When Do Luxury Cars Hit the Road? Findings by A Big Data Approach

Yang Feng
Computer Science
University of Rochester
Rochester, NY, 14627
yfeng23@cs.rochester.edu

Jiebo Luo
Computer Science
University of Rochester
Rochester, NY, 14627
jluo@cs.rochester.edu

Abstract—In this paper, we focus on a study of the timing of different kinds of cars on the road. This information will enable us to infer the lifestyle of the car owners. The results can further be used to guide marketing towards car owners and setting auto insurance policies. Conventionally, this kind of study is carried out by sending out questionnaires, which is limited in scale and diversity. To solve this problem, we propose a fully automatic method to conduct this study at scale. Our study is based on publicly available surveillance camera data. Images from the public traffic cameras are downloaded every minute. After obtaining the images, we apply faster R-CNN (region-based convolutional neural network) to detect the cars in the downloaded images and a fine-tuned VGG16 model is used to recognize the car makes. Based on the recognition results, we present a data-driven analysis on the relationship between car makes and their appearing times, with implications on lifestyles.

Index Terms—Data analytics; Car detection; Car make recognition

I. INTRODUCTION

It is well known that the car ownership rate is very high in America. There were about 1.8 vehicles per U.S. household in 2013 [1]. According to the report, the average number of minutes an American spends behind the wheel is 87 per day [2] and 85% workers commute to work in a car alone every day. Therefore cars are heavily used in America. From the use of the cars, we can find useful information about the lifestyle of American people. For example, such information can be provided to urban planning or car marketing.

Traffic jams are common in big cities. One interesting phenomenon about traffic jams is that we rarely see luxury cars in traffic jams. Do luxury car owners know when and where the traffic jam happens and how to avoid traffic jams? To answer this question, Zhang [5] carried out a study in one Chinese community. They manually recorded the leaving time and returning time of the cars in that community everyday. They also recorded the make of each passing car. They found out that the owners of many inexpensive cars usually leave their homes for work early, i.e. at 7 o’clock or 8 o’clock, while the owners of many expensive cars leave their homes no earlier than 9 o’clock. The luxury car owners, however, leave their homes at noon or in the evening. In fact, the luxury car owners usually do not go out during heavy traffic hours in the morning or afternoon.

While Zhang’s study is interesting, it only reports the result in one community. To draw a more convincing conclusion, their study needs to be carried out in more places. Their method is based on manually counting, which is labor expensive to scale up to many places. Motivated by Zhang’s study, we decide to design a fully automatic method to address the same question at scale.

We use the publicly available surveillance traffic cameras and computer vision methods to facilitate the study. To monitor the traffic condition on the road, the government has installed many surveillance traffic cameras. Some of them are configured as publicly accessible on the Internet. We search for them on Google and find several of these cameras. We
download the images captured by the cameras every minute. After obtaining the images, we need to detect the cars and recognize their makes. We use Faster R-CNN [3] for detecting cars. To recognize the car make, we transfer the knowledge in the CompCars dataset [6] to traffic camera images using the method proposed in [7]. Based on the recognition result, we find that different cars do appear at different times of a day according to their makes. Fig. 1 shows the entire framework of our method. Although some car makes produce both luxury cars and frugal cars, we regard these makes as frugal ones because the luxury cars sold by these makes are much less than the frugal ones.

II. RELATED WORK

A. Car User Study

The most related work we find is Choo and Mokhtarian’s research [8]. Based on a 1998 mail-out/mail-back survey, Choo and Mokhtarian studied the relation between vehicle type choices and lifestyle. The choice of individual’s vehicle type is affected by factors including the distance and frequency of trips, how much one enjoy traveling, pro-environmental considerations, family, money, gender and age. Zhang et al. [9] proposed a data-driven system to analyze individual refueling behavior and citywide gasoline consumption. Their method is more efficient and covers more people than the questionnaire-based methods.

B. Car Detection and Recognition

Ramnath et al. [10] proposed to recognize car make and model by matching 3D car model curves with 2D image curves. During recognition, the car pose is estimated and used to initialize 3D curve matching. Their method is able to recognize the make of a car over a wide range of viewpoints. Fraz et al. [11] represented the vehicle images using a mid-level feature for vehicle make and model recognition. Their mid-level representation is based on SURF and Fisher Vector while an Euclidean distance based similarity is used for classification. These two methods use high resolution car images, while the cars in our images are much smaller. So these methods are not suitable for our problem. Recently, He et al. [12] have used a part-based detection model to detect cars and designed an ensemble classifier of neural networks to recognize car models. Their method works on a single traffic-camera image and has obtained promising results on their dataset.

III. METHODOLOGY

A. Car Detection

We use Faster R-CNN [3] to detect cars. This detection method is developed from two previous work, namely R-CNN [13] and Fast R-CNN [14]. The object detection process in R-CNN can be summarized as:

- Generate object proposals using selective search [15].
- Extract a 4096-dimensional feature vector for each region proposal using AlexNet [16].
- Score each feature vector using object category classifiers.

The AlexNet used in the second step is pre-trained on ILSVRC2012 and fine-tuned on a specific detection data such as VOC or ILSVRC2013. The object category classifiers are trained using SVM. The detection results of R-CNN is good on the VOC dataset and ILSVRC2013 detection test set. However, it is a three-step procedure, in which the features need to be saved and training the object classifiers takes significant time. To solve these drawbacks of R-CNN, Fast R-CNN was proposed. In Fast R-CNN, feeding an image and multiple proposals to a CNN model, the model will output the softmax probabilities and refined bounding box offsets. The detection time reduced significantly in Fast R-CNN, which makes the proposal generating become the bottle neck in object detection. Although some car makes produce both luxury cars and frugal cars, we regard these makes as frugal ones because the luxury cars sold by these makes are much less than the frugal ones.
to use soft label for a few available labeled target domain samples. The idea of soft label is coming from distilling [21], which is able to distill the knowledge in an ensemble of models into a single model. The soft label in [7] is defined as the average over the softmax of all activations of samples in one category. It is able to provide more information than one-hot label. We do not have any labeled target sample, so we just use the first strategy.

IV. DATA

A. The CompCars dataset

The CompCars dataset [6] was collected by The Chinese University of Hong Kong. This dataset contains 208,826 images of 1,716 car models. These images are divided into web-nature images and surveillance-nature images. All the car models come from 163 car makes. This dataset was originally used for fine-grained car classification, car attribute prediction and car verification.

B. Traffic camera images

By searching “live traffic camera” in Google, we are able to find lists of available cameras from several cities on the Web. To cover different regions, we choose one city from each of east coast, west coast and the south. The cities chosen are New York, Seattle and Lafayette (Louisiana). They also happen to represent large, medium and small cities. For each city, we choose several cameras which are relatively clear. Images captured by these cameras are downloaded every minute. The image download URLs are listed in the appendix.

V. RESULTS

Nvidia GTX 980M is used for training and evaluating the CNN models.

The surveillance-nature scenario only contains the images captured in the front of a car, while the web-nature scenario and our download data contain images captured from different view angles. So we use the web-nature in CompCars dataset to train the car recognition model. There are 136,726 web-nature images in total. We randomly split them into a training part containing 122,995 images and a validating part containing 13,731 images. We reported the recognition accuracy on the 44,481 surveillance-nature scenario images. We first try to fine-tune the GoogLeNet model [22] and 16 layer VGG model [4]. Table I shows the recognition accuracy on the validating set and surveillance-nature scenario. VGG model achieves

|                  | GoogLeNet | VGG   |
|------------------|-----------|-------|
| Validating set, Top-1 | 84.4%     | 93.3% |
| Validating set, Top-5  | 95.9%     | 98.3% |
| Surveillance-nature, Top-1 | 49.2%     | 56.3% |
| Surveillance-nature, Top-5 | 72.4%     | 77.3% |

TABLE I
RECOGNITION ACCURACY ON THE VALIDATING SET AND SURVEILLANCE SCENARIO.

Fig. 2. The makes of all detected cars. The car makes with very small percentages are merged into “other”.

Detected cars in each hour on weekdays

Detected cars in each hour at weekend

Fig. 3. The percentages of detected cars in each hour: (top) weekdays and (bottom) weekends.
better results. So we train another VGG model using the domain confusion method proposed in [7], which gets a top-1 accuracy of 62.8% on the traffic-surveillance scenario. When predicting the traffic camera images, we train one additional model. The web-nature images are used as source domain and the traffic camera images are used as target domain. Because many car makes in the CompCars dataset are very rare to see in US, we manually prune some car makes when predicting the traffic camera images. After pruning, there are 30 major car makes left. They are Acura, Audi, Benz, BMW, Buick, Cadillac, Chevrolet, Chrysler, Dodge, Fiat, Ford, GMC, Honda, Hyundai, Infiniti, Jeep, Kia, Land Rover, Lexus, Lincoln, Mazda, Mini, Mitsubishi, Nissan, Porsche, Scion, Subaru, Toyota, Volkswagen and Volvo.

We remove the images captured from 9 o’clock p.m. to 6 o’clock a.m. because many of them are too dark for recognition. After cleaning, we have 128,768 images captured between Jul. 11 and Jul. 20 in 2016. The detector is applied to these images. After obtaining the detection result, we filter the cars less than 1000 pixels and the cars that stay in the same position. Fig. 2 shows the percentage of the car makes of all the detected cars. Fig. 3 shows the percent of detected cars in each hour. We observe that 40% of the cars appear between 6 o’clock and 13 o’clock on weekdays. On weekends, only 36% cars appear during that time. There are fewer percent of cars going out in the morning on weekends.

The next three figures show the sum of detected cars of some car makes with respect to hours. Fig. 4 lists the car makes which are detected most of the time. From Fig. 4, we observe that the owners of Lexus and Lincoln usually go to work earlier on weekdays and they go out much less on weekends compared with the owners of the other two car makes. Perhaps they work hard on weekdays and prefer to stay at home on weekends. The data for two Korean car makes are shown in Fig. 5. We can find that the owners of these two car makes go to work very early. And the owners of KIA also return to home quite late in the evening. This may be because some working-class people are likely to buy Korean cars because of their economical prices. Fig. 6 shows four luxury car makes. The owners of BMW also go to work quite early. But owners of the other three car makes go to work later than many of the owners of Japanese or Korean car makes. The owners of these luxury cars may have more flexibility in their work schedule.

VI. CONCLUSIONS

We have studied the timing when different classes of cars hit the road using a scalable data-driven approach. We have found that different brands of cars appear in the traffic at different times on weekdays, which suggests that the owners of different cars follow different lifestyles and work schedules.

Currently, we directly apply an object detector trained on PASCAL VOC dataset to the traffic camera images without adaptation. We plan to use the domain adaptation technique to adapt the object detector to the new domain to achieve a better detection performance so that the subsequent data analysis is less affected by detection noise and thus more
We acknowledge the support of New York State through the Goergen Institute for Data Science.

We also plan to significantly increase the size of the traffic image dataset in order to derive region-specific and population-specific patterns.

ACKNOWLEDGMENT

We thank the support of New York State through the Goergen Institute for Data Science.

REFERENCES

[1] “Car ownership in u.s. cities map,” http://www.governing.com/gov-data/car-ownership-numbers-of-vehicles-by-city-map.html, Accessed December 10, 2015.

[2] A. Quillen, “A snapshot of car-usage in america,” http://newurbanhabitat.com/2009/09/14/a-snapshot-of-car-usage-in-america/, September 14, 2009.

[3] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” arXiv preprint arXiv:1506.01497, 2015.

[4] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint arXiv:1409.1556, 2014.

[5] X. Zhang, “Finally understand why we cannot see luxury cars in traffic jams,” http://mp.weixin.qq.com/s?__biz=MzA4NjtA0TkzNw==&mid=26448367548&idx=2&sn=cdcadb38674ca7a3e4a3ca33c2869, July 15, 2015.

[6] L. Yang, P. Luo, C. C. Loy, and X. Tang, “A large-scale car dataset for fine-grained categorization and verification,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3973–3981.

[7] E. Tzeng, J. Hoffman, T. Darrell, and K. Saenko, “Simultaneous deep transfer across domains and tasks,” in Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 4068–4076.

[8] S. Choo and P. L. Mokhtarian, “What type of vehicle do people drive? the role of attitude and lifestyle in influencing vehicle type choice;” Transportation Research Part A: Policy and Practice, vol. 38, no. 3, pp. 201–222, 2004.

[9] F. Zhang, N. J. Yuan, D. Wilkie, Y. Zheng, and X. Xie, “Sensing the pulse of urban refueling behavior: A perspective from taxi mobility,” ACM Transactions on Intelligent Systems and Technology (TIST), vol. 6, no. 3, p. 37, 2015.

[10] K. Ramnath, S. N. Sinha, R. Szeliski, and E. Hsiao, “Car make and model recognition using 3d curve alignment,” in Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on. IEEE, 2014, pp. 285–292.

[11] M. Fraz, E. Edrissinghe, M. S. Sarfraz et al., “Mid-level-representation based lexicon for vehicle make and model recognition,” in Pattern Recognition (ICPR), 2014 22nd International Conference on. IEEE, 2014, pp. 393–398.

[12] H. He, Z. Shao, and J. Tan, “Recognition of car makes and models from a single traffic-camera image,” Intelligent Transportation Systems, IEEE Transactions on, vol. 16, no. 6, pp. 3182–3192, 2015.

[13] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on. IEEE, 2014, pp. 580–587.

[14] R. Girshick, “Fast r-cnn,” arXiv preprint arXiv:1504.08083, 2015.

[15] J. R. Uijlings, K. E. van de Sande, T. Gevers, and A. W. Smeulders, “Selective search for object recognition,” International journal of computer vision, vol. 104, no. 2, pp. 154–171, 2013.

[16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Advances in neural information processing systems, 2012, pp. 1097–1105.

[17] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

[18] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, “Large-scale video classification with convolutional neural networks,” in Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on. IEEE, 2014, pp. 1725–1732.

[19] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, “Deepface: Closing the gap to human-level performance in face verification,” in Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on. IEEE, 2014, pp. 1701–1708.

[20] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, “Learning and transferring mid-level image representations using convolutional neural networks,” in Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on. IEEE, 2014, pp. 1717–1724.

[21] G. Hinton, O. Vinyals, and J. Dean, “Distilling the knowledge in a neural network,” arXiv preprint arXiv:1503.02531, 2015.

[22] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” arXiv preprint arXiv:1409.4842, 2014.

APPENDIX

The image download urls are listed below.
http://www.bellevuewa.gov/TrafficCamImages/CCTV020.jpg http://www.bellevuewa.gov/TrafficCamImages/CCTV285.jpg http://www.bellevuewa.gov/TrafficCamImages/CCTV085.jpg http://www.bellevuewa.gov/TrafficCamImages/CCTV009.jpg http://www.bellevuewa.gov/TrafficCamImages/CCTV071.jpg http://www.bellevuewa.gov/TrafficCamImages/CCTV096.jpg http://www.bellevuewa.gov/TrafficCamImages/CCTV069.jpg http://207.251.86.238/cctv290.jpg http://207.251.86.238/cctv666.jpg http://207.251.86.238/cctv861.jpg http://207.251.86.238/cctv351.jpg http://207.251.86.238/cctv674.jpg http://207.251.86.238/cctv467.jpg http://207.251.86.238/cctv678.jpg http://www.lafayettela.gov/tcams/getTCamera.aspx?ptzid=113 http://www.lafayettela.gov/tcams/getTCamera.aspx?ptzid=210 http://www.lafayettela.gov/tcams/getTCamera.aspx?ptzid=151 http://www.lafayettela.gov/tcams/getTCamera.aspx?ptzid=186 http://www.lafayettela.gov/tcams/getTCamera.aspx?ptzid=3 http://www.lafayettela.gov/tcams/getTCamera.aspx?ptzid=4 http://www.lafayettela.gov/tcams/getTCamera.aspx?ptzid=5 http://www.lafayettela.gov/tcams/getTCamera.aspx?ptzid=6 http://www.lafayettela.gov/tcams/getTCamera.aspx?ptzid=7 http://www.lafayettela.gov/tcams/getTCamera.aspx?ptzid=8 http://www.lafayettela.gov/tcams/getTCamera.aspx?ptzid=9 http://www.lafayettela.gov/tcams/getTCamera.aspx?ptzid=10 http://www.lafayettela.gov/tcams/getTCamera.aspx?ptzid=11 http://www.lafayettela.gov/tcams/getTCamera.aspx?ptzid=12