margin to separate clusters. In cases where data may not have linear separability, mapping the data into a bigger space can help separate them linearly in this new space with bigger dimensions. Generally, SVM is a two-class classifier. For multi-class problems, the problem is divided into a number of two-class problems. Then the outputs are combined and a multi-class problem is solved. For this purpose a combination of one for one, and one against the rest classifiers are used.

KNN classifier:

One of the simplest types of classifiers is KNN. If the class of feature vector f is to be determined, first, the classes of K to the nearest vectors training data to vector f are considered, and then the class with the highest number of dependent vectors is chosen as the class depending on f .

Results evaluation

In this section results of KNN and SVM classifiers with different kernels, taking into account the type I and type II features (tissue features) for two sets of simulated and real data are presented. Energy statistical parameters, contrast, homogeneity, and correlation have been used for this purpose as type II features.

Simulated data:

In this study, two simulated images with dimensions of 256×256 are used. Since SAR images contain speckle noise, speckle noise with specific signal to noise ratio were added to the images. These images include three classes. In Figure 3 SAR simulation data has been shown to with the signal to noise ratio of 20 decibels (db), train classifier with the reference land.

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

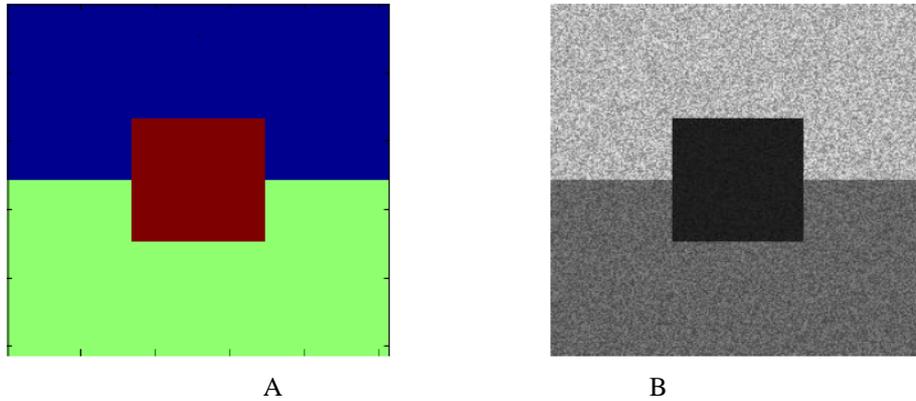


Figure 3: simulated SAR image to train the classifier. A: SAR image, B: land

In Figure 4 SAR simulation data has been shown with the signal to noise ratio of 20 decibels (db), o train classifier with the reference land.

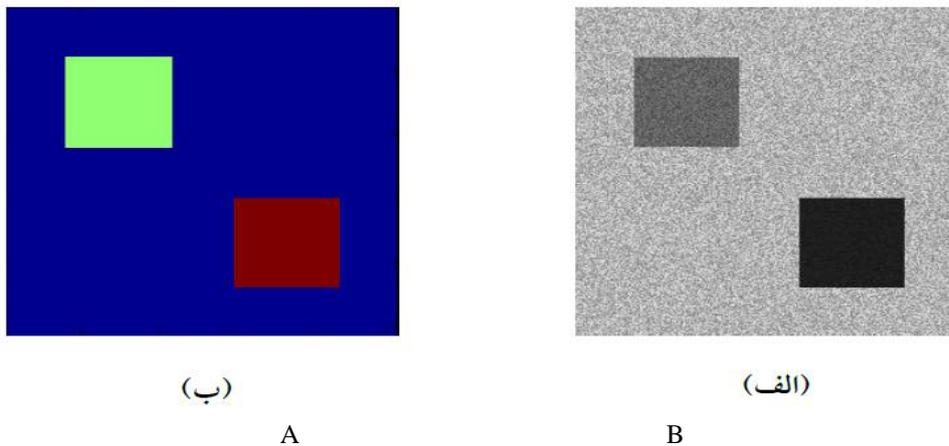


Figure 4: simulated SAR image to train the classifier. A: SAR image, B: reference land

Results of simulated image classification without the use of tissue features

In this section, with the training data reference land availability, some samples are extracted from each class as training samples to train the classifier. 5% of class 1, 5% of class 2, and 20% of class 3 are used. Then the system is trained using type I features that are mean, variance, and standard deviation. Next, evaluation data are classified using this system, and with the available land use, classification accuracy is measured. KNN and SVM classifiers with different kernels are used. Since the addition of speckle noise to simulated images is random and by each time adding noise to the picture, a different picture is presented, the algorithm is run ten times under SNR = 20db, and the average of results is calculated.

The accuracy of these classifiers is calculated by comparing their results with reference land. The results are arranged in Table 2-3.

	SVM-Linear	SVM-RBF	SVM-Polynomial	SVM-Quadratic	KNN
Accuracy	98/7097	98/7691	98/5402	98/7492	97/2641

Table 2-3: Classification accuracy of simulated data without the use of tissue features

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

Diagram of classification accuracy of simulated data without the use of tissue features is shown in Figure 5.

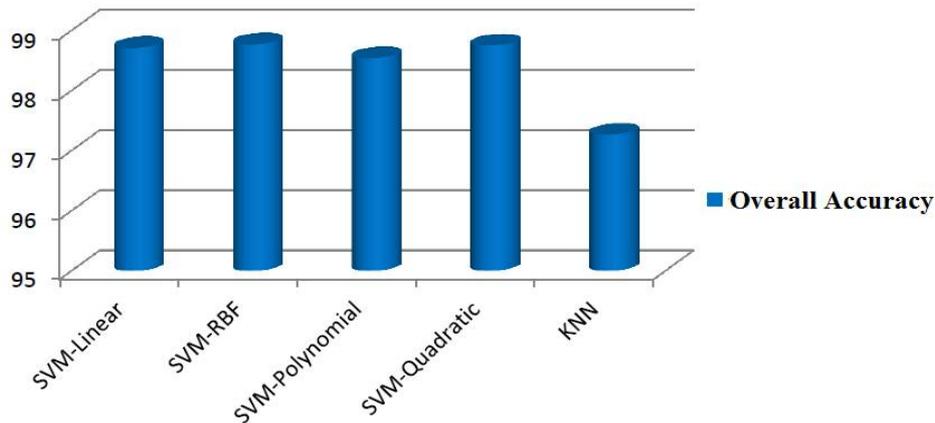


Figure 5: Diagram of classification accuracy of simulated data without the use of tissue features

As demonstrated, SVM classifier with RBF kernel function has achieved the best accuracy in this case. The resulting classification map is shown in Figure 6.

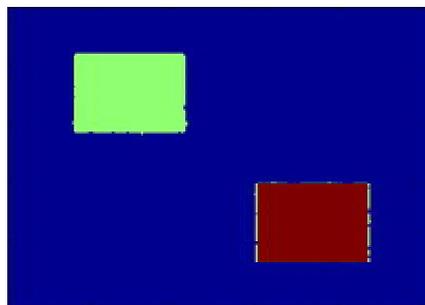


Figure 6: Map of SVM classifier with RBF kernel function without the use of tissue features for simulated data

Results of simulated image classification with tissue feature taken into account

Here, tissue features are also used for system training. Like before, 5% of class 1, 5% of class 2, and 20% of class 3 are used. Both types of features are used to train the system. Then the evaluation data are classified using this system, and with the available land reference classification accuracy is measured. Energy statistical parameters, contrast, homogeneity, and correlation have been used as features. These features were calculated using GLCM matrix for each window. Tissue extraction window size is considered 5×5 , vector distance with a radius of 1 and direction of 0 and 180, and the gray levels in the third area. KNN and SVM classifiers with different kernels are used here too. The algorithm is run ten times under SNR = 20db and the mean of results is calculated. The accuracy of these classifiers is calculated through comparing their results with reference land. The results are arranged in Table 2-4.

	SVM-Linear	SVM-RBF	SVM-Polynomial	SVM-Quadratic	KNN
Accuracy	99/2149	98/9116	99/2056	99/2438	99/0128

Table 2-4: simulated image classification accuracy with tissue feature taken into account

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

Diagram of classification accuracy of simulated data with the tissue feature taken into account is shown in Figure 7. Map of SVM classifier with linear kernel with tissue feature taken into account for simulated data is shown in Figure 8.

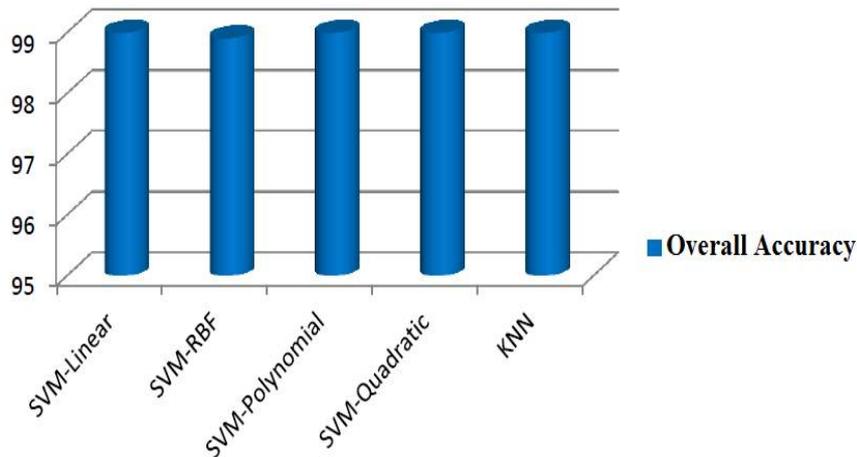


Figure 7: Diagram of classification accuracy of simulated data with the tissue feature taken into account

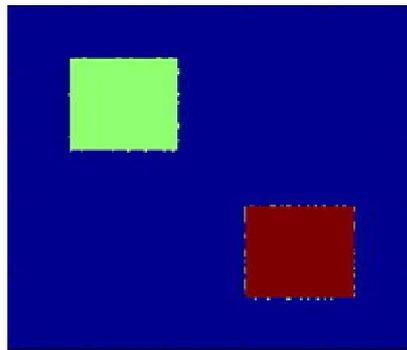


Figure 8: Map of SVM classifier with linear kernel with tissue feature taken into account for simulated data. In figure 9 classifiers accuracy with and without the features are compared. As demonstrated, GLCM tissue features have improved accuracy.

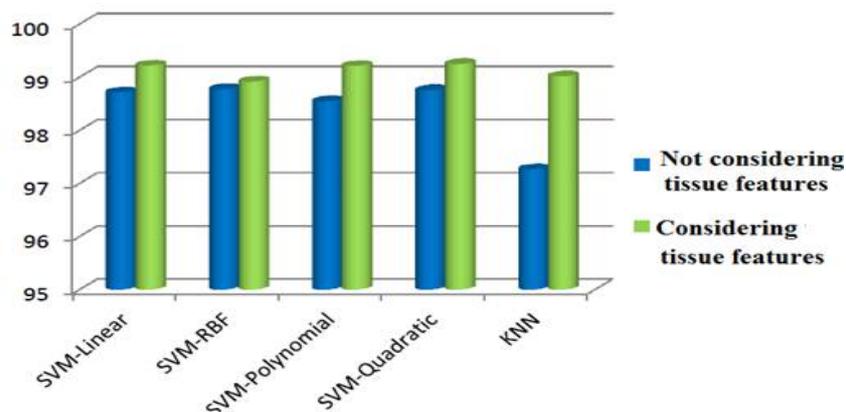


Figure 9: comparison of classification accuracy with and without the features for simulated data

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

Real data:

In this study, three SAR real data available on earth.esa.int website have been used. The data were extracted using PolSARpro application on the website. Training data is from Niigata urban area in Japan. The image is 1200x1200 pixels in terms of dimensions and contains urban areas, land coverage, and river. The related Pauli image is shown in Figure 10.

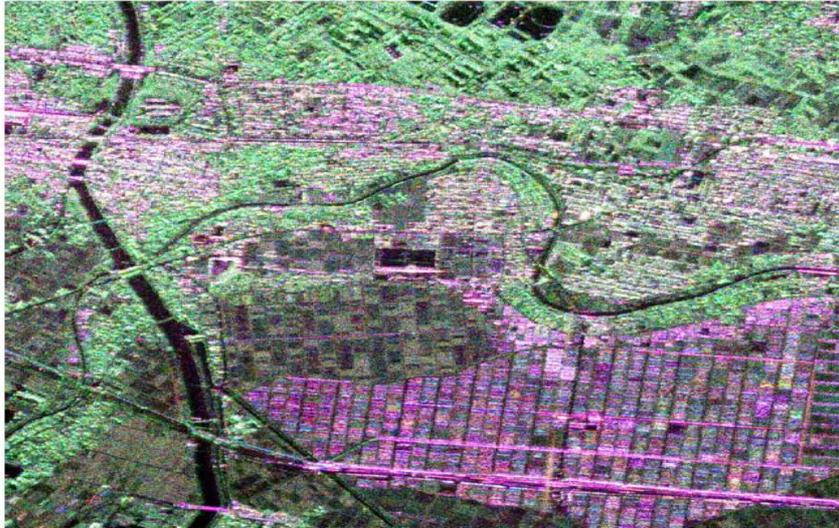


Figure 10: Pauli image of Niigata

The first test data is from Quebec in Canada, whose Pauli image is shown in Figure 11.

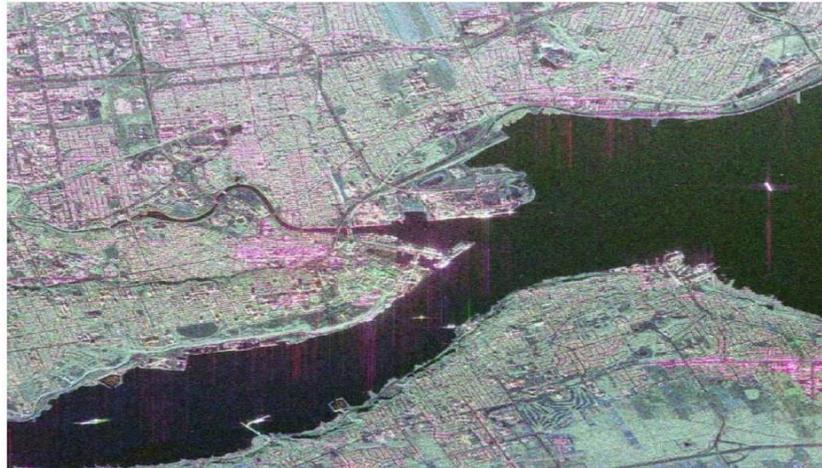


Figure 11: Pauli image of Quebec

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016



Figure 12: Pauli image of San Francisco Bay

The other test data used here was of San Francisco Bay. Its Pauli image is shown in Figure 12. Due to absence of reference land in these data, training data are clustered using FCM at first. The data are divided into three separate clusters and each cluster gets a label that shows the land coverage (water, plant, and city). Then some samples are selected from each cluster.

5% of class 1 (water), 1% of class 2 (plant) and 5% of class 3 (city) are used. Further, two images of evaluation are given to the system and classified. In the end, results of classification are compared to the results of FCM clustering and the classifier accuracy on real SAR data is calculated.

Since FCM results are used for classifier training, the classifiers FCM results are used for accuracy estimation.

Results of real image classification without use of tissue features:

In this section, only type I features are used for classifier training. Different classifiers with different kernels have been used to this end. Land coverage accuracy of these classifiers is estimated through comparing them with FCM image clustering results.

	SVM-Linear	SVM-RBF	SVM-Polynomial	SVM-Quadratic	KNN
Class 1 (Water)	91/8151	91/8810	91/7204	91/5575	94/7696
Class 2 (Brunette)	84/2960	85/1502	85/4413	85/4084	70/3331
Class 3 (City)	80/6695	80/1440	80/1800	80/2746	83/1518
Total	84/9559	85/0541	85/1209	85/1019	82/0800

Table 2-5: Quebec image classification accuracy without use of tissue features

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

	SVM-Linear	SVM-RBF	SVM-Polynomial	SVM-Quadratic	KNN
Class 1 (Water)	97/3600	97/3600	97/2948	97/1247	98/4675
Class 2 (Brunette)	82/4497	82/6785	81/6600	82/2078	71/8298
Class 3 (City)	92/8290	92/7386	93/5748	93/3139	92/4632
Total	91/7183	91/7450	91/7650	91/7509	89/1710

Table 2-6: San Francisco image classification accuracy without use of tissue features

As observed, SVM classifier with degree 3 polynomial core function has gained the highest accuracy for both Quebec and San Francisco Bay images.

The land coverage map provided by this classifier is shown in Figures 15 and 16.

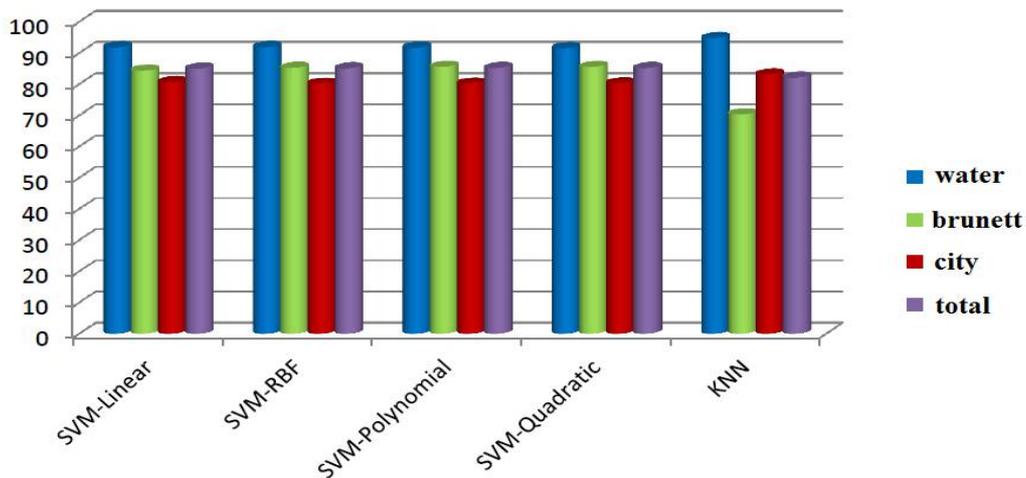


Figure 13: Quebec image classification accuracy diagram without use of tissue features

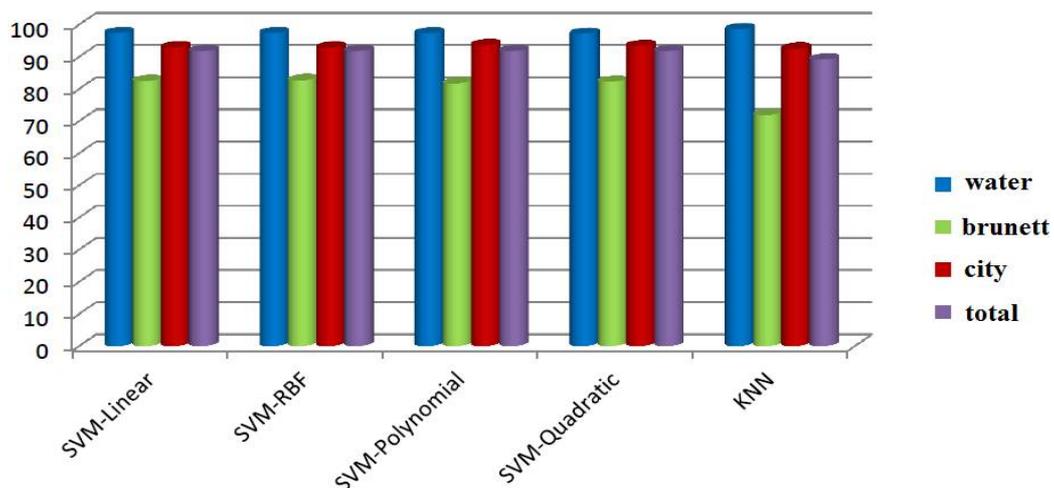


Figure 14: San Francisco image classification accuracy diagram without use of tissue features

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

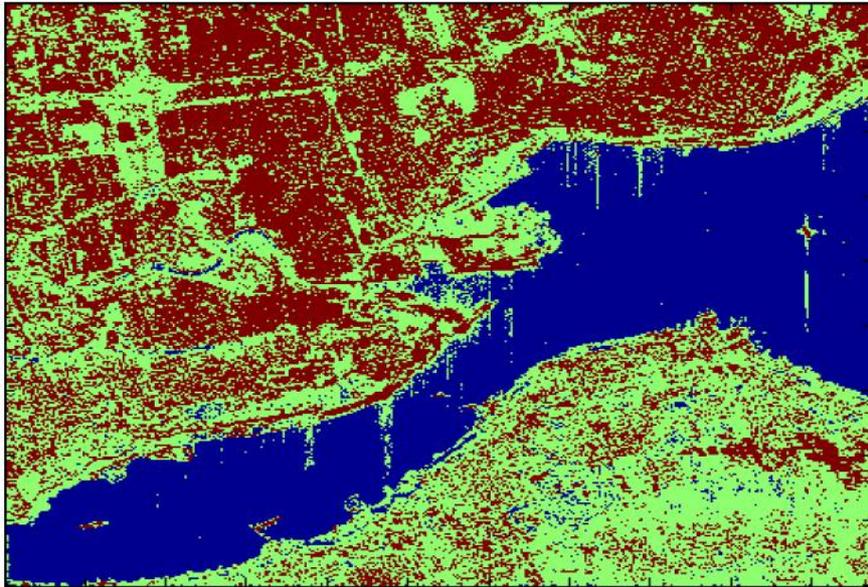


Figure 15: SVM classifier land cover map with the nucleus of a polynomial function of the third degree without the use of texture features for Quebec image

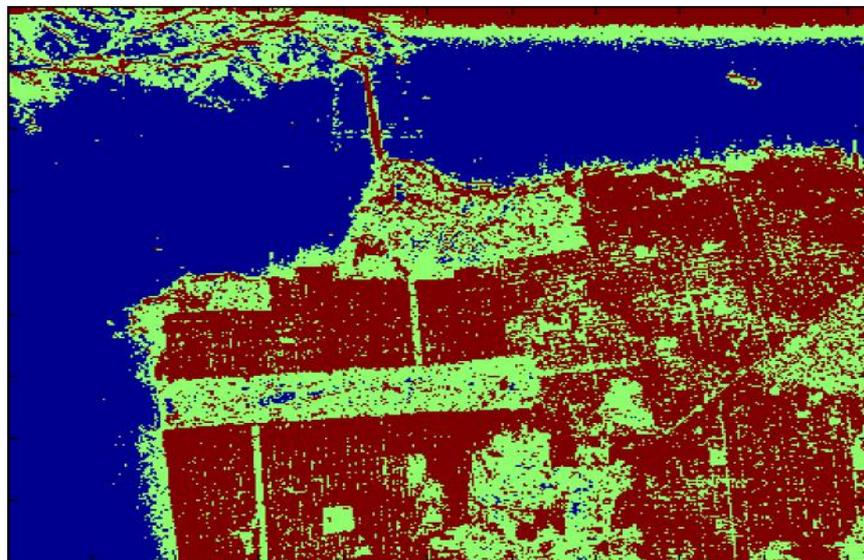


Figure 16: SVM classifier land cover map with the nucleus of a polynomial function of the third degree without the use of texture features for San Francisco image

Results of real image classification with tissue feature taken into account

Here, tissue features are used for system training alongside type I features. Energy statistical parameters, contrast, homogeneity, and correlation have been used as features. These features were calculated using GLCM matrix for each window. Tissue extraction window size is considered 5×5 , vector distance with a radius of 1 and direction of 0 and 180, and the gray levels in 8 areas.

KNN and SVM classifiers with different kernels are also used. The land coverage accuracy of these classifiers is calculated through comparing them with FCM clustering results.

Table 2-7 and Table 2-8, respectively, show results for Quebec and San Francisco. Accuracy diagram of the three classes and total accuracy are shown in Figures 17 and 18.

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

	SVM-Linear	SVM-RBF	SVM-Polynomial	SVM-Quadratic	KNN
Class 1 (Water)	93/5487	92/5686	89/5868	91/1496	93/8921
Class 2 (Brunette)	89/8377	90/3106	87/7482	90/1455	87/1511
Class 3 (City)	86/2584	86/2205	83/5570	87/3856	85/3876
Total	89/4650	89/3376	86/6199	89/3455	88/3233

Table 2-7: Quebec image classification accuracy with tissue feature taken into account

	SVM-Linear	SVM-RBF	SVM-Polynomial	SVM-Quadratic	KNN
Class 1 (Water)	98/4032	97/3134	91/3305	96/2417	98/2842
Class 2 (Brunette)	82/4001	84/4008	82/7573	81/4541	84/0887
Class 3 (City)	96/5090	96/1096	88/8731	97/6626	93/2194
Total	93/4594	93/5541	89/2250	92/8606	92/6318

Table 2-8: San Francisco image classification accuracy with tissue feature taken into account

As demonstrated, SVM classifier with linear kernel has achieved the highest accuracy for Quebec and SVM with RBF kernel function for San Francisco. The resulting coverage map is shown in Figures 17 and 18.

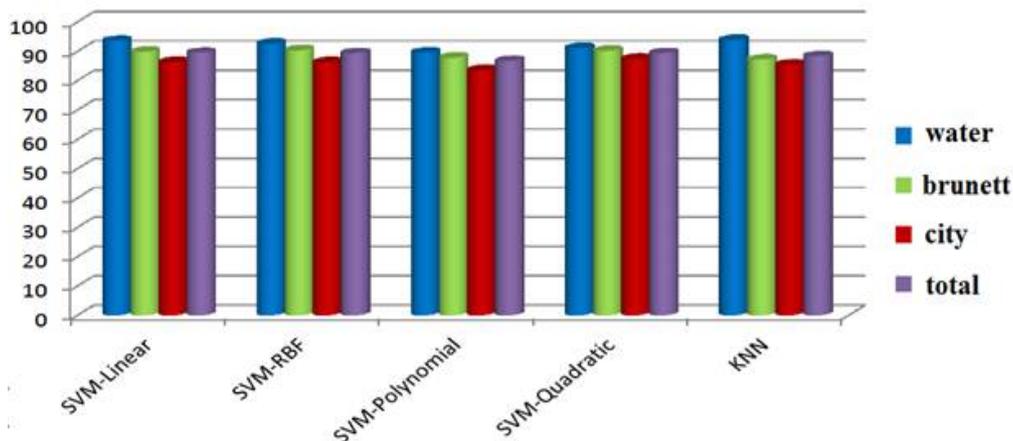


Figure 17: Quebec image classification accuracy diagram with tissue feature taken into account

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

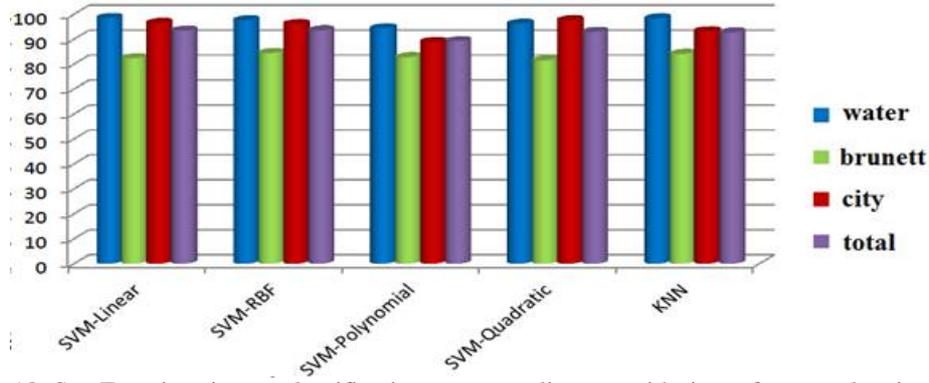


Figure 18: San Francisco image classification accuracy diagram with tissue feature taken into account

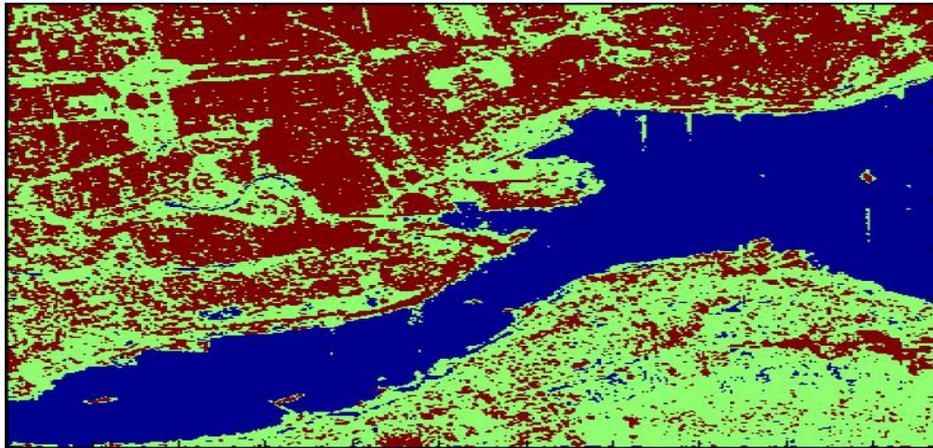


Figure 19: SVM classifier land coverage map with linear kernel, taking into account the tissue features for Quebec image

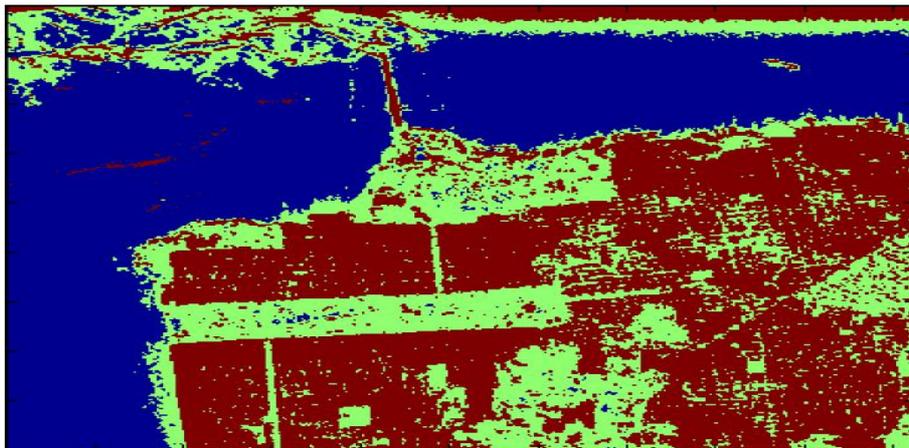


Figure 20: SVM classifier land coverage map with RBF kernel function, taking into account the tissue features for San Francisco

Total accuracy of these classifiers with and without the features are compared and results are shown in Figures 21 and 22.

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

As demonstrated, GLCM tissue features have improved accuracy and maps enjoy higher consistency in showing land coverage.

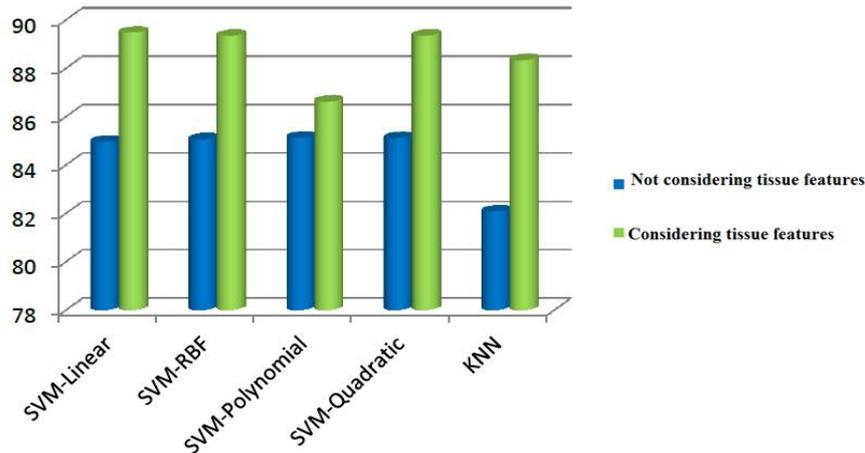


Figure 21: Total accuracy of classification with and without the features for Quebec image

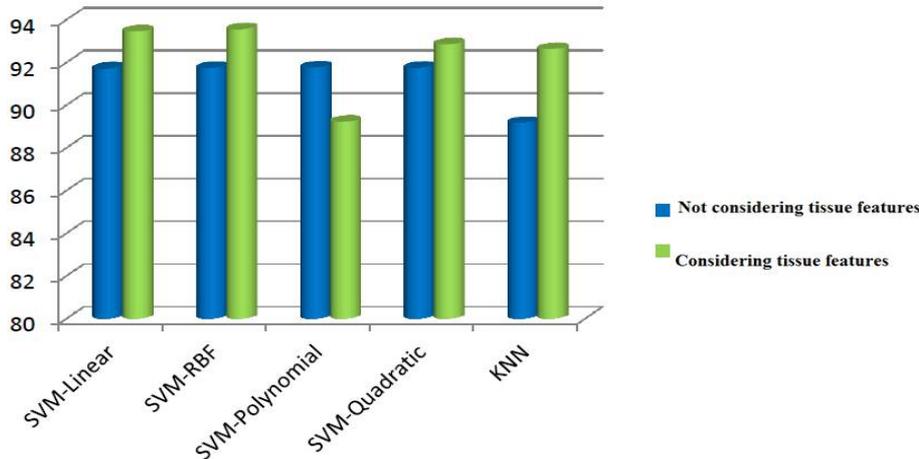


Figure 22: Total accuracy of classification with and without the features for San Francisco image

III. CONCLUSION

In this study performance of the suggested GLCM tissue feature extraction method for classification was evaluated. To do so, the two simulated SAR and real SAR data sets were used.

The two simulated data were used as one training data and one test data. Speckle noise of 20 db was added to them. With their land reference available, some samples of training data were used to train the classifier and then tested on evaluation data. Since the addition of speckle noise to images is random, the algorithm is run ten times and the mean of results was calculated. Results showed that using GLCM tissue feature lead to accuracy improvement.

Three real SAR images were also used to evaluate the suggested method's performance on real data. One of these images was used to train the classifier and the other two for evaluation. Due to absence of reference land in these data, training data are clustered using FCM at first and then samples are chosen.

After sample selection, the classifier was trained with type one and two features (tissue features) and tested on test images.

To evaluate the accuracy of the suggested method the land coverage maps were evaluated with their FCM results, since the system was trained with FCM clustering results in the training phase. For real SAR data, too, results showed improved land coverage map accuracy using GLCM tissue features.



ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 8, August 2016

REFERENCES

- [1]W. Wu , H. Guo , X. Li, and L. Zhang “Urban Land Use Information Extraction Using the Ultrahigh-Resolution Chinese Airborne SAR Imagery” IEEE Transaction on Geoscience and Remote Sensing, vol. 53, no. 10, Octobr 2015.
- [2]W. Wu, H. Guo and X. Li, “Urban Area SAR Image Man-Made Target Extraction Based on the Product Model and the Time–Frequency Analsi” IEEE Journal of Selected Topics in Applied Earth Observations and Rmote Sensing, vol. 8,no. 3, March 2015.
- [3]X. Cheng, W. Huang, and J. Gong, “An Unsupervised Scattering Mechanism Classification Method for PolSAR Image”, IEEE Geoscience and Remote Sensing Letter, vol. 11, NO. 10, Octobr 2014.
- [4]Koray Kayabol, “Histogram-Based Contextual Classification of SAR Images”, IEEE Geoscience AND Remote Sensing Letter, vol. 12, NO. 1, January 2015.
- [5]M. Jafari, Y. Maghsoudi, and M. J. Valadan Zoej, “A New Method for Land Cover Characterization and Classification of Polarimetric SAR Data Using Polarimetric Signatures”, IEEE Journal OF Selected Topics in Applied Earth Observations and Rmote Sensing, vol.8, July 2015 .
- [6]M. Salehi, M. R. Sahebi and Y.Maghsoudi, “Improving the Accuracy of Urban Land Cover Classification Using Radarsat-2 PolSAR Data,” IEEE Journal of Selected Topics in Applied Earth Observations and Rmote Sensing, vol. 7, no. 4, Apr 2014.
- [7]N.D .Cagatay and M.Datcu, “FrFT-Based Scene Classification of Phase-Gradient InSAR Images and Effective Baseline Dependence” IEEE Geoscience and Remote Sensing Society, vol. 12, May 2015.
- [8]B.Hou, B.Ren,G. Ju and H,Li “SAR Image Classification via Hierarchical Sparse Representation and Multisize Patch Features”, IEEE Geoscience and Remote Sensing Society, Vol.13 , Jan. 2016.
- [9]J. Yin and J. Yang “Novel Model-Based Method for Identification of Scattering Mechanisms in Polarimetric SAR Data”, Geoscience and Remote Sensing, IEEE Transactions on, Vol.54, Jan. 2016.
- [10]P. Gamba and M. Aldrighi, “SAR Data Classification of Urban Areas by Means of Segmentation Techniques and Ancillary Optical Data” IEEE Geoscience and Remote Sensing Letters, VOL. 12, NO. 1, January 2015.
- [11]Akbari, V., Anfinen, S. N., Doulgeris, A. P., and Etloft, T. “The Hotelling-Lawley trace statistic for change detection in polarimetric SAR data under the complex Wishart distribution.” in Proc. IEEE Int. Geosc. Remote Sens. Symp. IGARSS 2013, Melbourne, VIC, pp. 4162-4165,2013.
- [12]Akbari, V., S. N. Anfinen, A. P. Doulgeris, T. Etloft, G. Moser, and S. B. Serpico: Polarimetric SAR Change Detection with the Complex Hotelling-Lawley Trace Statistic, IEEE Trans. Geosci. Remote Sens, 2016.

BIOGRAPHY



Ahmad Reza Aghajani Fesharaki received the Associate degree in Electrical in 2007 from Kharazmi college of Technical & professional, Iran, and received the B.S. degree Electrical (electronic) in Engineering in 2009. Now I'm M.S student in Electronic engineering ACECR Institute of Higher Education (Isfahan Branch), Isfahan, Iran



Mehran Emadi received the B.S degree in Electrical Engineering (Electronic) in 1995,the M.S degree in Electrical Engineering (Electronics) in 1999 from Azad university , Najafabad branch ,Iran and the Ph.D. degree in electrical Engineering (Image processing) from university Technology Malaysia (UTM),Malaysia in 2013. His researchers include semiconductors in his M.S degree and image processing, computer vision, biometric applications in his Ph.D. degree period .There have been than 45 technical paper, published by him and he is an assistant professor in Azad University in Iran and now. He has been working as the program chair, and a committee member for some international conference, and he has been training several students in Universities.