A SOFT COMPUTING METHOD FOR AUTOMATED DAMAGE MAPPING USING VHR IMAGERY

Babak MANSOURI
Assistant Professor, Dept. of Emergency Management, International Institute of Earthquake Engineering and Seismology, Tehran. Iran
mansouri@iiees.ac.ir

Yaser HAMEDNIA
Ms student, International Institute of Earthquake Engineering and Seismology, Tehran, Iran
y.hamednia@iiees@ac.ir

Kambod AMINI-HOSSEINI
Associate Professor, Risk Management Research Center, International Institute of Earthquake Engineering and Seismology, Tehran. Iran
kamini@iiees.ac.ir

Keywords: Damage Map, Textural Features, Genetic Algorithm, Support Vector Machine, Fuzzy Inference System

ABSTRACT

In disaster management operations right after a disastrous earthquake, damage maps are essential as they reveal the exact location of hard hit areas in urban settings. This study investigates the feasibility of an automated soft computing algorithm in generating damage maps. The main idea is to classify different areas of each building roof as three classes relative to three distinct damage levels. Then, a fuzzy inference methodology is used to determine the damage grade of each building by the means of evaluating the percentages for three damage levels detected for individual building roofs. For implementation, satellite images of before and after the 2003 Bam, Iran earthquake, are used in addition to some available ancillary data. Firstly, a pre-processing step was completed involving the co-registration and enhancement of images. Thereafter, the roofs of buildings were extracted from the images by using the ancillary data. The Haralick's textural features were computed for the images where an optimum set of three such features were selected using the Genetic Algorithm. Then, the roof of the buildings were classified in three classes namely " intact", "partially-damaged" and "fully-damaged" using the selected optimum textural indices and by exploiting a Support Vector Machine (SVM) supervised classification algorithm. Thence, for individual building roofs, the percentage of pixels within each class was calculated as the input of a Fuzzy Inference System (FIS). Mamdani fuzzy inference engine was used to determine the damage grade of each building as to produce the damage map. The proposed algorithm was evaluated by comparing the produced damage map with a reference damage map as ground truth where the results demonstrated the efficacy of the method showing an overall accuracy of 76% for such rapid screening process.

1. INTRODUCTION

Natural disasters have affected lives of millions of peoples around the world each year. It is possible to reduce the impact of disasters such as earthquakes with effective disaster management strategies. A rapid screening procedure is essential to evaluate the magnitude and the extent of damage timely in disaster management activities such as search and rescue and relief planning. Remote Sensing and GIS technologies have been progressively advanced and routinely exploited in rapid disaster damage assessment for the purpose of disaster management. Very High Resolution (VHR) optical data and associated damage detection algorithms were proved to be effective in estimating earthquake damages in urban areas (Mansouri et al., 2005, Gusella et al., 2005, Eguchi et al., 2008). Commercially available VHR (Very High Resolution) satellites such as IKONOS and Quickbird provide images with 4m and 2.4m spatial resolution in multispectral bands and 1m and 0.61m in panchromatic band respectively. Such remote sensing systems have proved to be useful in acquiring data and processing crucial information for disaster management.

In this paper, a novel object-based method for automated damage detection using very high resolution (VHR) optical satellite imagery is proposed. After pre-processing steps needed to match before and after satellite images spectrally and spatially, eight Haralick textural features were extracted and an optimum set of three such features were selected using Genetic Algorithm. The classification of each building roof was completed for three classes relative to three distinct damage areas. These three classes (distinct damage area for individual building footprint) are determined according to a Support Vector Machine (SVM) scheme as a supervised classification method. Finally, a fuzzy inference methodology is used to determine the damage grade of each building (footprint) by the means of evaluating the percentages for distinct footprint damage areas. As follows, the details of the proposed method, the implementation steps and results are reported for the 2003 Bam earthquake as case study.

2. METHODOLOGY - DAMAGE MAP GENERATION

The proposed method for determining the extent of buildings damage is shown as the flowchart in Figure1. Data resources used in this study involve satellite images related to before and after the 2003Bam, Iran earthquake event in addition to a building vector map as ancillary data.

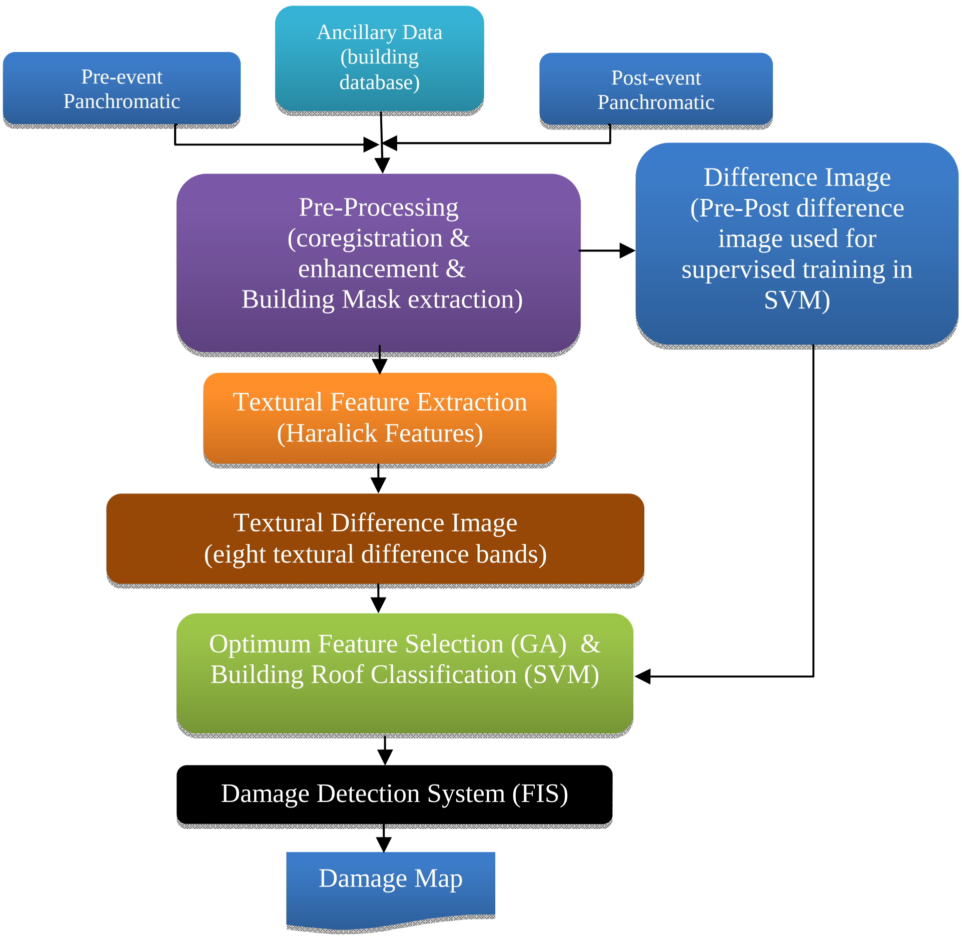


Figure 1. Flowchart of the proposed method

As seen in Figure1, the initial step involves the pre-processing of the satellite images and consequently, a difference image was made by subtracting the after from the before data. Also, a building roof mask is made from the vector map obtained as ancillary data. This vector layer was generated by a study conducted by Yamazaki et al., 2005 and complemented by some ground truth and 1:2000 digital map of the region (Mansouri et al., 2008). Textural features are extracted from images and as a result, the textural difference image was produced in eight textural bands. Then a set of optimum features was determined using an algorithm as a combination of Genetic Algorithm (GA) and Support Vector Machine (SVM). The roofs are classified in three classes as sub-objects using the supervised classifier (SVM). Finally, by considering the percentage of pixels within the aforementioned three classes (three sub-objects) in a fuzzy inference system, the damage for individual buildings is estimated and the damage map is created. Further details of the proposed method are described in the following sections.

2.1. PRE-PROCESSING

The panchromatic band of the images were used and the pre-processing step involved geo-referencing and image enhancement as histogram equalization and histogram matching. Also, a difference image was created by subtracting the post-event from the pre-event images. Moreover, roofs of the buildings (buildings footprints) were considered as main objects for damage evaluation, therefore; locating and extracting image features associated to building footprints were important. The building mask was generated from an ancillary vector layer where building roofs were extracted visually from satellite images (Yamazaki et al, 2005) and also from a 1:2000 scale digital map incorporated into a GIS (Mansouri et al., 2008).

2.2. TEXTURAL FEATURE EXTRACTION

Extracting and processing textural features in optical satellite images is useful in detecting damaged building in earthquake affected areas. Damaged buildings are distinctive from intact ones considering their textural features. Two-dimensional Grey-Level Co-occurrence Matrices (GLCM), are widely used in textural analysis because they are able to detect the spatial dependency of grey-level values within an image (Haralick et al., 1973). The mathematical representations for eight used features are shown in Table1.

Table 1. Haralick textural features extracted from GLCM matrix

Feature	Formula	Feature	Formula
Mean	$\frac{1}{2}\sum_{i}^{M}\sum_{j}^{M}(iP[i,j]+jP[i,j])$	Dissimilarity	$\sum_{i}^{M}\sum_{j}^{M}\lvert i-j\rvert P[i,j]$
Variance	$\frac{1}{2}\sum_{i}^{M}\sum_{j}^{M}((i-\mu)^2P[i,j]+(j-\mu)^2P[i,j])$	Sec-Moment	$\sum_{i}^{M}\sum_{j}^{M}P[i,j]^2$
Homogeneity	$\sum_{i}^{M}\sum_{j}^{M}\frac{P[i,j]}{1+\lvert i-j\rvert}$	Entropy	$-\sum_{i}^{M}\sum_{j}^{M}P[i,j]\,log\,P[i,j]$
Contrast	$\sum_{i}^{M}\sum_{j}^{M}(i-j)^2P[i,j]$	Correlation	$\sum_{i}^{M}\sum_{j}^{M}\frac{P[i,j]}{(i-j)^2}$

Haralick second order statistical features for both pre-event and post-event images were measured with a 5x5 window. Then, an eight band textural difference image was produced by subtracting relative textural images for individual footprint.

2.3. OPTIMUM FEATURE SELECTION AND ROOF CLASSIFICATION

The optimum feature selection is performed according to a combination of GA algorithm and SVM classification. As seen in Figure 2, the features are feed into the GA algorithm and the accuracy for selecting three feature out of eight candidate are decided by the fitness function in SVM procedure.

Genetic algorithm is widely used as an optimization method. In order to select an optimum set of features for image classification, the genetic optimization algorithm is used. Chromosomes are vectors in which the number of textural features was considered as the number of genes (8 genes). According to Figure 3, true value 1 for any gene means the presence of corresponding feature in the feature vector. The fitness function of the algorithm is the difference of Kappa coefficients measured in SVM classification for each generation (in GA).

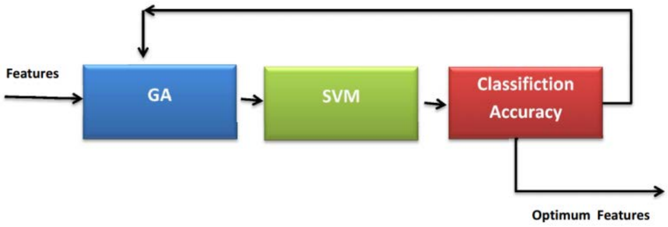


Figure 2. Genetic Algorithm for Optimal features selection

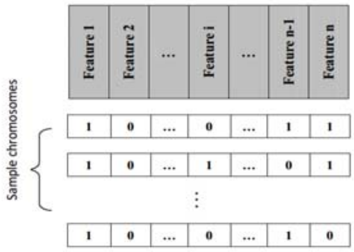


Figure 3. String of an n-bit chromosome

SVM is a powerful supervised classification method in which accurate results can be obtained using relatively limited number of training data as compared with other conventional supervised classification methods. SVM is a method that separate different classes within the training data space. Generally, SVM is a binary and a linear classifier in which by the use of kernel functions, multiclass and nonlinear data classes can be delineated. There are some prevalent kernel functions such as linear kernel, polynomial kernel, quadratic kernel, and RBF (Radial Basis Function) kernel that can be used in image classification. RBF kernel are widely used in land cover classification by satellite images and it has produced better results in comparison with other kernels as reported (Keuchel et al., 2003, Knorn et al., 2009). In this study, RBF kernel was used for classification.

Figure 4 depicts an arrangement for some neighbouring building footprints where their sub-region content variations are represented by three different colours. Each sub-region indicates the output of the SVM classification where the training data from the difference image was utilized as input. The main idea is to estimate the final building damage grade by computing the share of these sub-region components and relating them with the actual physical damage. This final phase of damage detection is performed by FIS as described in the next part.

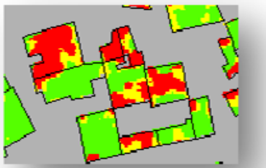


Figure 4. Sample arrangement of neighbouring roofs with some sub-region content variation

2.4. DAMAGE DETECION – FUZZY INFERENCE SYSTEM

Fuzzy inferencing is the process of mapping some input variable to one output variable using fuzzy logic. Generally a fuzzy inference system is comprised of three principal steps as fuzzification, inferencing and defuzzificaiton, as shown in Figure 5. The fuzzification phase involves the division of input feature space into fuzzy subspaces where each are specified by fuzzy membership function. Fuzzy rules are then generated by comparing the percentage of each damage region with the damage grade of each candidate building. The inferencing stage requires the calculation of the strength of each rule which is being triggered. Finally, the defuzzification step aggregates all triggered rules and generates a non-fuzzy output.

After classifying the textural difference image (an 8-band image) by SVM, three sub-region classes are identified. Also, the relative population of pixels representing the three sub-regions was calculated and set as the input to the fuzzy inference system (FIS). FIS is used to determine the actual physical damage grade of each building considering the three sub-regional contents. The membership functions for input and output variable were adjusted for the cases. Also, the fuzzy rules were determined by comparing the percentage of individual sub-region classes according to some validated samples. Finally, each building was labelled according to the EMS-98 damage grades.

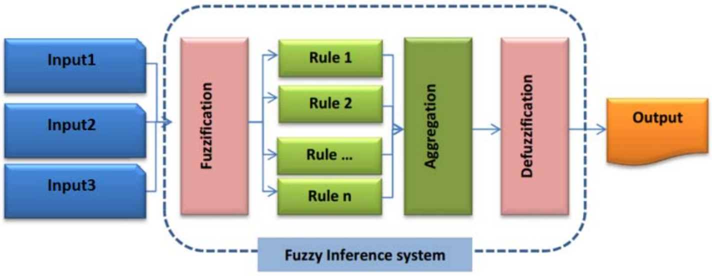


Figure 5. Schematic for Fuzzy Inference System

3. IMPLEMENTATION AND RESULTS

In this study, the panchromatic band of pre-event and post-event QuickBird satellite was used. A sample region of two mentioned images is shown in Figure 6. From the original building mask for the entire city of Bam, a total of dominant 11294 building footprints were used as reported (section 2.1). After extracting buildings footprint and making the difference image, eight Haralick second degree textural features (as described in part 2.2) were extracted from pre-event and post-event images and textural difference image was made. Then applying the GA optimization algorithm, three optimum features namely; Mean, Dissimilarity and Second Moment were selected. The kappa coefficient of the selected chromosome was computed as 0.85 which shows high degree of accuracy.

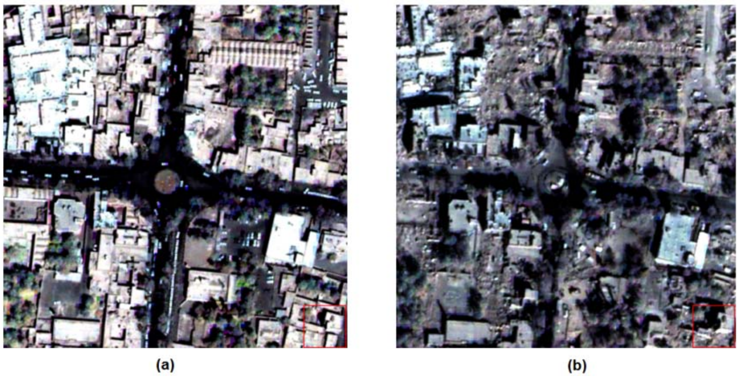


Figure 6. Pre-event(a) and post-event(b) sub-images of Bam (QuickBird)

In SVM, the training data of 500 pixels per each of the three damage regions were visually selected in random fashion. Each pixel represented spatially a sub-region (sub-object) associated to different parts of individual building footprints. These locations were visually selected from the difference panchromatic image for its better human perception. The test data comprised of 1500 pixels selected visually similar to the training data but associated to the textural difference image space. SVM classification was performed for each feature vector (chromosome) where the residual of the Kappa coefficients was used as the GA fitness function.

Using the SVM classification and the three optimum selected features and the training data, the textural difference image was classified into three classes of “intact region”, “partially-damaged region” and “fully-damaged region”. As shown in Figure7, three mentioned sub-region classes are presented in green, yellow and red respectively.

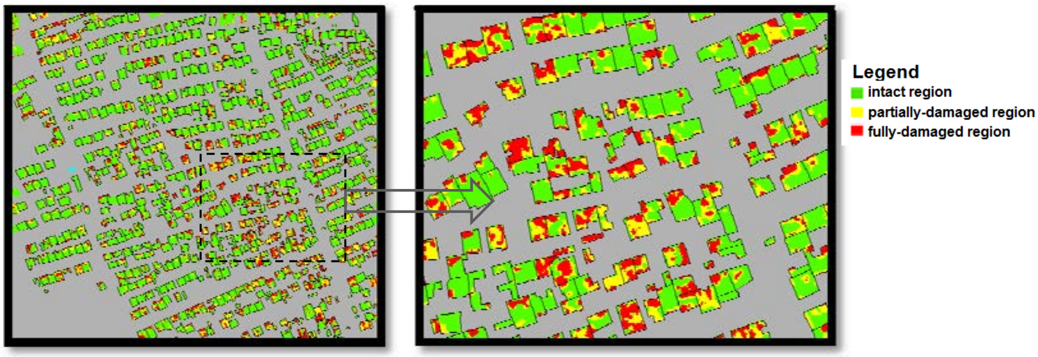
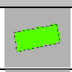
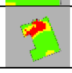


Figure 7. Result of image classification with SVM classification

Table 2. Fraction of pixels per different sub-region classes for three sample footprints

Building footprint	Intact Area	Partially-damaged area	Fully- damaged area
	1	0	0
	0	0.015384615	0.984615385
	0.650704225	0.169014085	0.18028169

As shown in Table 2, the percentage of each sub-region pixels was calculated for all buildings. These values were used as the input variables of FIS in the next step. To produce the desired damage map, the percentage of three classes (sub-regions) in each building footprint is regarded as input to FIS where the membership functions used for fuzzification and defuzzification are depicted in Figure 8. In order to derive the fuzzy rules, twenty five building footprint were selected randomly per each damage class as reference. Some of these rules are shown in Table 3. Mamdani fuzzy engine was chosen for this computation.

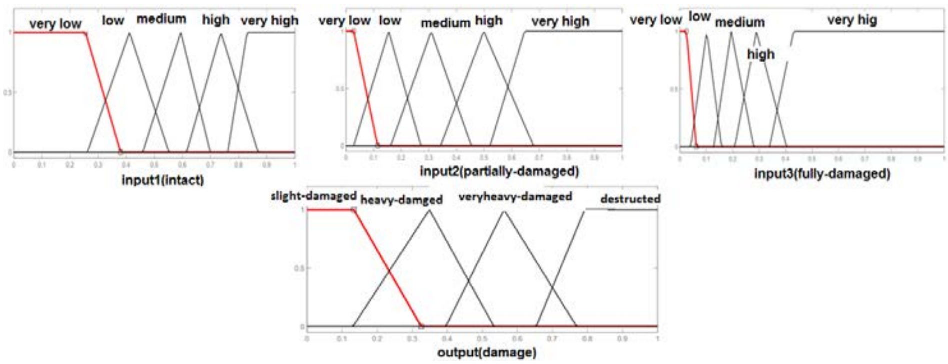


Figure 8. Membership functions for input and output variables

Table 3. Some rules in Fuzzy Inference System

-IF intact is very high AND partially-damaged is very low AND fully-damaged is very low THEN damage is slight- damaged
-IF intact is high AND partially-damaged is low AND fully-damaged is low THEN damage is heavy-damaged
-IF intact is very low AND partially-damaged is medium AND fully-damaged is very high THEN damage is destructed

The output of the FIS is a damage label assigned to each building. The confusion matrix was calculated using the results from this research and a reference dataset as ground truth (Table 5 and Table 6). In confusion matrices, D1, D3, D4 and D5 represent "negligible to slight damage", "substantial to heavy damage"," very heavy damage" and" destruction" grade respectively according to EMS98 damage classification (Table 4). In case1, D1 and D3 were aggregated together in one class namely D1&3. In case2, D4 and D5 were combined together in one class namely D4&5. The Overall Accuracy (O.A), the Kappa coefficient (K), the User Accuracy (U.A) and the Producer Accuracy (P.A) were calculated for these cases. The reference damage map is shown in Figure 9 and the damage maps are depicted in Figure10.

Table 4. EMS-98 building damage classification (Grunthal, 1998)

Damage Grade	Definition
Grade 1	Negligible to slight damage
Grade 2	Moderate damage
Grade 3	Substantial to heavy damage
Grade 4	Very heavy damage
Grade 5	Destruction

Table 5. Confusion Matrix (case1)

Confusion Matrix (D1&3,D4,D5)		Proposed Method				
		D1&3	D4	D5	sum	P.A
Reference Data	D1&3	4490	639	84	5213	0.86
	D5	259	1074	307	1640	0.65
	D5	312	1416	2713	4441	0.61
	sum	5061	3129	3104	O.A = 0.73 K = 0.58	
	U.A.	0.89	0.34	0.87		

Table 6. Confusion Matrix (case2)

Confusion Matrix (D1,D3,D4&5)		Proposed Method				
		D1	D3	d4&5	sum	P.A
Reference Data	D1	954	508	73	1535	0.62
	D3	893	2135	650	3678	0.58
	D4&5	32	539	5510	6081	0.91
	sum	1657	3214	6423	O.A = 0.76 K = 0.59	
	U.A.	0.58	0.66	0.86		

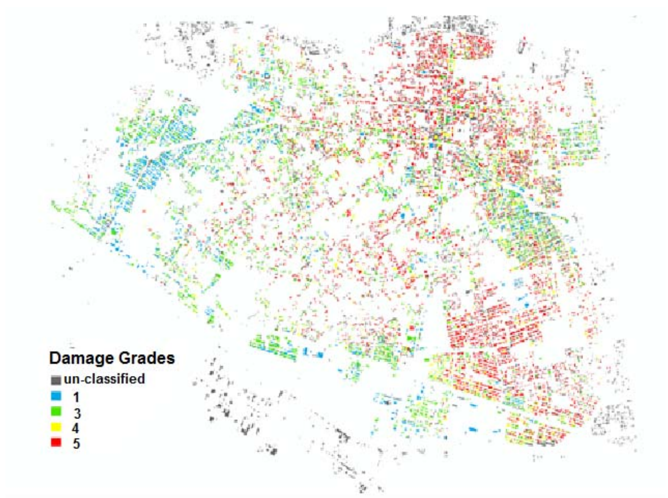


Figure 9. Reference damage map, courtesy of (Yamazaki el al., 2005)

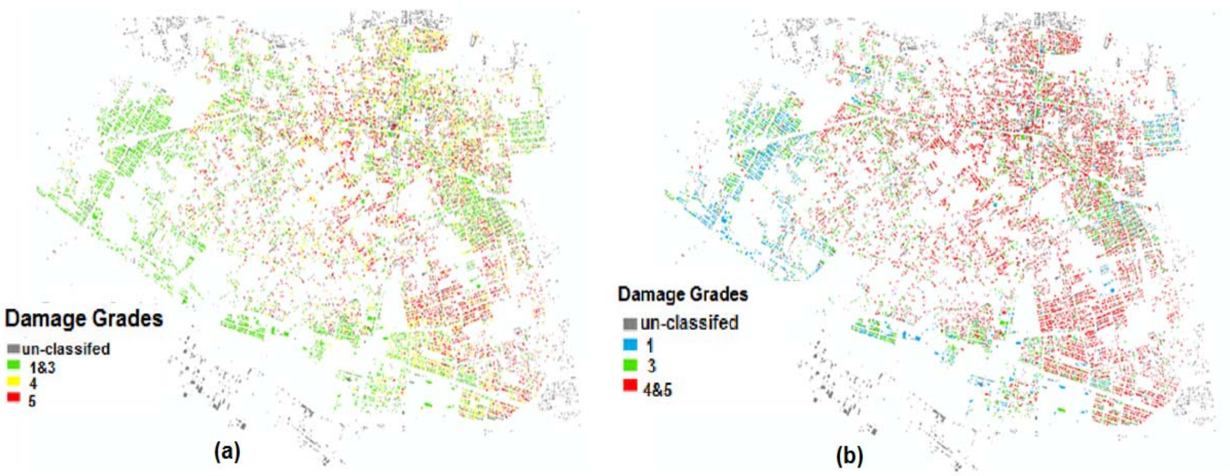


Figure 10. Modelled damage maps for case1 (a) and case2 (b)

4. CONCLUSION

In this study, a novel method for generating semi-automated damage map was proposed using eight Haralick textural features. The method primarily is based on soft-computing and in the first stage exploits GA (Genetic Algorithm) and SVM (Support Vector Machine) algorithms as a hybrid tool to select some optimum features. In the second stage, the Fuzzy Inference System (FIS) provide a means to produce damage maps. The process was applied to the urban area of Bam considering the 2003 earthquake data. In this research, a total of 11294 building roof were evaluated and their associated damage grades were determined. The proposed algorithm was evaluated by comparing the produced damage map with a reference damage map as ground truth. Considering the development of a rapid damage mapping scheme, the results demonstrated the efficacy of the method showing an overall accuracy of 73% and 76%, and the Kappa coefficient of 58% and 59% for case1 and case2 respectively.

REFERENCES

Eguchi R, Huyck C, Ghosh S and Adams B (2008) The application of remote sensing technologies for disaster management, The 14th World Conference on Earthquake Engineering

Gusella L, Adams J, Bittelli G, Huyck C and Mognol A (2005) Object-Oriented Image Understanding and Post-Earthquake Damage Assessment for the 2003 Bam, Iran, Earthquake, Earthquake Spectra, Vol. 21, No. , pp. 225-238

Haralick, R M, Shanmugan K and Dinesten I (1973) Textural feature For Image Classification. IEEE Transaction on System Man and Cyernetics , Volume 3, NO 6

Keuchel J, Naumann S and Heiler M (2003) Automatic land cover analysis for Tenerife by supervised classification usign remotely sensed data, Remote Sens,Environ., 86, pp. 536-541

Knorn J, Rabe A, Radeloff V, Kuemmerle T, Kozak J and Horsetert P (2009) Land cover mapping of large areas using chain classification of neighboring Landsat satellite images, Remote sens. Environ.,113,pp,957-964

Mansouri B, Mousavi M and Amini-Hosseini K (2005) Parcel-based Damage Detection using VHR Optical Data, 6th International Workshop on Remote Sensing for Disaster Applications

Yamazaki F and Yano Y (2005) Visual Damage Interpretation of Buildings in Bam City Using QuickBird Images Following the 2003 Bam, Iran, Earthquake, Earthquake Spectra, Spatial Issue 1, Vol.21, S329