A large-scale and global car dataset for verification

Lingji Hu[1], Xingcheng Luo[2]
School of Information and Software Engineering
University of Electronic Science and Technology of China
Country Chengdu, China P.R.C
{2013220302009[1], 2014220601014[2]}@std.uestc.edu.cn

Jianhua Deng[3], Fengjie Lai[4], Jian Hu[5], Yongbin Yu[6]
School of Information and Software Engineering
University of Electronic Science and Technology of China
Country Chengdu, China P.R.C
{Jianhua.deng[3], 2013220202014[4], ybyu[6]}@uestc.edu.cn

Abstract—Few researches focus on the larger-scale car model dataset compared with other objects in computer vision for verification, such as classification and face verification. In this paper, we present our on-going effort in collecting a large-scale and global dataset, *Gcars*, for improving related car model research. This dataset contains not only the most famous global car models but also the most locals in China, where all car images are collected from public website and the car hierarchy is a three-layer, make, model and type. We also demonstrate the most important application, car verification, exploiting the dataset. The performance in terms of verification accuracy is better than that of the benchmark dataset, namely *CompCars*, which is the similar and famous dataset, by using the deep learning framework Caffe.

Keywords—car model, dataset, Gcars, car verification

I. INTRODUCTION

Car, one of the most important innovation by human, brings us a large convenience and flexibility of moving from one place to another, and also gives significant contributions to community and society. Nowadays, car is not only employed as vehicle for traveling but also as a tool to make someone lives, and may to help project someone's economic status, or reflects our wealth. Furthermore, the car has evolved an interesting topic among some car fans in the world. Generally speaking, the car is required to have strong reliability, practicality, and high comfort. The car manufactured is very larger each year, for example, it is more than 90 million in 2015 including thousands models and types. The large quantity of car offers a potential to foster more sophisticated and robust computer vision models and algorithms.

There are some biggest challenges to propose novel research algorithms for object categorization and recognition for car as it presents several unique characteristics over other objects. While as the same reason, it also facilitates a range of new research topics in object categorization. Specially, one challenge of fine-grained task is enabled of the large number models as most other categories do not have such. Luckily, there exists two hierarchies for the car category: make, mode and released type; and make, mode and released year. Both can indicate a direction to address the fine-grained task in a hierarchical way.

The categorization and verification based on fine-grained can be employed in intelligent transportation system such as indexing, description and so on. For instance, the fine-grained car categorization can be used to develop some toll systems in highway based on the vehicle types. The car verification is used to track a car between various cameras in a network for the video application, and also can be employed in other applications. For example, it can help people to recognize a car over a photo by a car verification system where much detail information may be represented, which offers benefit both for manufacturers and consumers.

Despite the huge study and practical interest, a few attentions are attached of the car model analysis in the computer vision domain. It is obviously that the limitation of lacking high quality and large quantity dataset is the obstacle of exploration for the community in this domain. For ending this, we collect and organize a large-scale and global car image database called "Global Cars", with "Gcars" in short. The dataset of Gcars is very large scale compared with most current car image datasets, which includes 234,187 images of 2,237 car models from public website. In addition, the dataset contains a lot of Chinese local car models, and all images are carefully labelled with car-make-model-type. This new global car dataset provides a platform for validation the effectiveness of a wide range for computer vision algorithms, and is ready to be utilized for proposing some novel research topics and algorithms.

The remainder of this paper is organized as follows. Section II introduces some related work of current car datasets and in section III describes the organization principle of Gcars and some details, the application of car verification will be introduced in section IV, and some experiment results will be reported by employing the deep learning framework Caffe in section V. Finally, some conclusions will be shown in section VI.

II. RELATED WORKS

Recently, there are some researches have been carried about car model. In [28], an evolutionary computing framework is proposed which is used to meet a wireframe model for recognizing the car model based on car image dataset. Some research about car model also presented, such as tracking over videos [17], detection [22], hybrid pose estimation and detection [25], and so on. There are some researches about three dimensional car model. An approach

about using 2D car images to construct 3D curves for determining alignment car model is proposed in [7]. Jointly 3D

Fig. 1. The different overviews of a car in the Gcars dataset.

car view and fine-grained for classifying car model is shown in [14,9].

To our best knowledge, few researches doing on the verification task of car model except [31]. Most studies of verification task are about face verification [8,11,29]. Furthermore, there is no previous attempt on the car model verification task by using deep learning algorithm. Deep learning algorithms are widely used for face verification [21, 32, 33].

Similar to Gcars, the CompCars dataset [31] also is a fine-grained dataset for the car verification, which contains images from different viewpoints and different car parts. Apart from the large-scale, our Gcars dataset have many significant benefits over the CompCars dataset. Firstly, our dataset includes large Chinese local car models which are very popular in China while absent in other countries. Secondly, we use the make-model-type car hierarchy which is better than that of CompCars dataset, make-mode-year. As the car hierarchy of CompCars dataset, the third lever structure, year, is insignificance classification for some car modes as the same model cars made in different years are slightly difference or may be nothing change. The car hierarchy of Gcars dataset is more practicality, and useful for computer vision and customers. Thirdly, our dataset offers better performance over than that of the CompCars dataset by using the benchmark framework Caffe.

III. GCARS DATASET

The images of Gcars dataset are collected from the public websites and search engines. It includes 643 car makes with 2,237 car models which contains most of the home and aboard car models. These images captures different views of car, including left, right, front, beside, front-side, rear, rear-side and overview. Usually, for a car model there are serval images to represent from different viewpoints. Some examples of car image are illustrated in Fig. 1, where the car is observed from front, rear, side and front-side with color. Overall, the unique

feature in comparing to existing car image dataset is that the Gcars dataset offers special car hierarchy and car range.
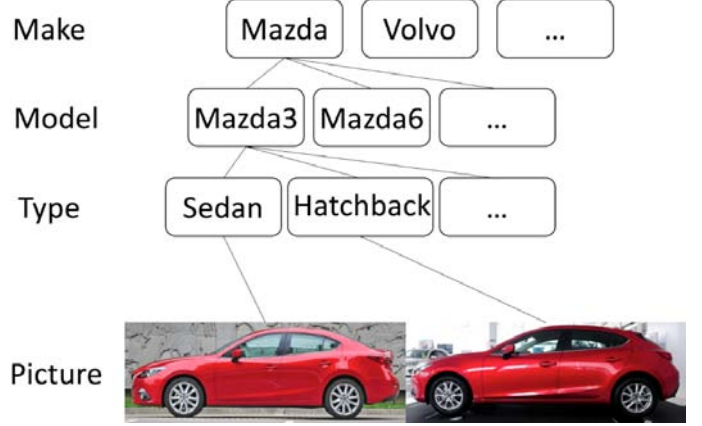

Fig.2. The car hierarchy of Gcars dataset.

A. Car Hierarchy

As we known that the car models can be organized into a large tree structure, here, three layers is consisted, namely car make, car model and type of manufacture. Fig. 2 depicts the structure of our dataset. The third level is type, which indicates that each car mode has serval types to meet the different requirement of customers. For example, the "AEOLUS" has many types. The car hierarchy of Gcars dataset has more advantages than that of the CompCars dataset. It has stronger practicability, easier classification and understanding for customer.

B. Car Range

We collect 2,237 car models which contains most of the car models in China. As China has the most rapidly increasing car market, it has the most car models all of the world including not only the famous car models such as Audi and Benz, but also some local car models, e.g. BYD, Geely and Alto. These local car models are not famous in the world, but popular and widely used by the Chinese customers. As the hybrid global and local car range, our Gcars dataset has more particularity than other previous datasets. It is obviously, our datasets can be very easily extended to add more other local car models for meeting some special requirement.

IV. CAR VERIFICATION

In this section, we investigate the car verification by using Gcars dataset. We select 90,990 images from the Gcars dataset and divide into two parts, training and verification of 78,741 and 12,249 images, respectively. This can be implemented by tune the parameter of *shuffle*. The DCNN is very popular in computer vision problem, especially in deep learning area. The DCNN is used for object classification, detection, and face verification. In this paper, the Caffe framework, where the DCNN algorithm is realized, is employed for car verification. Caffe is a most popular deep learning framework made with expression, speed, and modularity in mind, which is developed by the Berkeley Vision and Learning Center (BVLC) and by community contributors [34]. We adopt Caffe framework as a

baseline framework to realize the Deep Convolutional Neural Networks (DCNN) [35] in car model verification. Actually, it is used as the baseline framework for comparison the performances between the CompCars dataset and the Gcars dataset.
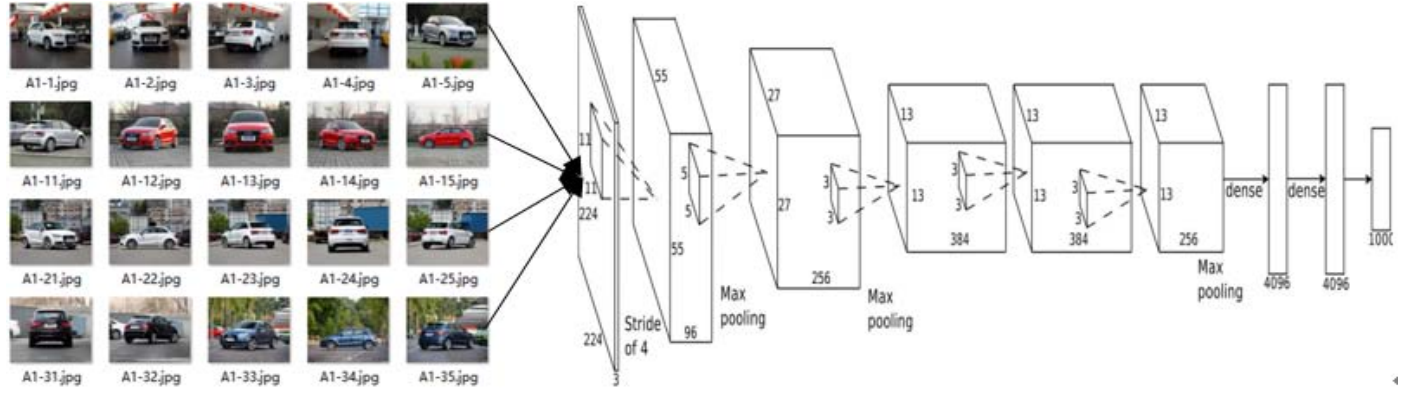

Fig. 3. The structure of Caffe framework

TABLE I. THE PARAMETERS OF CAFFE NET

Parameter	Value
batch size	128
dropout	0.5
learning rate	0.0001
maximum iterate	450,000
momentum	0.9

TABLE II. THE VERIFICATION ACCURACY OF TOP-1 OF THE TWO DATASETS

Dataset	Verification Accuracy
CompCars	54.12%
Gcars	80.156%

As shown in fig. 3, it has 10 neuron layers including input and output layers, and 8 hidden layers. The number of neurons in each hidden layer is given by 253440, 186624, 64896, 64896, 43264, 4096, 4096 and 1000. All images will be divided into two parts for training and verification, where the stochastic gradient descent and the backpropagation algorithm are employed. Car image as the raw image input the first layer where it will be preprocessed. Max-pooling will be used in the first, second and fifth convolutional layers.

Based on the DCNN algorithm, the one output unit per neuron is defined as

$$f(x_i) = \frac{e^{x_i}}{\sum_{j=i}^{n} e^{x_j}}$$

Where, x_i is the total input to a neuron i in the *Kth* hidden layer;

x_j is the input to a neuron j in the same layer;

And n is the number of neurons in the *(K-1)th* hidden layer.

The training images, total 78,740, selected from the Gcars dataset is the input of the first neuron layer of the Caffe net.

And the same activates is done for the CompCars dataset based on the same platform, where also contains 78,740 images. The parameters are shown in Table I.

The performance of the two datasets are presented in Table II. We observe that our dataset outperforms CompCars dataset with the Caffe net and DCNN algorithm.

V. CONCLUSIONS

As fewer researches focus on the larger-scale car model dataset compared to other objects in computer vision for verification, such as classification and face verification. In this paper, we present our on-going effort in collecting a large-scale and global dataset, Gcars, where 234,187 images of 2,237 car models from public website are collected including home and aboard car models. The dataset has a three-layer car hierarchy, make, model and type. We also demonstrate the most important application, car verification, exploiting the dataset. The performance in terms of verification accuracy is 80.156%, which is over than that of the benchmark CompCars dataset, 54.12%, by using the deep learning framework Caffe. The verification accuracy has improved more than 26%.

There are many other potential tasks than can be exploit Gcars dataset. Car alignment is a very interesting research topic, which can help for tracking, matching or matching. The 3 level hierarchy offers a strong and meaningful relationship graph for car models, and it will be more practicality combing with the other hierarchy, make-model-year.

REFERENCES

[1] Z. Zhang, T. Tan, K. Huang, and Y. Wang. Threedimensional deformable-model-based localization and recognition of road vehicles. IEEE Transactions on Image Processing, 21(1):1–13, 2012.

[2] B. C. Matei, H. S. Sawhney, and S. Samarasekera. Vehicle tracking across nonoverlapping cameras using joint kinematic and appearance features. In CVPR, 2011.

[3] Z. Sun, G. Bebis, and R. Miller. On-road vehicle detection: A review. T-PAMI, 28(5):694–711, 2006.

[4] L. Yang, J. Liu, and X. Tang. Object detection and viewpoint estimation with auto-masking neural network. In ECCV, 2014.

[5] E. Hsiao, S. N. Sinha, K. Ramnath, S. Baker, L. Zitnick, and R. Szeliski. Car make and model recognition using 3d curve alignment. In IEEE Winter Conference on Applications of Computer Vision, 2014.

[6] Y.-L. Lin, V. I. Morariu, W. Hsu, and L. S. Davis. Jointly optimizing 3D model fitting and fine-grained classification. In ECCV. 2014.

[9] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, October 2007.

[10] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar. Attribute and simile classifiers for face verification. In ICCV, 2009.

[11] Z. Zhu, P. Luo, X. Wang, and X. Tang. Multi-view perceptron: a deep model for learning face identity and view representations. In NIPS, 2014.

[16] Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems, 2012.

[7] J. Krause, M. Stark, J. Deng, and L. Fei-Fei. 3D object representations for fine-grained categorization. In ICCV Workshops, 2013.

[8] Yang L, Luo P, Loy C C. A large-scale car dataset for fine-grained categorization and verification. In CVPR, 2015.

[12] Y. Sun, X. Wang, and X. Tang. Deep learning face representation from predicting 10,000 classes. In CVPR, 2014.

[13] Zhao W, Chellappa R, Phillips P J, et al. Face Recognition: A Literature Survey. Acm Computing Surveys, 2000, 35(4):399-458.

[14] John W, Yang A Y, Arvind G, et al. Robust face recognition via sparse representation. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 44(12):2368 - 2378.

[15] http://caffe.berkeleyvision.org/.