Random Forest Feature Selection for SAR-ATR

Pouya Bolourchi
Department of Electrical and Electronic Engineering
Final International University
Kyrenia, via Mersin 10, Turkey
pouya.bolourchi@final.edu.tr

Masoud Moradi, Hasan Demirel and Sener Uysal
Department of Electrical and Electronic Engineering
Eastern Mediterranean University
Famagusta, via Mersin 10, Turkey
masoud.moradi@cc.emu.edu.tr

***Abstract* - In this paper, a novel approach for selection of relevant features in SAR-ATR is proposed. The main concern of all studies in this filed is the accuracy. For this reason, many researchers have worked on feature extraction phase. Just a few studies focus on feature selection stage. The goal of working on feature selection is twofold. Firstly, the dimensionality of feature space can be reduced and secondary the accuracy can be further improved by eliminating the redundant features. Random Forest is the technique that can be easily implemented over the alternative algorithms such as Genetic Algorithms to SAR-ATR. The easy and fast implementation are the main advantages over the alternative methods. The experimental results show that by selecting just a few features, the accuracy is reaches to saturation.**

Keywords - component; Syentetic Aperture radar; random forest; moment methods

I. Introduction

The applications of Synthetic Aperture Radar (SAR) images increase drastically over the optical images because of their capability to work independently from any weather condition and any solar illumination [1]. SAR images can be seen in many applications such as fire detection [2], flood detection [3], earthquake detection [4] ship detection [5], wave forecasting and marine climatology [6], agricultural industry [7], homeland security applications [8] and military surveillance systems [9]. Automatic Target Recognition (ATR) is one of the application of SAR images that widely studied in many researches. Preprocessing, feature extraction, feature selection and classification are four different phases for target recognition. Due to the noisy background, preprocessing is required in the first phase, to remove the background. As in [10] and [11] two common methods are introduced for noise removal. However, the focus of many researches are in the second and fourth phase, which are feature extraction and classification respectively. In the feature extraction phase, many different algorithms are introduced in the literature. Linear Discriminant Analysis (LDA) [12], Principal Component Analysis (PCA) [12] and Independent Component Analysis (ICA) [13] are the classical method for extracting features. The main drawback associated with those methods is that they are not scale, translation and rotation invariant. Template matching [14] is another technique used in this field. However, for effective recognition rate many templates should be used that arises the computational cost. Recently, many algorithms such as radon transform [15], Linear Binary Pattern (LBP) [15], and wavelet [15] are introduced in the literature. Although the accuracy is improved, many features should be extracted. In order to overcome this problem moment methods are introduced in Cartesian and polar coordinates. Moments in polar coordinates have the property of being scale, rotation and translation invariant [16]. By extracting only a few features, the accuracy is in the acceptable range. Many moments such as Zernike moments (ZM) [17], Pseud Zernike Moments (PZM) [18], and Radial Chebyshev (RCM) [19] are adopted in this paper for a feature extraction. In the classification phase, many classifiers are successfully used in SAR-ATR. Two most commonly used classifiers are Support Vector Machine (SVM) and *k*-Nearest Neighbor (*k*-NN). In this paper for comparison purposes, we utilized both classifiers. Feature selection is a significant phase in classification to minimize the classification errors and decrease the dimensionality of feature space. However this phase is ignored by many researchers. In [20] Genetic Algorithm (GA) is used for selecting the most dominant features. The number of selected features are reduced and the accuracy is improved. In this study the alternative feature selection algorithm - Random Forest followed by feature level fusion- that has the capability to improve the accuracy while reducing the dimensionality of features is proposed. Easy and fast implementation are the main advantages over other techniques such as GA.

The reminder of this study is as follows. In the next section, proposed method is explained in details. It consists of SAR database, preprocessing, feature extraction, feature selection and classification phases. In section III, we discuss the experiment results achieved by proposed method and we compare our results with alternative methods in the literature. Conclusions appear in the final section.

II. Proposed approach

SAR images consist of noisy background. First, we apply some preprocessing to remove the noise and keep the Region of Interest (ROI). Then we apply different moment methods for extracting more informative features. ZM, PZM, and RCM are three method of moments used for feature extraction. After feature extraction - for each

DOI 10.1109/UKSim.2018.00028

moment method- some features in the training set are selected using Random Forest (RF). The selected features of each moment method then are fused and are fed to the classifiers. Finally, two classifiers are used for evaluating accuracy: *k*-NN and SVM classifier. The summary of proposed approach is illustrated in Fig. 1.

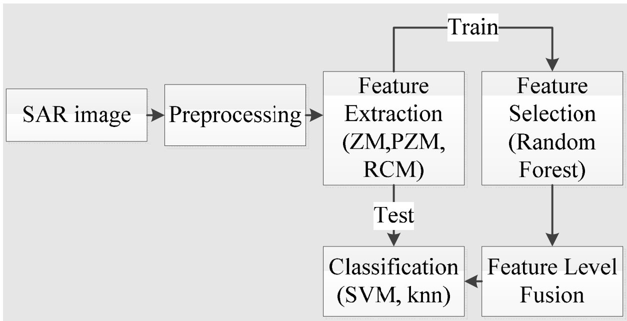

Figure 1. Framework of proposed method

A. *SAR dataset*

Moving Stationary Target Acquisition and Recognition (MSTAR), database [21] are used for the testing and validation of different algorithms. It collected by Sandia National Laboratory (SNL). The number of training and test samples used for each type is listed in Table I. All images with the size of 128 × 128 pixels have one-foot resolution. Totally, 3671 images are collected at 17◦ depression angle for training and 3203 images at 15◦ for testing. Database consists of vehicles from ten different types of ground vehicles. BTR70 is an armored personnel carrier, BMP2 is an infantry fighting, and T72 is a tank

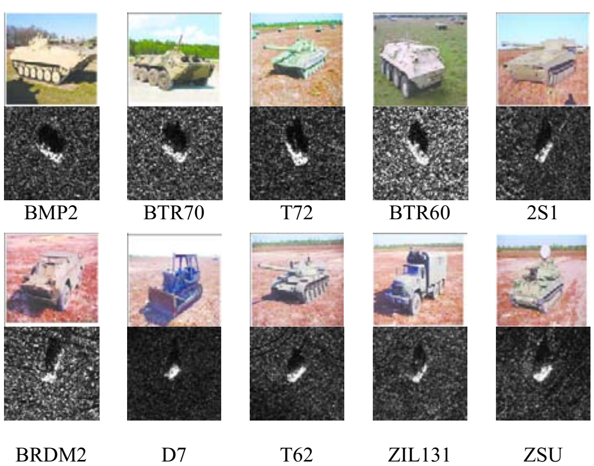

Figure 2. Optical images and SAR images of different vehicles. Figure 3.

These three types of vehicles are commonly used for 3-class problem.10-class problem contains the previous three types (BTR70, BMP2, and T72) and the other seven types of vehicles includes D7 a bulldozer, BDRM2, ZIL131 two types of trucks, 2S1, ZSU-23-4, two types of cannon, BTR60 an armored car and T62 a tank. The optical images of all 10 types of vehicles and their corresponding SAR representation are illustrated in Fig. 1.

B. *Preprocessing Phase*

Each SAR image, consists of noisy background, target and shadow regions. In preprocessing phase, our aim is to remove noisy background while maintaining the target and shadow regions. Histogram equalization followed by averaging filter used to reduce the noise artifacts.

TABLE I. TABLE TYPE STYLE SAR DATA

Vehicle Class	# of samples at 17° of depression angle	# of samples at 15° of depression angle
BMP2	698	587
BTR70	233	582
T72	691	196
BTR60	256	195
2S1	299	274
BRDM2	298	274
D7	299	274
T62	299	273
ZIL131	299	274
ZSU-23-4	299	274
Total	3671	3203

For detecting target and shadow region, we should apply different threshold values respectively. In order to retrieve target region, which is brighter in the image a threshold value (τ) closer to one should be selected. Experimentally τ=0.8 is used for extracting target region. Since shadow is based on electromagnetic waves, which is caused by the depression angle of the aerial vehicle acquiring the images, it can be considered as an extra features that can be used to improve the recognition rate. In order to detect shadow region, a threshold value (τ) closer to zero should be selected. Experimentally τ=0.2 is used for extracting shadow region. Detected target and shadow region called binary of target and shadow respectively. Afterward, we combine them to have a binary target and shadow (binary mask). Finally we multiply the input SAR image with its binary mask to get the texture. This texture consist of target and shadow regions. The schematic of segmented algorithm is illustrated in Fig.3.

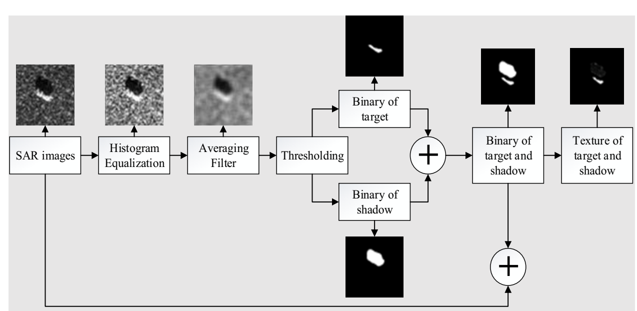

Figure 3. Segmentation algorithms

C. Feature Extraction Phase

In this paper, moment methods in polar forms are used for feature extraction due to their capability to be scale, translation and rotation invariant. Moments also are be able to represent a detected target with a few features. They are scalar quantities, which are used to capture dominant features of an image and mathematically can be presented as [16]:

$$M_{p,q} = \int_0^{2\pi}\int_0^{1} R_{pq}(r)e^{-iq\theta}\tilde{f}(r,\theta)rdrd\theta \qquad (1)$$

where p and q are the order and repetition of moment. $\tilde{f}(r,\theta) = f(r\cos\theta, r\sin\theta), R_{pq}(r)$ called radial part of polynomial and $e^{-iq\theta}$ indicates angular part of polynomials. r,θ refer to transformation from Cartesian to polar coordinates which is defined as:

TABLE II. RADIAL PART OF POLYNOMIAL BASIS FUNCTION OF DIFFERENT MOMENTS IN POLAR COORDINATE

Method	Radial part of Polynomial Basis Function
ZM	$R_{pq}(r) = \sum_{s=0}^{(\frac{p-\lvert q\rvert}{2})} \frac{(-1)^s(p-s)!r^{p-2s}}{s!(\frac{p+\lvert q\rvert}{2}-s)!(\frac{p-\lvert q\rvert}{2}-s)!}$
PZM	$R_{pq}(r) = \sum_{s=0}^{(p-\lvert q\rvert)} \frac{(-1)^s(2p+1-s)!r^{p-s}}{s!(p-\lvert q\rvert-s)!(p-\lvert q\rvert+1-s)!}$
RCM	$R_p(r) = \frac{p!}{\rho(r,N)}\sum_{s=0}^{N}(-1)^{p-s}\binom{N-1-s}{p-s}\binom{p+s}{p}\binom{r}{s}$ $where: (r,N) = \frac{N(1-\frac{1}{N^2})(1-\frac{2^2}{N^2})....(1-\frac{p^2}{N^2})}{2p+1}$ $p = 0,1,...,N-1$

$$r = \sqrt{x^2+y^2}\ , \qquad and \ \ \theta = \tan^{-1}(\frac{y}{x}) \qquad (2)$$

By substituting different polynomial basis functions of Table II into (1), Zernike moments, Pseudo Zernike moments, and Radial Chebyshev moments can be generated. In this paper for each method, 100 features are extracted for comparison purposes.

D. Feature Selection and Feature Level Fusion Phase

The main contribution of this paper is dimensionality reduction by selecting most dominant features using RF. In RF, n decision trees are created by selecting the observations randomly. RF is an ensemble technique in which the final prediction is based on the majority voting among decision trees. A final prediction is computed based on the results of the individual predictions.

For each moment method, we apply RF to select dominant features. In this regard, each feature received a RF score. For each moment, features are ranked in the descending order according to their scores. Finally we select features whose score values are higher than threshold value. The threshold value is the mean of all score values. Fig. 4 shows the RF-score of RCM. Each feature contains a RF-score. Threshold value is achieved by finding the mean value of these scores. The features with higher score value than mean are selected. In this experiment 47, 48 and 58 features are selected for ZM, PZM, and RCM respectively. Then these selected features are fused to form a single vector of 151 elements. Data fusion is used for improving recognition rate. The fused feature vector then is fed to the different classifiers.

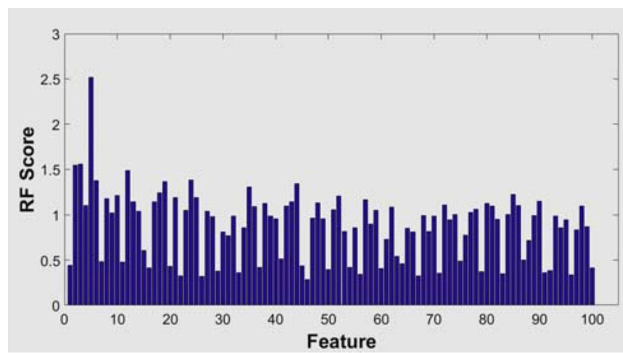

Figure 4. RF score of RCM.

E. Classification Phase

Two classification are used for evaluating the accuracy. Support vector machine and *k*-nearest neighbor. All samples at 15◦ depression angle, consider to be a test set. In the training stage which consists of all samples at 17◦ depression angle, we applied 10-fold class validation technique, in which the data in the training set are divided in to 10 equal subsets. In each fold 9 subset considered as a training set and the remaining one considered as a validated set.

III. RESULTS AND DISCUSSIONS

In this section, we discuss our proposed method and we compare our results with Zernike Moment, Pseudo Zernike Moments and Radial Chebyshev Moments, as well as state-of-the-art methods. The comparisons are used by different metrics as Accuracy (ACC), True Positive Rate (TPR) /sensitivity, True Negative Rate (TNR)/specificity, and Receiver Operating Characteristic (ROC). They are defined as [27]:

$$Sensitivity = TPR = \frac{TN}{TN + FP} \quad (3)$$

$$Specificity = TNR = \frac{TN}{TN + FP} \quad (4)$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (5)$$

Table III shows the accuracy (ACC), sensitivity (TPR) and specificity (TNR) of Zernike moments, Pseudo Zernike moment and radial Chebyshev moments both with and without applying random forest. For fair comparison we applied two classifiers as *k*-NN and SVM. The comparison shows that for both classifiers, as we apply the random forest, the accuracy is either changes slightly or improves significantly. The reason is that by applying feature selection approach, we select the features with no redundancy of information that can affect the accuracy. In PZM for example, when SVM is applied, the accuracy is improved significantly by 3.8% and it reaches to 96.9%. Also in this case the dimensionality of features is reduced to only 48 features.

TABLE III. TABLE TYPE STYLES

Method	ACC (%)	TPR (%)	TNR (%)
ZM+SVM	96.9	96.8	97.2
ZM+*k*-NN	92.9	93.0	94.2
PZM+SVM	93.1	91.7	91.7
PZM+*k*-NN	85.1	82.2	82.9
RCM+SVM	93.0	91.5	91.7
RCM+*k*-NN	85.2	82.6	83.3
ZM+RF+SVM	96.1	95.9	96.1
ZM+RF+*k*-NN	92.1	92.6	93.3
PZM+RF+SVM	96.9	96.4	96.8
PZM+RF+*k*-NN	94.4	94.2	94.8
RCM+RF+SVM	92.6	91.1	91.4
RCM+RF+*k*-NN	84.3	81.1	81.7
ZM+PZM+RCM+RF+SVM	**98.6**	**98.4**	**98.5**
ZM+PZM+RCM+RF+k-NN	95.7	95.6	96.5

Even the higher accuracy is achieved by concatenating the selected features from different three methods (ZM, PZM and RCM) to a single vector and evaluate the performance. Experimental results indicate that by utilizing random forest followed by data fusion the accuracy, sensitivity and specificity is improved in both k-NN and SVM classifiers. The accuracy of proposed method is reaches to 98.6% when SVM classifier is used. Fig. 5 shows the confusion matrix of proposed method. The elements on the main diagonals indicate the correct recognition rate.

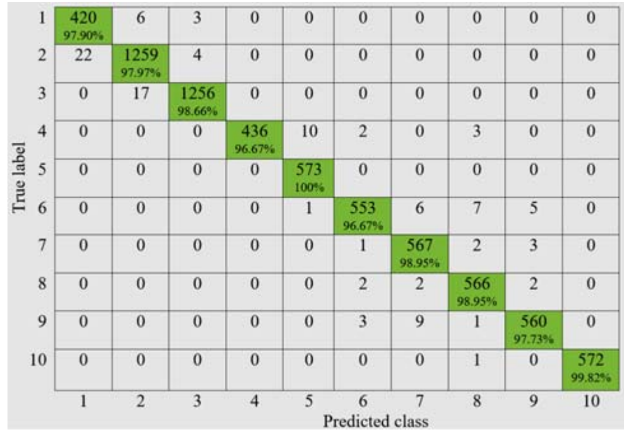

Figure 5. Confusion matrix of proposed method.

ROC also is another metric that shows how well the recognition rate is. ROC is generally designed for two class problems. For multi-class problems first sensitivity and specificity of each class versus others are evaluated. Then, the average sensitivity and specificity are reported as the overall sensitivity and specificity. The ROC can be plotted by using the overall sensitivity and specificity. ROC generated by plotting the sensitivity against 1- specificity. Fig. 6 and Fig. 7 show the ROCs of ZM, PZM and RCM for SVM classifier and *k*-NN classifier respectively. As the Area under Curve (AUC) is larger, then the recognition rate is higher. By comparing ROCs, proposed method has the largest AUC for both classifiers and therefore it has the highest recognition rate.

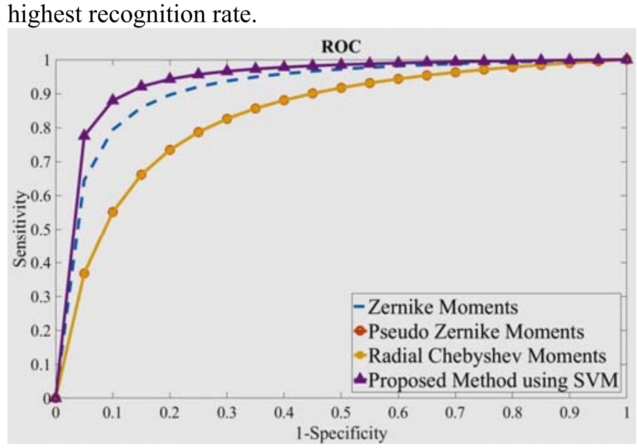

Figure 6. ROCs of ZM, PZM and RCM for SVM classifier

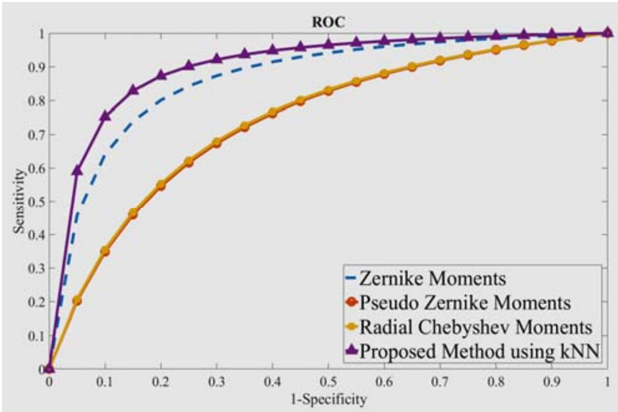

Figure 7. ROCs of ZM, PZM and RCM for k-NN classifier.

Table IV shows the comparison of the performance of the proposed technique with other techniques in the literature. Linear Discriminant Analysis (LDA), Principle Component Analysis (PCA) and Independent Component Analysis (ICA) are a rotational variant, and therefore the recognition rate is poor, even though all methods are fused.

TABLE IV. COMPARISON OF PROPOSED METHOD WITH OTHER TECHNIQUES

Method	Accuracy (%)
LDA [22]	87.4
ZM [23]	89.4
Template Matching [24]	90.4
PCA+LDA+ICA [25]	90.6
MINACE [26]	90.6
PCA [22]	93.3
Seven EFS Coefficient [26]	93.5
QP normalized Image [20]	94.1
RHFM+LBP+HWT+RT+PCA+SVM [15]	98.1
Proposed method	**98.6**

Template matching requires huge amounts of data to be analyzed. A better choice of feature extraction technique is the moment enforced by fusion. In [15] Radial Harmonic Fourier Moment (RHFM) followed by Local Binary Pattern (LBP), Haar Wavelet Transform (HWT), Radon Transform (RT), PCA and SVM and in [19] Radial Chebyshev Moment (RCM) followed by fusion techniques are successfully applied. However, the proposed method outperforms all other techniques in the literature.

IV. CONCLUSIONS

In this paper, a unique approach for feature selection based on random forest is proposed. So far, in order to improve the performance of SAR images, many works are done in the feature extraction phase. Only a few researches focused on feature selection phase. By applying feature selection approaches followed by feature level fusion, the accuracy increases, while the dimensionality of the feature space decreases. In the current study, in the selection phase, we applied random forest in which by selecting a few features from three different moment method and then fusing them to a single vector, the accuracy improved significantly over different method of moments and different classifiers.

REFERENCES

[1] P. Mangalraj and A. Agrawal, "Despeckling of SAR images by directional representation and directional restoration," *Opt. - Int. J. Light Electron Opt.*, vol. 127, no. 1, pp. 116–121, 2016.

[2] F. Siegert and A. Hoffmann, "A quantitative evaluation using high resolution, multitemporal ERS-2 SAR images and NOAA-AVHRR hotspot data." *The 1998 forest fires in East Kalimantan (Indonesia): Remote Sensing of Environment.*, vol.72, no. 1, pp. 64-77, 2000.

[3] J. Cohen, H. Riihimäki, J. Pulliainen, J. Lemmetyinen and J. Heilimo, "Implications of boreal forest stand characteristics for X-band SAR flood mapping accuracy," *Remote Sensing of Environment.*, vol. 186, pp. 47-63, 2016.

[4] M. Chini, N. Pierdicca and W. J. Emery, "Exploiting SAR and VHR optical images to quantify damage caused by the 2003 Bam earthquake," *IEEE Transactions on Geoscience and Remote Sensing.*, vol. 47, no. 1, pp. 145-152, 2009.

[5] S. Wang, M. Wang, S. Yang and L. Jiao, "New hierarchical saliency filtering for fast ship detection in high-resolution SAR images," *IEEE Transactions on Geoscience and Remote Sensing.*, vol. 55, no. 1, pp. 351-362, 2017.

[6] D. Fustes, D. Cantorna, C. Dafonte, B. Arcay, A. Iglesias and M. Manteiga, "A cloud-integrated web platform for marine monitoring using GIS and remote sensing. Application to oil spill detection through SAR images," *Future Generation Computer Systems.*, vol. 34, pp. 155-160, 2014.

[7] E. Erten, J. M. Lopez-Sanchez, O. Yuzugullu and I. Hajnsek, " Retrieval of agricultural crop height from space: A comparison of SAR techniques," *Remote Sensing of Environment.*, vol. 187, pp. 130-144, 2016.

[8] E. J. Baranoski, "Through-wall imaging: Historical perspective and future directions," *Journal of the Franklin Institute.*, vol. 345, no 6, pp. 556-569, 2008.

[9] S. Uḡur and O. Arıkan,"SAR image reconstruction and autofocus by compressed sensing," *Digital Signal Processing.*, vol. 22, no. 6, pp. 923-932, 2012.

[10] Y. Sun, Z. Liu, S. Todorovic and J. Li, "Adaptive boosting for SAR automatic target recognition," *IEEE Transactions on Aerospace and Electronic Systems.,* vol. 43, no. 1, 2007.

[11] L. Hu, J. Liu, H. Liu, B. Chen, B and S. Wu, "Automatic target recognition based on SAR images and two-stage 2DPCA features," *In Synthetic Aperture Radar., APSAR 2007. 1st Asian and Pacific Conference on IEEE.*, pp. 801-805 , November. 2007.

[12] A. K. Mishra, "Validation of PCA and LDA for SAR ATR," *In TENCON 2008-2008 IEEE Region 10 Conference,* pp. 1-6, 2008. November.

[13] L. Maokuan, G. Jian, D. Hui and G. Xin, "SAR ATR based on support vector machines and independent component analysis," *In Radar, 2006. CIE'06,. International Conference on* pp. 1-3. IEEE, 2006. October

[14] A. Mahalanobis, D.W. Carlson, D. W and B.V. Kumar, "Evaluation of MACH and DCCF correlation filters for SAR ATR using MSTAR public data base," *Algorithms for Synthetic Aperture Radar Imagery.*, vol. 3370, pp. 460-468, 1998.

[15] P. Bolourchi, M. Moradi, H. Demirel and S. Uysal, "Feature Fusion for Classification Enhancement of Ground Vehicle SAR Images," *UKSim-AMSS 19th International Conference on Modelling & Simulation.,* European, pp. 111-115, April. 2017.

[16] J. Flusser, B. Zitova and T. Suk, "Moments and moment invariants in pattern recognition," *John Wiley & Sons.*, 2009.
[17] A. Khotanzad and Y. H. Hong, "Invariant image recognition by Zernike moments," *IEEE Transactions on pattern analysis and machine intelligence.*, vol. 12, no. 5, pp. 489-497, 1990.
[18] C. Singh and R. Upneja, "Accurate calculation of high order pseudo-Zernike moments and their numerical stability," *Digital Signal Processing.,* vol. 27, pp. 95-106, 2014.
[19] P. Bolourchi, H. Demirel and S. Uysal, "Target recognition in SAR images using radial Chebyshev moments," *Signal, Image and Video Processing.*, vol. 11, no. 6, pp. 1033-1040, 2017.
[20] M. Amoon and G.A. Rezai-Rad, "Automatic target recognition of synthetic aperture radar (SAR) images based on optimal selection of Zernike moments features," *IET Computer Vision.*, vol. 8, no.2, pp. 77-85, 2013.
[21] Sensor Data Management System (SDMS) Public website. (2017, June 22). Retrieved from https://www.sdms.afrl.af.mil
[22] B. Wang, Y. Huang, J. Yang and J Wu, "A feature extraction method for synthetic aperture radar (SAR) automatic target recognition based on maximum interclass distance," *Science China Technological Sciences.*, vol. 54, no. 9, pp. 2520-2524, 2011.
[23] P. Bolourchi, H. Demirel, H and S. Uysal, "Continuous Moment-Based Features for Classification of Ground Vehicle SAR Images," *In Modelling Symposium (EMS).,* 2016, European, pp. 53-57, 2016, November.
[24] Q. Zhao and J.C. Principe, "Support vector machines for SAR automatic target recognition," *IEEE Transactions on Aerospace and Electronic Systems.,* vol. 37, no. 2, pp. 643-654, 2001.
[25] X. Yuan, T. Tang, D. Xiang, Y. Li and Y. Su, "Target recognition in SAR imagery based on local gradient ratio pattern," *International journal of remote sensing.*, vol. 35, no. 3, pp. 857-870, 2014.
[26] R. Patnaik and D. Casasent, "SAR classification and confuser and clutter rejection tests on MSTAR ten-class data using Minace filters," *In Optical Pattern Recognition XVIII International Society for Optics and Photonics.* vol. 6574, p. 657402, April, 2007.
[27] A. Tahmasbi, F. Saki and S.B. Shokouhi, "Classification of benign and malignant masses based on Zernike moments," *Computers in biology and medicine.,* vol. 41, no. 8, pp. 726-735, 2011.