Specialized Ensemble of Classifiers for Traffic Sign Recognition

M.P. Sesmero[1], J.M. Alonso-Weber[2], G. Gutiérrez[1], A. Ledezma[1], and A. Sanchis[1]

Universidad Carlos III de Madrid, Computer Science Department
Avda. de la Universidad 30, 28911, Leganes, Madrid, Spain
[1]{msesmero,ggutierr,ledezma,masm}@inf.uc3m.es,
[2]jmaw@ia.uc3m.es
http://www.caos.inf.uc3m.es*

Abstract. Several complex problems have to be solved in order to build Advanced Driving Assistance Systems. Among them, an important problem is the detection and classification of traffic signs, which can appear at any position within a captured image. This paper describes a system that employs independent modules to classify several prohibition road signs. Combining the predictions made by the set of classifiers, a unique final classification is achieved. To reduce the computational complexity and to achieve a real-time system, a previous input feature selection is performed. Experimental evaluation confirms that using this feature selection allows a significant input data reduction without an important loss of output accuracy.

Keywords: Traffic Sign Recognition, Artificial Neural Networks, Feature Selection, Binary Classifier.

1 Introduction

Road signs carry essential information for successful driving: they define right-of-way, prohibit or permit certain maneuvers, warn about risky factors, etc. Therefore, for developing an autonomous Driver Support System (DSS), detection and classification of road signs are essential tasks. In spite of the increasing interest in the last years, traffic sign detection and classification are some of the less studied subjects in the field of Intelligent Transport Systems. Approaches in this area have been mainly focused on the resolution of other problems, such as road border detection [1, 2] or the recognition of obstacles in the vehicle's path such as pedestrians [3, 4] or other vehicles [3, 5].

In this work, we describe the implementation of a system whose main task is to classify prohibition road signs into several categories, once the traffic sign has been detected. In order to improve the classification performance, the system is composed by an ensemble of specialized classifiers. Each classifier is associated with a particular kind of road sign and its goal is to distinguish this sign from the others.

* The research reported here has been supported by the Ministry of Education and Science under project TRA2004-07441-C03-C02.

The system presented in this paper is a part of a Driver Support System. An important requirement is that it should be implemented in hardware in order to use it for real-time traffic sign categorization. In order to fulfill this constraint, an input feature selection step is needed.

This paper is organized as follows. Section 2 presents the problem overview. In section 3 the proposed architecture is explained. The experimental results are shown in section 4. Finally, in section 5, some conclusions and future work are summarized.

2 Problem Overview

The goal of the research presented in this paper is to build a system whose task is to classify prohibition road signs into several categories. This task can be described as a supervised learning problem in which the input information comes from a set of road signs arranged in a fixed number of categories (classes) and the goal is to extract, from the input data, the real knowledge needed to classify correctly new signs.

Since the signs to be classified are embodied in images captured with cameras, before applying learning techniques, an image preprocessing phase is required. This phase consists of a color thresholding, a border detection and extraction [6] and a subsequent normalization to 32x32 pixels. Then, a grayscale conversion is performed averaging each pixel's RGB components. Afterwards, learning examples (patterns) are generated. Each pattern is the result of transforming the normalized image into a 1024 element vector in which each grayscale pixel value is represented with a real number in the range [0.0 , 1.0].

The next step is to encode the class associated to each pattern. Since the classification of a sign is unique, the standard way of encoding it consists of using a vector C, which has as many components c_i, as existing classes. The component value c_i will be *1* if the sign belongs to class i, and *0* in any other case.

The selected learning method to solve the problem is a Multilayer Perceptron (MLP) trained with the Back Propagation algorithm (BP).

3 Our Approach

In a previous work [7] we showed that MLP is a good learning method to deal with the classification task described in the previous section. However, as we mentioned before, this work is a part of a complete DSS. In order to integrate this classification system into a DSS capable to perform real-time traffic sign categorization, a hardware implementation on FPGA (Field Programmable Gate Array) is necessary. Due to strict size limitations of Artificial Neural Networks (ANN) implementations on FPGAs [8], solving this classification problem with a unique MLP is not a feasible approach. With the aim of reducing both the classification problem complexity and the ANN size, an ensemble of specialized neural networks, combined with a feature selection process applied to the input data is proposed.

The general framework of this approach (shown in Figure 1) is composed of two modules: the *Data Preprocessing Module* (DPM) and the *Classification Module* (CLM).

The DPMs function is to select from among the 1024 attributes that describe a sign the subset that each specialized neural network inside the CLM must receive. On the other hand, the CLMs function is to classify each input data set as one of the available prohibition road sign-types. Since this module is composed of several independent classifiers, in order to obtain the final classification, an integration of the individual predictions is required.

In the following subsections each one of these modules are described.

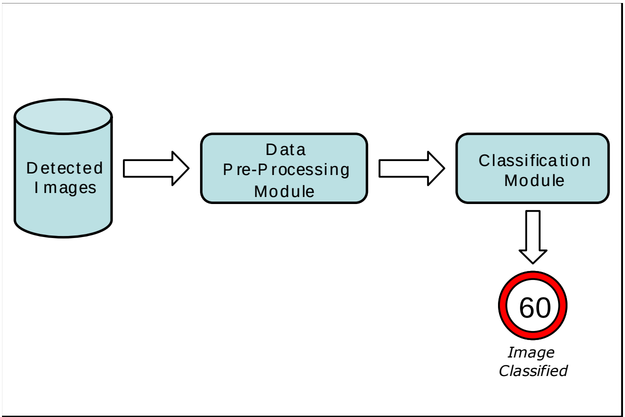

Fig. 1. General framework of the proposed system

3.1 DPM Construction

Practical experience shows that using as much as possible input information (features) does not imply higher output accuracy. Feature subset selection [9, 10] is the procedure of selecting just the relevant information, avoiding irrelevant and redundant information and reducing the learning task dimensionality.

The proposed architecture adopts a model in which the feature subset that describes an example is not unique but depends on the task associated to each classifier. In other words, since the classification problem is divided in n binary sub-problems, n feature selection procedures are necessary.

In this work, the feature selection module has been built using the Weka tool [9]. Subsets of attributes are obtained from the data by applying search algorithms (named *search method*), and each feature subset is evaluated by another algorithm (named *attribute evaluator*).

At first, several search-method/attribute-evaluator pairs from those included in Weka were considered. After analyzing the features sets obtained for each pair, three of them were selected, those that obtained an attribute subset according to the ANN size restrictions. The three selected search-method/attribute-evaluator pairs were:

- *Best First* [11] and *Correlation-based Feature Selection* (CFS) [10];
- *Ranker* (it ranks attributes using their individual evaluations) and "*GainRatioAttributeEval*" (it evaluates each attribute individually by measuring the gain ratio with respect to the class) and
- *Ranker* and "*InfoGainAttributeEval*" (it evaluates each attribute individually by measuring the information gain with respect to the class).

Although, the methods mentioned above, *a priori,* seem to be a good choice to resolve the Feature Selection process, experimental evaluation showed that the combination of Best First and CFS gives better results for this classification task [12]. For this reason, this method has been used as a base for the DPM construction.

After the DPM has been built, it is able to select the significant attributes for each binary classification task.

3.2 CLM Construction

In order to build the CLM, we propose an approach based on an ensemble of specialized ANN. In the next subsections, we present the details of the CLM construction.

3.2.1 Specialized ANN

The CLM is composed of a collection of binary classifiers where each of them is specialized in discriminating a specific road sign type from the others. Therefore, for a classification problem where *n* road sign types have to be separated, the system is composed of *n* Multilayer Perceptrons. The ease of the task associated to each neural network allows us, among other things:

- Designing minimal neural nets. The simpler the learning task is the fewer nodes will be needed in its design. Moreover, each NN can have its own architecture (number of hidden nodes, activation function, learning factor,…)
- Improving the convergence of the training process and, therefore, reducing the training time.
- Obtaining better classification performance compared with a single large classifier.
- Selecting the feature subset that, in each particular case, better captures the information included in the patterns to be classified.

To build these classifiers, a new output encoding schema is necessary. In this case, the class associated with each pattern is encoded using an output whose value is *1* if the example is a positive instance, or *0* if the example is a negative instance. This encoding is equivalent to select the c_i component from the previous codification.

3.2.2 Neural Network Topology

The neural networks architecture and topology were adjusted using experimental methods comprising constructive, pruning and analytical techniques.

The structure of the resulting neural networks is the following one:

- An input layer containing as many neurons as the number of relevant attributes associated with the class.
- One hidden layer. The number of neurons in this layer has been fixed to 50 when the net has 10 output neurons and to 36 when the net has a single output neuron. These numbers were determined by trial-and-error methods.
- An output neuron whose value is in the interval [*0,1*]. An input sign is classified as belonging to the encoded class when the corresponding output value is higher than 0.5. Otherwise, the input will be considered as a negative example (not belonging to the associated class).

3.2.3 Learning Phase

Once the structure of the neural networks and their input information are known, the next step consists of a training phase. Since there is no connection between the individual networks, the training can be performed distributing the work on several processors. This allows us to reduce the global system training time.

To overcome the overfitting problem and the influence of the initial weights, a 10-fold cross-validation approach was used which was run 10 times. A different weight initialization was performed on each run.

3.2.4 ANNs Output Combination

When the system receives an unlabeled road sign to be classified in some of the fixed categories, such sign is sent to each classifier's input module. The DPM selects the pixel subset according to its relevant attribute list. The chosen pixels are used as the input for the associated ANN, which applies its knowledge to make a prediction. The individual predictions are sent to the combining module that carries out an integration of the received information and produces a unique final classification.

The final classification can be established selecting the Neural Network with the highest output value or applying the following rules [13]:

- If for all classifiers the individual prediction is *<=0.5*, the example will be considered as an unclassified pattern.
- If the individual prediction of exactly one classifier is *>= 0.5* and the output value of all other classifiers is *<0.5*, then the example will be considered an instance of the class associated to the ANN with the maximum output value.

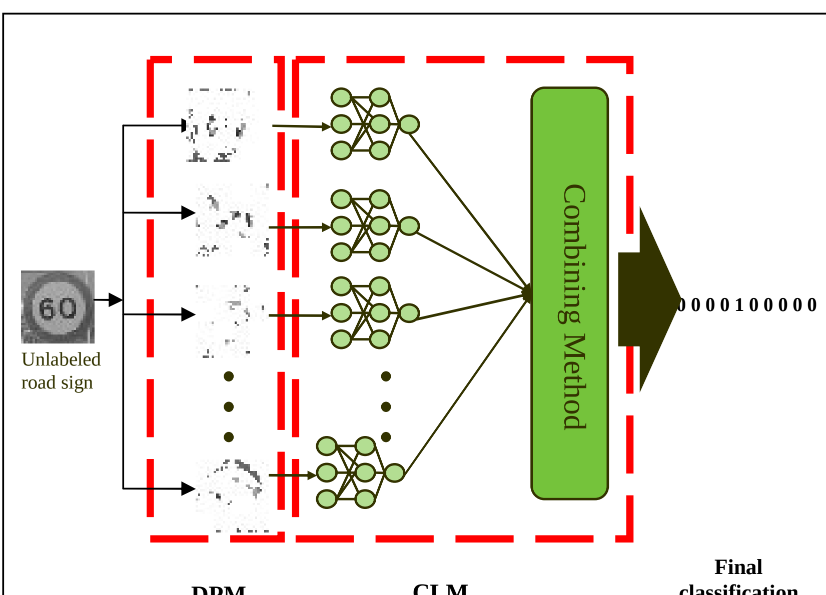

Fig. 2. Classification Process

- If there are two o more classifiers whose output value is *>=0.5*, the final classification will be established combining the individual predictions and each networks precision.

Experimentally it is found that, for the proposed classification task, the first criterion is more efficient than the second one. Therefore, the formula used in the combining module is:

$$F(\bar{x}, y_1, y_2, \ldots\ldots. y_n) = \arg\max_{i=1,\ldots,n}(y_i), \qquad (1)$$

where y_i is the output value of the neural network associated to the i^{th} class.

Figure 2 shows the classification process of an unlabeled road sign.

4 Empirical Evaluation

The whole system has been validated over 5000 examples arranged in ten generic kinds of prohibition road signs: no pedestrians, no left/right turn ahead, no stopping and no parking, no passing, and 20-30-40-50-60 and 100 km speed limits.

In order to evaluate our approach, first we train a unique ANN with 1024 (32 × 32) input nodes, one hidden layer with 50 neurons and one output layer with 10 neurons, (experiment 1).

In a second step, we carried out a set of experiments (experiment 2) in which the previous net is split into ten binary ANN in order to divide the global signal recognition problem into smaller individual signal recognition tasks. Using this kind of ensembles should improve the classification accuracy and allow a size reduction of the hidden layers.

At last, a third kind of experiment (experiment 3) was carried out reducing the input neuron number as much as possible in order to fulfill the hardware implementation requirement. In this case, the ensemble of classifiers contains 10 binary neural networks with 36 hidden nodes in each. The input unit number is equal to the attribute number selected by the DPM. This number is shown in Table1.

Once the relevant input features associated to each class were known, the nets were built, trained and tested according to them.

In Table 2 we show the accuracy for the described experiments when a 10-fold cross validation process is used.

Table 1. Number of Selected Features. In the first column appears the label of each class.

Class	Prohibition road sign	Number of selected features
C1	no pedestrians	116
C2	no (left, right) turn ahead	91
C3	stopping and no parking	44
C4	no passing	114
C5	60 km speed limit	114
C6	50 km speed limit	110
C7	40 km speed limit	100
C8	30 km speed limit	114
C9	20 km speed limit	103
C10	100 km speed limit	87

Table 2. Train Set Accuracy

Experiment	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	Global
1	100	100	100	99.96	98.00	99.81	99.53	98.98	99.27	100	99.56
2	100	100	100	100	97.69	99.56	99.54	98.80	99.67	100	99.52
3	100	99.83	99.28	99.93	94.98	95.19	97.81	95.80	97.24	99.95	98.00

Table 3. Test Set Accuracy

Experiment	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	Global
1	99.96	99.38	99.22	98.10	95.58	97.48	94.64	98.22	95.90	98.94	97.74
2	100	99.44	99.72	98.46	96.06	98.48	95,62	98.36	96.28	99.02	98.14
3	99.74	98.72	97.46	98.22	93.06	93.58	92.78	95.52	94.66	97.96	96.17

The experimental evaluation reflects that dividing the classification task into binary subtasks increases slightly the classification accuracy on the test set.

On the other hand, the loss of classification accuracy when the feature selection process is performed is not very significant compared with the benefits of the drastic input data reduction.

5 Conclusions and Future Work

In this paper, an architecture for traffic sign classification has been described. Very high recognition rates have been obtained when the global learning task is divided into simple tasks and, for each task, *Best First* and *Correlation-based Feature Subset Selection* are applied, as a *relevant feature selection method*.

The features of this architecture make it possible to implement this system on FPGAs and to use it in real time problems.

The future work will be mainly focused on extending the system in order to cope with regulatory, warning, indication, etc, signs. This task will allow us to investigate and develop new procedures that will contribute to the design of a more versatile system.

References

1. Dickmanns, E.D., Zapp, A.: A Curvature-Based Scheme for Improving Road Vehicle Guidance by Computer Vision. In: Proceedings of the SPIE Conference on Mobile Robots, pp. 161-168 (1986)
2. Pomerleau, D., Jochem, T.: Rapidly Adapting Machine Vision for Automated Vehicle Steering. IEEE Expert Magazine 11(2), 19–27 (1996)
3. Franke, U., Gavrila, D., Görxig, S., Lindner, F., Paetzold, F., Wöhler, C.: Autonomous Driving Goes Downtown. IEEE Intelligent Systems 13(6), 40–48 (1998)
4. Handmann, U., Kalinke, T., Tzomakas, C., Werner, M., Seelen, W.: An Image Processing System for Driver Assistance. Image and Vision Computing 18(5), 367–376 (1999)
5. Bertozzi, M., Broggi, A.: GOLD: A Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection. IEEE Transactions on Image Processing 7(1), 62–81 (1998)

6. Escalera, A., Moreno, L.E., Salich, M.A., Armingol, J.M.: Road Traffic Sign Detection and Classification. EEE Transactions on Industrial Electronics 44(6), 848–859 (1997)
7. Granados, A., Ledezma, A., Gutierrez, G., Sanchis, A.: Reducing the Amount of Input Data in Traffic Sign Classification. Modeling Decisions for Artificial Intelligence (2006)
8. Zhu, J., Sutton, P.: FPGA Implementations of Neural Networks – A survey of a Decade of Progress. In: Field-Programmable Logic and Applications. Lecture Notes in Computer Science, vol. 2778, pp. 1062-1066 (2003)
9. Witten, I.H., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, Washington (2005)
10. Hall, M.A.: Correlation-based Feature Selection for Machine Learning. Ph.D diss. Hamilton, NZ: Waikato University Department of Computer Science (1998)
11. Russell, S.J., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice Hall, Englewood Cliffs (2003)
12. Sesmero, M.P., Alonso-Weber, J.M., Gutiérrez, G., Ledezma, A., Snachis, A.: Testing Feature Selection in Traffic Signs. In: Proceedings of 11th International Conference on Computer Aided Systems Theory, pp. 396-398 (2007)
13. Ou, G., Murphey, Y.L.: Multi-Class Pattern Classification using neural Networks. Pattern Recognition 40(1), 4–18 (2007)