Available online at www.sciencedirect.com

ScienceDirect

Expert Systems with Applications 36 (2009) 1391–1402

Expert Systems with Applications

www.elsevier.com/locate/eswa

Selecting of the optimal feature subset and kernel parameters in digital modulation classification by using hybrid genetic algorithm–support vector machines: HGASVM

Engin Avci *

Firat University, Department of Electronic and Computer Education, 23119 Elazig, Turkey

Abstract

The support vector machines is a new technique for many pattern recognition areas. The digital modulation classification is one of these pattern recognition areas. In SVM training, the kernels, kernel parameters, and feature selection have very important roles for SVM classification accuracy. Therefore, most appropriates of these kernel types, kernel parameters and features should be used for SVM training. In this study, a hybrid of genetic algorithm–support vector machines (HGASVM) approach is presented in digital modulation classification area for increasing the support vector machines (SVM) classification accuracy. This HGASVM approach proposed in this paper selects of the optimal kernel function type, kernel function parameter, most appropriate wavelet filter type for problem, wavelet entropy parameter, and soft margin constant C penalty parameter of support vector machines (SVM) classifier. The classification accuracy of this HGASVM approach is tried by using real digital modulation dataset and compared with the SVMs, which has kernel function type, kernel function parameter, wavelet filter type, wavelet entropy parameter, and C parameter are randomly selected. Here, discrete wavelet transform (DWT) and adaptive wavelet entropy are used in feature extraction stage of this HGASVM approach. The digital modulation types used in this study are *ASK-2, ASK-4, ASK-8, FSK-2, FSK-4, FSK-8, PSK-2, PSK-4, and PSK-8*. The experimental studies conducted in this study show that the classification accuracy of this HGASVM approach is more superior than SVM, which has constant parameters.

Keywords: Digital modulation classification; Pattern recognition; Support vector machines; Genetic algorithm; Parameter optimization; Kernel function; Intelligent system; DWT; Wavelet entropy

1. Introduction

Nowadays, automatic modulation classification (AMC) have been become very important topic. The automatic classification of digital modulated signals is used for many areas such as civilian and military communication systems, signal interception, electronic warfare, radar target recognition (Mustafa & Doroslovaki, 2004; Wu, Wang, Gao, & Ren, 2005). So far, many modulation classification studies have been presented. The some of most important of these studies are conducted by Nandi and Azzouz who have suggested analog modulation classification algorithms, digital modulation classification algorithms, and analog and digital modulation classification algorithms by using decision-theoretic and artificial neural network (ANN) (Wu et al., 2005; Azzouz & Nandi, 1996a, 1996b). In other automatic modulation classification studies, different methods such as statistical pattern classification, spectrum monitoring applications, and intercepted signal filtering are used for automatic modulation classification (Dubuc & Boudreau, 1999; Lopatka & Pedzisz, 2000; Wong & Nandi, 2001; Wu et al., 2005).

The ANNs is used for automatic modulation classification in many AMC studies. Nandi and Azzouz (1997) suggested a single hidden layer ANN structure for automatic

* Tel.: +90 4242370000; fax: +90 4242367064.
E-mail address: enginavci@firat.edu.tr

0957-4174/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2007.11.014

modulation classification. This ANN has a 4-node input layer, a 25-node hidden layer and a 7-node output layer. The most important drawback of ANNs is difficulties of local minimums. ANN cannot estimate one global minimum due to the initial weights is randomly selected. On the other hand, the SVM is trained as a convex optimization problem, which can be estimated a global solution (Fernández Pierna, Baeten, Michotte Renier, Cogdill, & Dardenne, 2004). Namely, while the ANNs suffer from local minimums, the SVMs do not. The SVM can overcome the overfitting much successfully than ANN (Osareh, Mirmehdi, Thomas, & Markham, 2002). Therefore, SVM is preferred to ANN for automatic digital modulation classification (ADMC) in this study.

In modulation classification literature, the SVM classifiers are used in several studies (Gang, Jiandong, & Donghua, 2004; Mustafa & Doroslovaki, 2004; Wu et al., 2005). In reference Gang et al. (2004), it is presented a new algorithm for digital modulated signal classification based on Higher Order Cumulants (HOCs) and SVM. In reference Wu et al. (2005), it is suggested a method based on SVM for digital modulation signal classification in the presence of additive white Gaussian noise. In reference Mustafa and Doroslovaki (2004), it is proposed four features to classify ASK-2 and ASK-4 digital modulation classification.

Discrete wavelet transform (DWT) types are commonly used in many areas such as signal processing, image processing, speech processing, computer vision, and radar target classification owing to many advantages of DWT (Wu, Ren, Wang, & Zhao, 2004). In this study, DWT is used for obtaining the compact set of features which capture all the major characteristics of the modulated signals in a relatively small number of the components.

The using of SVM has three difficulties, which are how to select the optimal kernel function type, most appropriate kernel function parameter value, and best feature subset for SVM training and testing stages (Huang & Wang, 2006). The solutions of these difficulties are very important because of the optimization of kernel function type, kernel function parameters, and feature subset of SVM influence the classification accuracy of this SVM (Fröhlich, Chapelle, & Scholkopf, 2003; Huang & Wang, 2006). For this reason, optimal kernel parameters, and feature subset of SVM should be estimated simultaneously (Huang & Wang, 2006).

So far, the optimization of SVM parameter and feature subset have been realized in some studies (Fröhlich et al., 2003; Jack et al., 2002; Mao, Zhou, Pi, Sun, & Wong, 2005; Pai, 2006; Pai & Hong, 2005; Wu, Tzeng, Goo, & Fang, 2007). In these studies, the penalty parameter C, kernel function parameters, and feature subset are generally optimized by using various techniques. The first, an appropriate kernel function for SVM should be selected. Second, the kernel function parameters such as the sigma (σ) should be set. Third, value of a soft margin constant C penalty parameter should be determined.

So far, many techniques have been used for optimization of these parameters. The most commons of these optimization techniques are Grid algorithm and genetic algorithm. The Grid algorithm is a method for estimating the appropriate C and sigma when using the radial basis function (RBF) kernel function. Nevertheless, this Grid algorithm method is time consuming and does not perform well (Hsu & Lin, 2002; Huang & Wang, 2006; LaValle & Branicky, 2002). In addition to the Grid algorithm method is not able to the feature selection task (Huang & Wang, 2006). On the other hand, the genetic algorithms are able to perform both optimal feature extraction and optimization of SVM parameters simultaneously (Jack et al., 2002; Huang & Wang, 2006; Mao et al., 2005; Pai, 2006; Pai & Hong, 2005; Wu et al., 2007). In this area, there are a few studies in SVM literature (Bradley & Mangasarian, 1998; Bradley, Mangasarian, & Street, 1998; Guyon, Weston, Barnhill, & Bapnik, 2002; Huang & Wang, 2006; Mao, 2004; Weston et al., 2001). In some of these studies, the genetic algorithms are used for optimum feature extraction of SVM, but many of in these studies is not performed the SVM parameter optimization (Huang & Wang, 2006; Raymer, Punch, Goodman, Kuhn, & Jain, 2000; Salcedo-Sanz, Prado-Cumplido, Pe'rez-Cruz, & Bousoño-Calzo'n, 2002; Yang & Honavar, 1998).

On the other hand, optimal kernel type and kernel parameters, and feature subset of SVM are estimated simultaneously in this study. Therefore, the HGASVM approach presented in this study is comprehensive than algorithms used in other genetic algorithm-SVM studies. This state is one of superiorities of this study than other previous studies performed in genetic algorithm-SVM area. The other different of this study than others is the firstly using of HGASVM in automatic digital modulation classification (ADMC) area. In addition to, the other advantages of this study than other studies performed in genetic algorithm-SVM area can summarize as follows:

- The effectiveness of the DWT features used for ADMC feature extraction in this study.
- The using of wavelet entropy (WE), which effectively reduces size of feature vector is used for increasing the effectiveness of HGASVM proposed in this study.

In HGASVM structure, 16 differently wavelet decomposition filter types, which are db2, db4, db5, db10, sym2, sym3, sym5, sym8, bior1.3, bior2.2, bior3.5, bior6.8, coif1, coif2, coif3, and coif5, eight differently kernel function types, which are linear, polynomial, radial basis function (RBF), sigmoid, spline, Fourier, bspline, and exponential radial basis function (ERBF) are used.

This paper is organized as follows: In Section 2, DWT and Wavelet Entropy (WE) are explained. In Section 3, SVM classifier is briefly reviewed. In Section 4, the proposed method based on HGASVM is presented. In Section 5, experimental results are given. In Section 6, conclusions are given.