World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:6, No:6, 2012

Site Selection of Traffic Camera based on Dempster-Shafer and Bagging Theory

S. Rokhsari, M. Delavar, A. Sadeghi-Niaraki, A. Abed-Elmdoust, B. Moshiri

***Abstract*—**Traffic incident has bad effect on all parts of society so controlling road networks with enough traffic devices could help to decrease number of accidents, so using the best method for optimum site selection of these devices could help to implement good monitoring system. This paper has considered here important criteria for optimum site selection of traffic camera based on aggregation methods such as Bagging and Dempster-Shafer concepts. In the first step, important criteria such as annual traffic flow, distance from critical places such as parks that need more traffic controlling were identified for selection of important road links for traffic camera installation, Then classification methods such as Artificial neural network and Decision tree algorithms were employed for classification of road links based on their importance for camera installation. Then for improving the result of classifiers aggregation methods such as Bagging and Dempster-Shafer theories were used.

***Keywords*—**Aggregation, Bagging theory, Dempster-Shafer theory, Site selection

I. INTRODUCTION

OPTIMUM site selection of traffic camera could help to obtain real time data without human intervention with high accuracy and helps to decrease road accidents. For doing Optimum site selection following steps should be followed:

A. Identification of criteria

For site selection of traffic sensors and cameras the first step are determination of criteria, some criteria such as annual traffic distance from critical places such and parks and that need more controlling.

B. Classification of road links

In the next step, some classification methods such as artificial neural network and Decision tree were employed for classification of road links based on their importance for camera installation. Section-2 demonstrates the principal of the methodology of classification based on these classifiers.

C. Aggregation of single classifiers

For improving the results of single classifiers aggregation methods such Bagging and Dempster-Shafer theories were used.

S. Rokhsari is Student of GIS Engineering in University of Tehran, Iran (e-mail: so_rokhsari@yahoo.com).

Dr.Delavar is with Department of Surveying and Geomatic Engineering, College of Engineering, University of Tehran, Iran.

Dr.Sadeghi is with Department of Geoinformatic Engineering at Inha University, South Korea.

A.Abed-Elmdoust is PhD Student of Civil Engineering in University of Tehran.

Dr.Moshiri is with Department of Electrical and Computer Engineering, University of Tehran, Tehran, Iran.

Section-3 demonstrates the principal of the aggregation concept for improvement of classification results based on single classifiers such as Decision tree and neural network.

D. Identification of critical links

After using classification based on artificial neural network and Decision tree classifiers and aggregation method for improvement the result of classifiers, road links were classified in two classes based on their importance for camera installation so links in high risk class will be selected for camera installation in future; these steps of camera site selection are shown on Fig. 1.

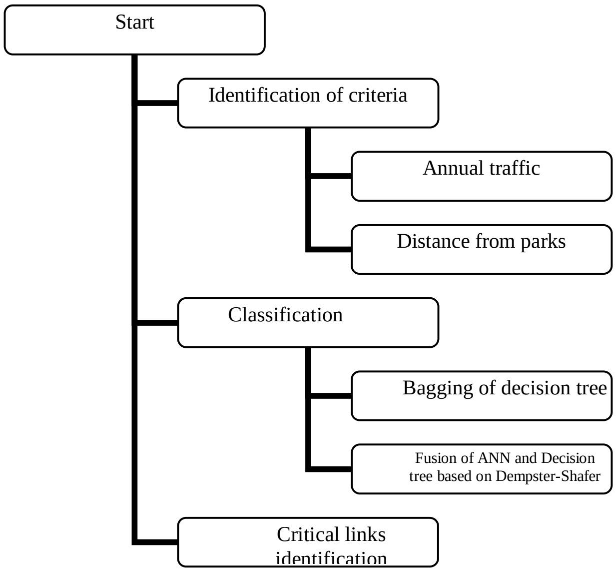

Fig. 1 The steps of the research

II. CLASSIFICATION

In supervised learning or classification we use learning data for identification of category of data. Two famous classifiers are decision tree and artificial neural network.

A. Artificial neural network

The usage of neural network could help experts to perform intelligent tasks in a manner similar to human brain.ANNs resemble the human brain in two ways:

(1) They obtain knowledge based on learning.

(2) Their knowledge is saved within interneuron connection strengths known as synaptic weights.

The aim of any neural network is to identify a relationship between the inputs and outputs.

ANN can be classified in to supervised and unsupervised based on learning algorithm. A supervised network has its output compared to known answers in training process.

A famous type of neural network model is the multi-layer perceptron (MLP), it needs a desired output for learning.

ANNs are easy to use, and they can approximate any input/output although require a relatively large data set for training, although they are slow to train, strongly correlated inputs can be a problem, continuous and discrete inputs can be difficult to handle.

The MLPs learn using an algorithm called back propagation. Basically, back propagation is a supervised learning that during the learning, the input data is repeatedly presented to the neural network. With each presentation, the output of the neural network is compared to the desired output and an error is calculated. This error is then fed back (back propagated) to the neural network and utilize to adjust the weights such for decreasing error so this could improve weight of network. Forward pass and backward estimations are shown on (1), (2): [1].

$$y_i = f(\sum w_{ij} Y_j) + X_i \tag{1}$$

Where: y_i = System output , W_{ij} = Weight, X_i = Current input

$$e_i = -\varepsilon_i + \sum w_{ij} \sigma_j \tag{2}$$

Where: e_i = Produced error , ε_i = Injected error ,
σ_j = Error computes by topology

In ANN data set should be divided into three types of data

B. Training set

Training set is the portion of the data used to train the network. This is the largest portion of the data at least is 50%.

C. Cross validation set

Cross validation set is the data set aside to test the network during training.

D. Testing set

Testing set is used for validation of the result.

E. Measuring Performance of Classification

Performance of classification could be estimated by learning curve, if a learning curve is noisy or (jumping up and down), this means the network is not training well, Fig. 2 shows a good learning curve for a given dataset.

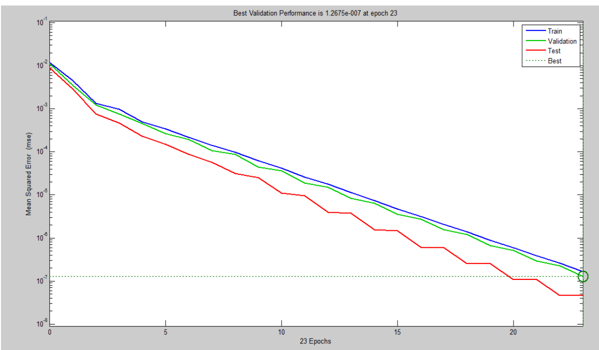

Fig. 2 The learning curve for a given dataset

F. Decision tree

Decision tree is a famous classifier that does not require any knowledge or parameter setting. The approach is supervised learning. Given a training data, we can extract a decision tree to make rules about the data; the structure of a Decision tree is shown on Fig. 3.

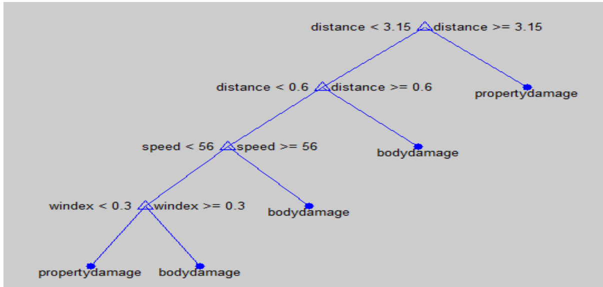

Fig. 3 Decision trees structure on a given dataset for severity analysis

III. Aggregation of Single Classifiers

Aggregation of some classifiers could improve the result of classification result two famous methods for aggregations are Bagging and Dempster-Shafer.

A. Bagging tree

Bagging predictors is a famous method for generating multiple versions of a predictor and using these to acquire an aggregated predictor. The aggregation averages over the versions when predicting a numerical outcome and does a plurality vote when predicting a class. The multiple versions are formed by making bootstrap replicates of the learning set and employing these as new learning sets.

For learning set $L := \{(y_n, x_n), n = 1,2,..., N\}$ where: $x_n \in X$ is the input and $y_n \in Y$ is the response from a predictor denoted by $\Phi(X, L)$.When consider k learning sets $\{(L_k)\}$ of N_{iid} (independent identical distribution) observation from the similar distribution of L, it is obvious to use the average to get to a better result for classification, this concept is shown on (3): [2, 3].

$$\phi_A(X) = E_L \phi(X, L) \tag{3}$$

Where: E_L Denotes the expectation over L.

When there is just one learning set L, one resolvable way is to take bootstrap samples {L(B)} of N(B) cases, randomly, but with replacement from L, and form predictor set { ϕ(X, L(B))}.

Then the predictor $\Phi_A(X)$, which gives us the most voting score of test set, is the best result for classification.

B. Dempster-Shafer

Dempster-Shafer theory offers an alternative to traditional probabilistic theory for the mathematical representation of uncertainty, it let us to allocate probability mass to sets or intervals. It gives combination rules are the special types of aggregation methods for data acquired from multiple sources to obtain the better result than single source.

The most important assumption in this theory is independency of sources. In a simple example, suppose m_1

and m_2 are two mass functions formed based on information obtained from two different information sources, combination rule for these sources is shown on (4):[4]

$$m(c) = (m_1 \oplus m_2)(c) = \frac{\sum_{A \cap B = c} m_1(A) m_2(B)}{1 - K} \quad (4)$$

Where K represents a basic probability mass associated with conflicts among the sources of evidence. The conflict K is defined in (5): [4].

$$K = \sum_{A \cap B = \phi} m_1(A) m_2(B) < 1 \quad (5)$$

IV. Case Study

The optimum site selection of traffic camera was calculated with usage of Matlab functions for fusion and clustering. To evaluate the steps of process we used a part of USA road map (north of Washington), Meta data information is mentioned as below:

Projected_Coordinate_System: NAD_1983_HARN_StatePlane_Washington_North_FIPS_4601_Feet.

Projection: Lambert_Conformal_Conic.

Geographic_Coordinate_System:GCS_North_American_1983_HARN.

Datum: D_North_American_1983_HARN.

The data was prepared in vector formats that contain some non spatial parameters of traffic data like annual traffic.

Parcels of parks were prepared for calculation of distance criteria for site selection. Fig. 4 shows the structure of the data in vector format.

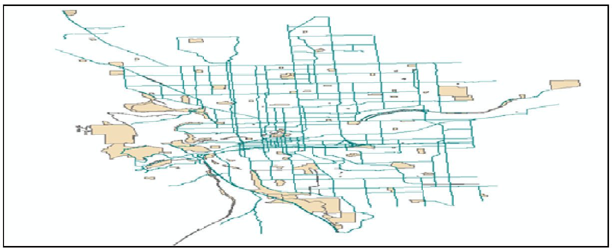

Fig. 4 Input data for camera site selection in a vector format

V. Result and Discussions

In the first step, important criteria such as annual traffic, distance from critical places such as parks that need more traffic controlling were identified for selection of important road links for traffic camera installation, Then classification methods such as Artificial neural network and Decision tree algorithms were employed for classification of road links based on their importance for camera installation.

In order to classify road links, learning data set was prepared,

60% of data set was used as a training data, 20% was used for checking the reliability of these methods.

For increasing the accuracy of Decision tree, Bagging tree was used instead of that.

Fig. 5, 6 shows estimated mean square error in ANN and Bagging tree.

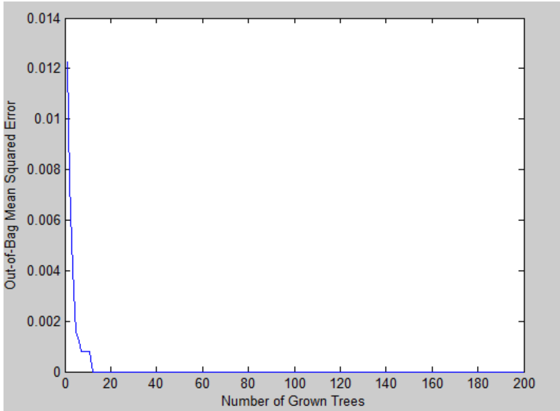

Fig. 5 Out-of-bag error graph in bagged tree

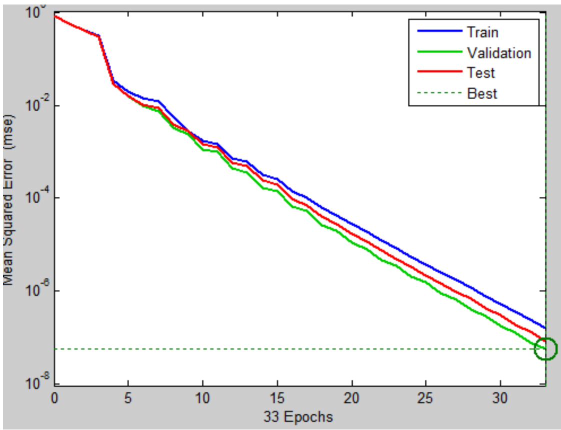

Fig. 6 The mean square graph for classification with ANN

Then Dempster-Shafer were used to aggregate the result of ANN and Decision tree classifiers to improve the result of classification, the combination rule for classification is shown on (6),(7).Where $m_N(L)$ is mass function of low risk class for neural network classifier and $m_D(L)$ is mass function of low risk class for Decision tree classifier, $m_N(S)$ is mass function of high risk class for neural network classifier and $m_D(S)$ is mass function of high risk class for Decision tree classifier

$$m(L) = \frac{m_N(L)m_D(L) + m_N(L)m_D(L \cup S) + m_N(L \cup S)m_D(L)}{1 - (m_N(L)m_D(S) + m_N(S)m_D(L))} \quad (6)$$

$$m(S) = 1 - m(L) \quad (7)$$

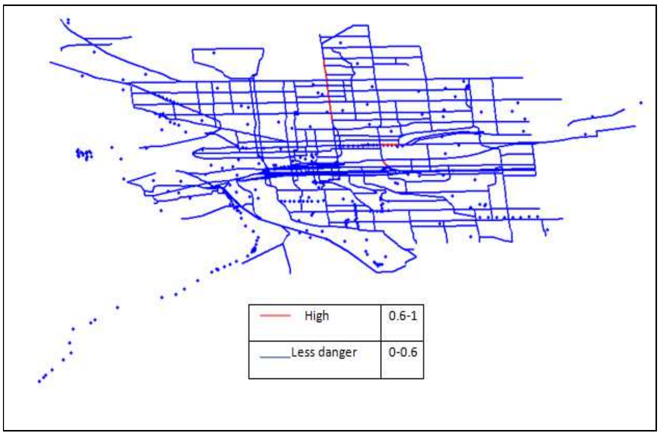

Fig. 7 The result of Bagging tree

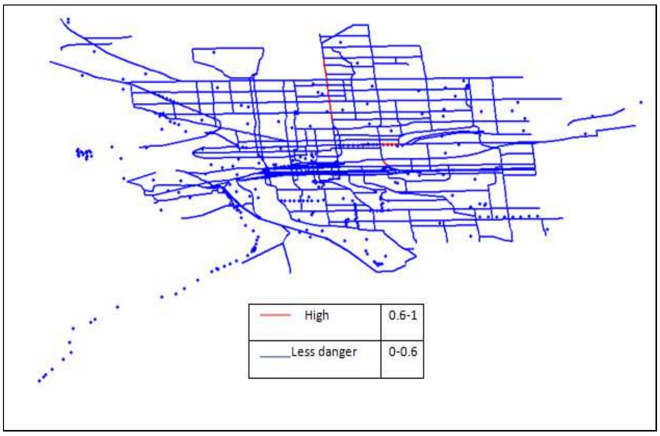

Fig. 8 The result of aggregation of ANN and Decision tree classifiers with Dempster-shafer theory

VI. CONCLUSION

The main objective of this study was to use some aggregation methods such as Bagging and Dempster-Shafer theories to improve the result of single classifiers such as ANN and Decision tree for optimum site selection of traffic cameras so the result classified road links in two classes, class with high risk and low risk, as a result links in high risk class could be selected for camera installation. So this study could help us to overcome the problem uncertainty for classification with single classifiers such as ANN and Decision tree.

REFERENCES

[1] N. Shoukry, "Artificial Neural Network in Classification of Severity Levels in Crashes with Guardrail", USA, 2005.
[2] L. Dong, "Using Bagging Classifier to Predict Protein Domain Structural Class", China,2006
[3] L. Breiman, "Bagging Predictors", USA, 2006.
[4] D. Koks, "An Introduction to Bayesian and Dempster-Shafer Data Fusion", Australia, 2005.
[5] J.Dezert,"Evidential Reasoning for Multi-Criteria Analysis based on DSmT-AHP", France, 2000.
[6] S.Lehegarat-Mascle,"Application of Dempster–Shafer Evidence Theory to Unsupervised Classification in Multisource Remote Sensing", IEEE VOL. 35, NO. 4, 1997.
[7] S.rokhsari,"Optimum site selection of traffic camera for efficient control of roads to decrease traffic incidents based on AHP and Voronoi diagram and fusion concepts", ITAC conference, Iran, 2012.