Ford vehicle identification via shallow neural network trained by particle swarm optimization

Jingyuan Yang1,2
1The Faculty of Computer Science and Engineering, Xi’an University of Technology, Xi’an 710048, China; 2Department of Informatics, University of Leicester, Leicester LE1 7RH, UK

Lei Wang1,*, Qiaoyong Jiang1
1The Faculty of Computer Science and Engineering, Xi’an University of Technology, Xi’an 710048, China; wangleei@163.com (Lei Wang)

Abstract—Automatic identification of the car manufacturer is difficult to achieve because of the similarity among the different brands. In this work, we propose a new system of Ford vehicle identification. Firstly, we captured the side view of the car image. Secondly, we employed the wavelet entropy (WE) to extract efficient features from car images. Thirdly, we employed a shallow neural network (SNN) as a classifier. Finally, we used the particle swarm optimization to train the classifier. The 10 10 - fold cross validation on a dataset containing 220 vehicle images showed that our Ford vehicle identification system obtained the overall sensitivity of 83.27±1.61%. The overall specificity is 83.91±1.87%, the overall accuracy is 83.59±0.94%. Experiment result show that the proposed system is effective for Ford vehicle identification.

Keywords—Ford vehicle; identification; wavelet entropy; shallow neural network; particle swarm optimization; cross validation

I. INTRODUCTION

Recently, the amount of traffic accidents is continue increasing [1, 2]. How to extract car images from surveillance video is a major problem after accidents. However, most studies in the field of extract images have only focused on the front view and are too difficult for us.

A search of the literature revealed few studies which identify the car brand from the side view [3]. Hence, there is still a huge research space in extracting images of the side of cars, so we decide to do some studies in this space.

There are various studies using various method to extract image features, Mustapha, Mohammed (2012) [4] used local binary patterns (LBP), but it only covers a small area within a fixed radius, which can’t satisfy the needs of different size and frequency textures. Ghoualmi, Draa (2016) [5] employed scale invariant feature transform (SIFT) to extract feature, but this method will cause the feature matching spread error, and the matching will be failing. To overcome these limits, a new method for this problem is proposed in this paper.

In this paper, the main contribution is to propose a novel Ford vehicle identification system. We employed wavelet-entropy (WE), which is a relatively accurate feature descriptor, to extract efficient features from the side view of car images [6]. Meanwhile, we used a shallow neural network (SNN) as the classifier. We also utilized cross validation (CV) to verify the classifier performance. Then, we employed particle swarm optimization (PSO) for classifier training because of its excellent performance.

II. METHOD

A. Wavelet Entropy

We employed WE to extract features from car images. WE combines wavelet transform and entropy calculation [7]. The continuous wavelet transform (CWT) is defined as follows:

$$C_{\psi}(k,j) = \int_{-\infty}^{+\infty} x(t) \psi(\frac{t - k}{j}) \, dt$$  \hspace{1cm} (1)

where

$$\psi(t; k, j) = \frac{1}{\sqrt{|k|}} \psi(\frac{t - k}{j}), k \neq 0, j \in R$$  \hspace{1cm} (2)

C represents the coefficients of CWT. The wavelet $\psi(t; k, j)$ is generated by translation and dilation, using the scale and translation factor $k$ and $j (k, j \in R^+)$, of the parent wavelet $\psi(t)$. The wavelets are discretely sampled is called discrete wavelet transform (DWT). Eq. (1) can be discretized by restraining $k$ and $j$ to a discrete lattice $(k = 2^l \& j > 0)$ to give the DWT which can be defined as follows:

$$L(n|a,b) = DS[\sum_n x(n) \delta^b (n - 2^a)]$$  \hspace{1cm} (3)

$$H(n|a,b) = DS[\sum_n x(n) \delta^b (n - 2^a)]$$

Figure 1 Diagram of two-level wavelet entropy (WE)
Here \( L \) and \( H \) represent the coefficients of the approximation components and detail components respectively. \( l \) and \( h \) represent the low pass and high pass filters respectively. \( a \) and \( b \) stand for the wavelet scale and translation factors, respectively. The DS is regards as the downsampling operator. In this paper, the DWT is employed to extract the feature of the side view of cars images because of the scale factor and the translation factor are dispersed [8-12]. Figure 2 shows the diagram of two-level wavelet entropy. First, generated four sub-bands (LL1, LH1, HL1, and HH1), and the LL1 was further decomposed into four smaller sub-bands. Then, we acquired seven sub-bands for a two level decomposition.

### B. Shallow Neural Network

![Figure 2: Structure of the shallow neural network](image)

After reducing and extracting the image features, we put them into the shallow neural network (SNN). The reason we employed SNN is as follows: (1) it has been widely applied to pattern classification; (2) it can avoid complicated training. The model of SNN is shown in Figure 2.

Shallow neural network include three layers: an input layer, an output layer, and only one hidden layer \([13-16]\). Two nodes of each adjacent layer are connected completely. Each link is assigned with a weight value that is stand for the relational degree between two nodes. Sigmoid and linear functions are applied to the activation functions for hidden layer and output layer, respectively.

### C. Particle Swarm Optimization

PSO is a swarm based meta-heuristic algorithm which is inspired by the social behavior of a swarm of birds, flocking bees, and fish schooling \([17-21]\). PSO algorithm can find the optimal solution by the cooperation and information sharing among individuals in the group. Due to its simple structure and easy implementation, PSO has captured much attention and has been has been widely used in functional optimization, neural network training, clustering, and so on. Next, the process of PSO is presented in brief. Figure 3 gives a pipeline of PSO.

To start the optimization process, an initial population is randomly generated \([22-25]\). Then find the optimal solution by iteration. In each iteration, the particle updates itself by tracking two best positions (\( p_{\text{best}}, g_{\text{best}} \)). After finding the two optimal values, the particle’s speed and position were updated as follows:

\[
v_{i} = \omega \times v_{i} + c_{1} \times \text{rand()} \times (p_{\text{best}} - x_{i}) + c_{2} \times \text{rand()} \times (g_{\text{best}} - x_{i}) \tag{4}
\]

\[
x_{i} = x_{i} + v_{i} \tag{5}
\]

Here, \( \text{rand()} \) is a random number between 0 and 1, \( c_{1} \) and \( c_{2} \) are positive constant, called “acceleration coefficients”. \( \omega \) is called the “inertia weight” which controlled the impact of the previous velocity of the particle on its current one. If \( \omega > 1 \), the particle favors exploration than exploitation, if \( \omega < 1 \), it favors exploitation than exploration. The PSO algorithm iteratively traverses these processes until the termination criterion was satisfied. In the future, we shall test other swarm intelligence methods \([26-28]\).

### D. Cross Validation

To improve the generation performance of the classifier, we employed the cross validation technique on the training set \([29-32]\). CV is a statistical analysis method used to assess the statistical relevance of the classifiers. The basic idea of CV is to group the datasets in a sense, one part as a training set and the other part as a validation set. First, the training set is applied to train the classifier. Second, the validation set is applied to test the model obtained by training, which is used as the performance index of the evaluation classifier.

There are three common cross validation methods:
random subsampling, K-fold cross validation (K-fold CV), leave-one-out validation. In this paper, we used the K-fold CV because of its simple and easy properties, and used all the data for training and validation. The dataset is divided into K groups, one is set as a validation set and the rest K-1 are used as training set [33-36]. In this way, the cross-validation process would be repeated k times, finally, average the error rates of k experiments to obtain the final evaluation result.

It is important to decide the number of folds which would influence the computation time, bias of the estimator and the variance of the estimator. If K was too large, the bias of the estimator would be small but the variance of the estimator would be large, causing the computation time is too long. On the contrary, if K was too small, the computation time would reduce, the variance of the estimator would be small, but the bias of the estimator would be large [37]. In this method, we assigned K as 10 considering the best compromise between computational cost and reliable estimates.

III. DATASET

220 vehicle images are obtained, 110 are Ford vehicles, and the rest 110 images include Buick, Shanghai Volks, Hyundai, and Toyota. Two senior cameramen with more than five-year experiences were requested to shoot the vehicles from side view.

We removed the background and left the vehicle in the center of the picture by watershed algorithm and manual revision. Each image was resized to 256×256. The color car images were converted to gray-level images, i.e., we discard the color information. Several samples of preprocessed dataset are listed in Figure 4.

IV. EXPERIMENTS AND RESULTS

A. Parameter Setting

We used trial-and-error method to determine the parameters in this method. The final chosen parameters are listed in Table 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelet Family</td>
<td>db2</td>
</tr>
<tr>
<td>Wavelet decomposition</td>
<td>4</td>
</tr>
<tr>
<td>Number of hidden neurons</td>
<td>11</td>
</tr>
<tr>
<td>Activation function</td>
<td>Sigmoid</td>
</tr>
<tr>
<td>Acceleration coefficients</td>
<td>(c_1 = c_2 = 1)</td>
</tr>
<tr>
<td>Maximum iteration</td>
<td>200</td>
</tr>
</tbody>
</table>

B. Statistical results of our method

Here the sensitivities, specificities, and accuracies of 10 runs of 10-fold are listed below in Table 2, Table 3, and Table 4, respectively. The overall sensitivities of identifying a Ford motor as a Ford motor is 83.27±1.61%. The overall specificities, i.e., identifying a non-Ford motor as a non-Ford motor is 83.91±1.87%. The overall accuracies, i.e., identifying an image as its corresponding class is 83.59±0.94%.

C. Structure of Neural Network

There are 13 input neurons from 4-level wavelet decomposition, and there is merely one output neuron, with value true indicating Ford or false indicating non-Ford. We here use grid searching method to validate the number of hidden neurons as 11 is the optimal for our system. Let the number of hidden neurons change from 5 to 15. The results are shown in Figure 5, were we can observe 11 hidden neurons yield the best performance.
classify different vehicles into different brands; (2) test other
effectiveness.

In the future, we will carry out the following researches:

This work is partly supported by the National Natural

V. CONCLUSION

In this study, we developed a Ford vehicle identification
method based on SNN and PSO. The result showed its
effectiveness.

In the future, we will carry out the following researches:
(1) find a better combination of the method that can be used to
classify different vehicles into different brands; (2) test other
advanced wavelet transforms to reduce the computation time;
(3) test other swarm intelligence methods; (4) collect more car
images.

ACKNOWLEDGMENT

This work is partly supported by the National Natural
Science Foundation of China under Project Code (61803301,
61272283, 11361001, 61573281, U1334211), the China
Postdoctoral Science Foundation (2014M562435), and
the Natural Science Research Program of the Educational Office
of Shaanxi province (15JK1518).

REFERENCE

1. Zhang, G., et al., Traffic accidents involving fatigue driving and their
extent of casualties. Accident Analysis & Prevention, 2016. 87: p. 34-42.

2. Thomas, R.W., et al. Toward detecting accidents with already available
passive traffic information. in IEEE 7th Annual Computing and

Image Based on Convolutional Neural Network. in 18th International
Conference on Intelligent Data Engineering and Automated Learning

4. Mustapha, D., et al., Multi-Metric Based Face Identification with Multi
Configuration LBPH Descriptor: International Journal of Image, Graphics

5. Ghoualmi, L., et al., An ear biometric system based on artificial bees and
the scale invariant feature transform. Expert Systems with Applications,
2016. 57: p. 49-61

6. Atananga, A., Application of stationary wavelet entropy in pathological
3701-3714

Heisenberg's uncertainty principle and Shannon's entropy criterion to
monitor power system small signal oscillations. IEEE Transactions on

8. Li, Y.-J., Single slice based detection for Alzheimer's disease via wavelet
entropy and multilayer perceptron trained by biogeography-based
10393-10417
34. Rao, R.V., Abnormal Breast Detection in Mammogram Images by Feed-forward Neural Network trained by Jaya Algorithm. Fundamenta Informaticae, 2017. 151: p. 191-211