CS766 Mid-term Report
A Robust License Plate Recognition System based on Domain Adaptation

Zhikang Hao
UW-Madison
hao45@wisc.edu

Wendi Li
UW-Madison
wendi.li@wisc.edu

Zhikang Meng
UW-Madison
zmeng34@wisc.edu

1 REVISIONS AND UPDATED TIMELINE

Previously we were focusing on SISR tasks on domain shifting. However, after a more comprehensive study of the models and datasets described in the papers related to SISR, we decide not to continue this topic. The reason is that we found evaluating the result of SISR tasks is very difficult, given the state-of-art models (e.g. ESRGAN) have already provided somewhat promising results. In other words, humans can hardly distinguish or quantitatively describe which model is the best generally. In addition, evaluating SISR results with machine learning is an emerging topic, where no conclusion on how to evaluate has been made. The challenges and solutions in these are focused on proposing new ideas which achieve better results after shifting to a wider range of real-world scenarios. The improvements could be fixing unexpected blur or other degradation in certain areas varying on the individual picture.

Below is the proposed timeline updated for the topic change

April 16	finalize implementation
April 24	evaluate overall performance
April 28	finalize our report and presentation slides

2 CREATE OUR NEW DATASETS

The Chinese City Parking Dataset (CCPD) [13] that contains 250 thousand license plate pictures will be used. The CCPD images are sampled from different scenarios, the description of this CCPD dataset is described in Table 1.

Table 1: THE DESCRIPTION OF DATA SET

Type	Description
CCPD-Base	The only common feature of these photos is the inclusion of a license plate.
CCPD-DB	Illuminations on the LP area are dark, uneven or extremely bright.
CCPD-Weather	Images taken on a rainy, snow, or fog day.
CCPD-Fn	The license plate is relatively far or near to the camera
CCPD-Challenge	The most challenging images

The overall license plate image processing workflow is as follows:

(1) Converting the pictures to gray-scale images.
(2) Augmenting the image with the following three techniques:
(3) Applying Median filtering on gray-scale images to remove single noisy pixels on the image.
(4) Using Histogram Equalization to dynamically enhance contrast and spread of pixel value within each image.
(5) Binarizing the images dynamically with the following steps: Let $h(i)$ be the equalized histogram of gray-scale image, the mean of gray-scale value is

$$u_r = \sum i * h(i)$$

The zeroth cumulant w(k) and first cumulant v(k) are:

$$w(k) = \sum_{i=0}^{k} h(k), \quad \forall k = 0, 1...255$$

$$v(k) = \sum_{i=0}^{k} ih(k), \quad \forall k = 0, 1...255$$

Binarization adopts the threshold k which satisfies:

$$k = \underset{k}{\operatorname{argmax}} \frac{[u_r w(k) - v(k)]^2}{w(k)(1 - w(k))}$$

Figure 1 shows the transformation of raw image into binarized image.

The next step is separating the image into segments containing single characters. The license plates contain all numbers and all capital English characters except `O` and `I`. An accurate segmentation directly helps boost the performance of classification by providing legitimate and clear inputs. We adopt Vertical Projection to separate each character on the license plate.

For each column j, sum up its pixel value

$$sum(j) = \sum_{i} d(i, j), \; d(i, j) = 0, 255$$

Scanning the image by columns and starts segmenting individual character by a threshold Q_s such that:

$$sum(j) \geq Q_s$$

and ends if

$$sum(j) < Q_s$$

The processed result contains five sets of images captured during different climates/environments/scenarios, as illustrated in 1. Each set includes images with exactly one English character or digit separated by Vertical Projection. Each image is labeled by the English character or digits in it. Although the above process has given a relatively well segmentation of plate characters, there are still outliers that need manual separation and classification. We review all the datasets and amend labels manually and conduct new processed CCPD datasets for license plate character recognition.

The pictures in CCPD-Base are the collection of the most common ones. They are shot right in front of the car when the light

(a) Raw pictures

(b) License plate cropping

(c) Gray-scale

(d) Binary

Figure 1: License plate image processing workflow

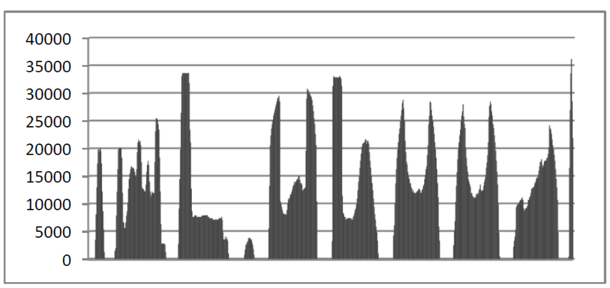


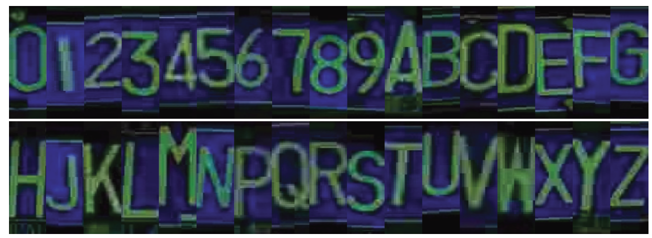
Figure 2: Accumulated Gray Scale Value by Column

is relatively abundant. The angle of the shooting is relatively positive, the focus is good, and there is almost no blur. While building the license plate image recognition system, CCPD-Base would be the ideal dataset for training models. However, not all the images captured by the system are like this. There would be images shown from other CCPD scenarios. Our goal is to build a robust license plate recognition system, which means robust to all of the scenarios.

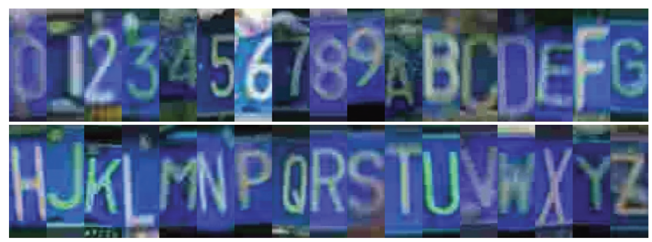
3 PROBLEM FORMULATION & FUTURE PLANS

Our next step will focus on doing a robust recognition process for all the scenarios in our processed CCPD datasets.

(a) License plate samples in *CCPD-Base*.



(b) License plate samples in *CCPD-DB*.



(c) License plate samples in *CCPD-Weather*.

(d) License plate samples in *CCPD-Fn*.

(e) License plate samples in *CCPD-Challenge*.

Figure 3: Five types of license plates compared.

Consider the case when the amount of training data is limited and there are many unseen situations, where the generalization capability of a trained license plate recognition system is usually limited. Sometimes, the license plate image distortion can be serious due to either weather conditions or technical reasons for photographing. In this case, the accuracy of the originally trained license plate recognition system may be greatly reduced.

We think solving such a problem would involve the concept, of transfer learning. Or more precisely, the domain adaption. It means that we regard the original dataset (CCPD-Base) as the domain of the training set, and without knowing the labeled data in other scenarios (CCPD-DB, CCPD-Weather, CCPD-Fn, CCPD-Challenge), we want our license plate recognition system to be robust to those new domains. The criteria of "robust" would be the evaluation performance in those scenarios. We have reviewed the state-of-the-art research papers on transfer learning and concluded some possible directions that we think are useful. In the next step, we will verify these methods to see whether they could help us achieve our goal. If not, we will analyze the reasons why they do not work (differences in assumptions, models, etc.). We will build a robust license plate recognition system as our final production.

3.1 Transfer Learning: An overview

In general, the environment where a classifier works can be different from the environment of training. For example, a plate recognition classifier is supposed to work under various conditions, including sunny days, rainy days, or even a snowy night. In the assumption of the transfer learning paradigm, we can assume the source and target domains have different distributions rather than the same distribution [6]. Most statistical models need to be rebuilt from scratch based on newly collected training data when the distribution changes. However, it may be expensive or impossible to recollect the labeled training data under the new distribution. For example, we need to deploy the license plate recognition system right now with a limited original dataset, but we need our system available to deal with the other scenarios.

Transfer learning allows the domains used in training and test to be different. There are very common examples even in daily real-world life. Children can learn to recognize shepherd dogs once they learn sled dogs. The study on transfer learning is motivated by the fact that humans can apply knowledge learned previously to solve new problems faster or with better solutions.

Transfer learning includes multiple categories, such as learning to learn [10], multi-task learning [7], life-long learning [2], knowledge distillation [4], meta-learning [11], and domain adaptation [5], etc. According to our problem situation, we will choose the domain adaptation methods to deal with our problem.

3.2 Domain Adaptation

Domain adaptation allows knowledge from a source domain to be transferred to a different but related target domain [5]. It aims to solve a learning problem in the target domain by utilizing training data in the source domain, allowing these domains have different distributions.

Domain adaptation is considered a special set of transfer learning which transfers shared knowledge across different but related domains. In order to reduce the difference between the distributions of the source and target domain, discovering a good feature representation is important [1, 3].

3.2.1 Problem Formulation in Domain Adaptation. Consider a domain as a feature space $\mathcal{X}$ and a marginal probability distribution of input $P(X)$, where $X = \{x_1, x_2, \cdots, x_n\} \in \mathcal{X}$ is the set of training samples.

Let D_S and D_T represent the source domain and target domain, respectively. Notice that we have labeled data of the source domain (training set), but we do not have the label of the target domain (testing set) D_T. Therefore, the source domain data would be $D_S = \{(x_{S_1}, y_{S_1}, \cdots, x_{S_{n1}}, y_{S_{n1}}\}$, and the target domain would be $D_T = \{x_{T_1}, \cdots, x_{T_{n2}}\}$.

Let P and Q denotes the marginal distributions of $X_S = \{x_{S_i}\}$ and $X_T = \{x_{T_i}\}$. The target of the domain adaptation is to predict the labels in the target domain y_{T_i} given the features in the target domain x_{T_i} and the source domain D_S. Note that according to the assumption, $P \neq Q$ while $P(Y_S|X_S) = Q(Y_T|X_T)$.

3.3 Some SOTA Methods Considered

Currently, we have reviewed the academic papers on domain adaptation, including traditional domain adaption methods and machine-learning-related domain adaption methods. Here is a conclusion of some possible state-of-the-art (SOTA) approaches that we would like to try:

(1) Sequential DA using distributionally robust experts [9] uses the Bernstein online aggregation algorithm on the proposed family of robust experts to generate predictions for the sequential stream of target test samples.
(2) Discriminative joint probability Maximum Mean Discrepancy (MMD) [14] uses joint adaptation with different weights to provide a simpler and more accurate method to compute the distribution discrepancy. It also increases the transferability and discriminability simultaneously.
(3) Federated Knowledge Alignment [8] tries to align features from different clients and those of the target task. It also devise a federated voting mechanism to provide labels for samples from the target domain via a consensus from querying local models and fine-tune the global model with labeled samples.
(4) Graph-relational domain adaptation [12] uses a domain graph to encode domain adjacency, and allows domains to align flexibly based on the graph structure.

We will use the SOTA methods mentioned above and compare whether they can handle our license plate recognition scenarios. The experiments are still carrying out and will be included in our final presentation.

3.4 Final Project Website

Since our project topic is creating a robust classification under shifting domains, we plan to build a website to illustrate our works with interactions. In addition to explanation, we are going to build online interfaces which accept license plate pictures uploaded by users and return recognized plate values. The training data is solely

CCPD-Base, but the user inputs may be licence plate pictures taken in any scenario.

REFERENCES

[1] Shai Ben-David, John Blitzer, Koby Crammer, and Fernando Pereira. 2006. Analysis of representations for domain adaptation. *Advances in neural information processing systems* 19 (2006).

[2] Zhiyuan Chen and Bing Liu. 2018. Lifelong machine learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning* 12, 3 (2018), 1–207.

[3] Joumana Ghosn and Yoshua Bengio. 2003. Bias learning, knowledge sharing. *IEEE Transactions on Neural Networks* 14, 4 (2003), 748–765.

[4] Yoon Kim and Alexander M Rush. 2016. Sequence-level knowledge distillation. *arXiv preprint arXiv:1606.07947* (2016).

[5] Sinno Jialin Pan, Ivor W Tsang, James T Kwok, and Qiang Yang. 2010. Domain adaptation via transfer component analysis. *IEEE transactions on neural networks* 22, 2 (2010), 199–210.

[6] Sinno Jialin Pan and Qiang Yang. 2009. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering* 22, 10 (2009), 1345–1359.

[7] Sebastian Ruder. 2017. An overview of multi-task learning in deep neural networks. *arXiv preprint arXiv:1706.05098* (2017).

[8] Yuwei Sun, Ng Chong, and Ochiai Hideya. 2022. Multi-Source Domain Adaptation Based on Federated Knowledge Alignment. *arXiv preprint arXiv:2203.11635* (2022).

[9] Bahar Taskesen, Man-Chung Yue, Jose Blanchet, Daniel Kuhn, and Viet Anh Nguyen. 2021. Sequential domain adaptation by synthesizing distributionally robust experts. In *International Conference on Machine Learning*. PMLR, 10162–10172.

[10] Sebastian Thrun and Lorien Pratt. 2012. *Learning to learn*. Springer Science & Business Media.

[11] Joaquin Vanschoren. 2018. Meta-learning: A survey. *arXiv preprint arXiv:1810.03548* (2018).

[12] Zihao Xu, Guang-He Lee, Yuyang Wang, Hao Wang, et al. 2022. Graph-Relational Domain Adaptation. *arXiv preprint arXiv:2202.03628* (2022).

[13] Zhenbo Xu, Wei Yang, Ajin Meng, Nanxue Lu, and Huan Huang. 2018. Towards End-to-End License Plate Detection and Recognition: A Large Dataset and Baseline. In *Proceedings of the European Conference on Computer Vision*. 255–271.

[14] Wen Zhang and Dongrui Wu. 2020. Discriminative joint probability maximum mean discrepancy (DJP-MMD) for domain adaptation. In *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 1–8.