Ensemble of ANN for Traffic Sign Recognition

M. Paz Sesmero Lorente
Universidad Carlos III de Madrid, Spain

Juan Manuel Alonso-Weber
Universidad Carlos III de Madrid, Spain

Germán Gutiérrez Sánchez
Universidad Carlos III de Madrid, Spain

Agapito Ledezma Espino
Universidad Carlos III de Madrid, Spain

Araceli Sanchis de Miguel
Universidad Carlos III de Madrid, Spain

INTRODUCTION

"Machine Learning (ML) is the subfield of Artificial Intelligence conceived with the bold objective to develop computational methods that would implement various forms of learning, in particular mechanisms capable of inducing knowledge form examples or data" (Kubat, Bratko & Michalski, 1998, p. 3).

The simplest and best-understood ML task is known as supervised learning. In supervised learning, each example consists of a vector of features (**x**) and a class (y). The goal of the learning algorithm is, given a set of examples and their classes, find a function, *f*, that can be applied to assign the correct class to new examples. When the function *f* takes values from a discrete set of classes $\{C_1, \ldots, C_K\}$, *f* is called a *classifier* (Dietterich, 2002).

In the last decades it has been proved that learning tasks in which the unknown function *f* takes more than two values (multi-class learning problems) the better approach is to decompose the problem into multiple two-class classification problems (Ou & Murphey, 2007) (Dietterich, & Bakiri, 1995) (Massulli & Valentini, 2000).

This article describes the implementation of a system whose main task is to classify prohibition road signs into several categories. In order to reduce the learning problem complexity and to improve the classification performance, the system is composed by a collection (ensemble) of independent binary classifiers. In the proposed approach, each binary classifier is a single-output neural network (NN) trained to distinguish a particular road sign kind from the others.

The proposed system is a part of a Driver Support System (DSS) supported by the Spanish Government under project TRA2004-07441-C03-C02. For this reason, one of the main system requirements is that it should be implemented in hardware in order to use it aboard a vehicle for real time categorization. In order to fulfill this constraint, a reduction in the number of features that describe the instances must be performed. As consequence if we have *k* generic road sign types we will use *k* binary NN and *k* feature selection process will be executed.

BACKGROUND

It is known that road signs carry essential information for safe driving. Among other things, they permit or prohibit certain maneuvers, warn about risk factors, set speed limits and provide information about directions, destinations, etc. Therefore, road sign recognition is an essential task for the development of an autonomous Driver Support System.

In spite of the increasing interest in the last years, traffic sign recognition is one 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 (Dickmanns & Zapp, 1986) (Pomerlau & Jochem, 1996) or the recognition of obstacles in the vehicle's path such as pedestrians (Franke, Gavrilla, Görxig, Lindner, Paetzold & Wöhler, 1998) (Handmann, Kalinke, Tzomakas, Werner & Seelen, 1999) or other vehicles (Bertozzy & Broggi, 1998).

When the number of road sign types is large, road sign recognition task is separated in two processes: detection and classification. Detection process is responsible for the localization and extraction of the potential signs from images captured by cameras. Only when the potential signs have been detected they can be classified as one of the available road sign-types.

In the published researches, detection is based on color and/or shape of traffic signs (Lalonde & Li, 1995). On the other hand, to solve the classification task several ML algorithms have been used. Among the used techniques it is worth mentioning: The Markov Model (Hsien & Chen, 2003), Artificial Neural Networks (Escalera, Moreno, Salich & Armingol, 1997) (Yang, Liu & Huang, 2003), Ring Partitioned Method (Soetedjo & Yammada 2005), the Matching Pursuit Filter (Hsu & Huang 2001) or the Laplace Kernel classifier (Paclík, Novovicová, Pudil, & Somol, 1999).

A NEURAL NETWORK BASED SYSTEM FOR TRAFFIC SIGN RECOGNITION

In this work, we present the architecture of 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.

The proposed system is a Multilayer Perceptron (MLP) based classifier trained with the Back-Propagation algorithm. In order to integrate this classification system into a DSS capable to perform real-time traffic sign categorization, a hardware implementation on Field Programmable Gate Array (FPGA) is necessary.

With the aim of reducing the problem complexity, an ensemble of specialized neural networks is proposed. In addition and due to the strict size limitations of ANN implementation on FPGAs (Zhu & Sutton, 2003) the construction of each specialized MLP is combined with a specific reduction in the number of features that describes the examples.

Traffic Sign Pre-Processing

Since the signs to be classified are embodied in images acquired by a camera attached to a moving vehicle, it can be assumed that the signs have a varying size (signs get bigger as the vehicle moves toward them). Therefore, once the traffic signs have been detected, the first step is to normalize them to a specific size. The aim of this process is to ensure that all the signs (examples) are described by the same number of pixels (features). In our approach we have used 32x32 pixel signs.

Once the signs have been normalized, a grayscale conversion is performed. Since the original images are represented in the RGB (Red, Green and Blue) color space, this conversion is done by adding the red, green and blue values for each pixel and dividing by three. As result of both processes, each road sign is transformed into a 1024 element vector in which each pixel is represented by a real number in the range [0.0, 1.0].

System Architecture

The general framework of the proposed system (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.

To build both, the DPM and the CLM, a new data encoding schema is necessary. In particular, the multiclass problem has to be decomposed into a set of binary subproblems.

Data Preprocessing Module

Practical experience shows that using as much as possible input information (features) does not imply higher output accuracy. Feature subset selection (Witten & Frank, 2005) (Hall, 1998) is the procedure of selecting just the relevant information, avoiding irrelevant and redundant information and reducing the learning task dimensionality.