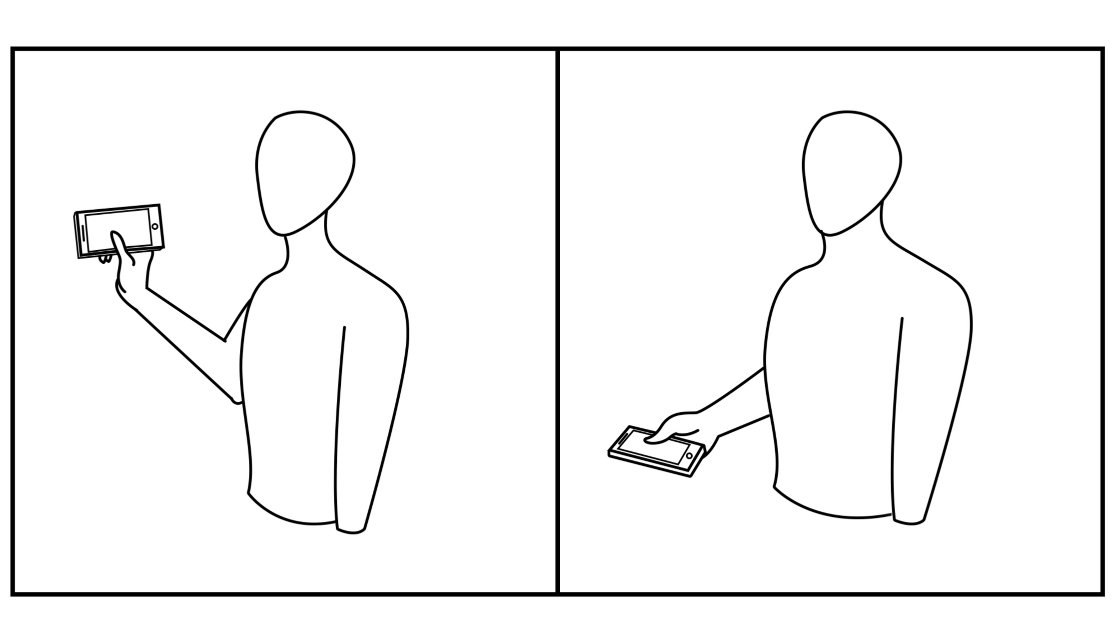
Cross-People Mobile-Phone Based Airwriting Character Recognition

Yunzhe Li, Hui Zheng, He Zhu, Haojun Ai, Xiaowei Dong
{liyunzhe, zhenghui_cs, zhuhe_cs, aihj, xwdong}@whu.edu.cn

Problem

Airwriting using mobile phones has many applications in human-computer interaction. However, the recognition of airwriting character needs a lot of training data from user, which brings great difficulties to the pratical application. The model learnt from a specific person often cannot yield satisfied results when used on another person.The data gap between people is mainly caused by the following factors: personal writing styles, mobile phone sensors, and ways to hold mobile phones.



Algorithm

Train with Modified AdaBN

Input:
Unlabeled Target Dataset D_T, Pre-trained model $\hat{f}$, Labels Y

Initialization:
Data Mapping and AE Conversion;

Output:
A new model $\hat{f}^a$ adaptable to D_T.

while training $\hat{f}^a$:
  for neuron j in $\hat{f}$ **do**
    Collect the neuron responses $\{x_j(m)\}$ on all tensors of D_t, where $\{x_j(m)\}$ is the response for tensor m.
    Compute the mean and variance of the target domain: μ_j^t and σ_j^t using Eq. 1, Eq.2, Eq.3 and Eq.4.
  end for
end while
while testing with $\hat{f}^a$:
  for neuron j in $\hat{f}^a$ **do**
    Compute BN output $y_j(m) := \gamma_j \frac{(x_j(m-\mu_j^t))}{\sigma_j^t} + \beta_j$
  end for
end while
while Stop condition not reached:
  Ask for some labels.
  Fine-tune the model $\hat{f}^a$ and testing it.
end while
return $\hat{f}^a$

Parameter update

$$d = \mu - \mu_j \tag{1}$$

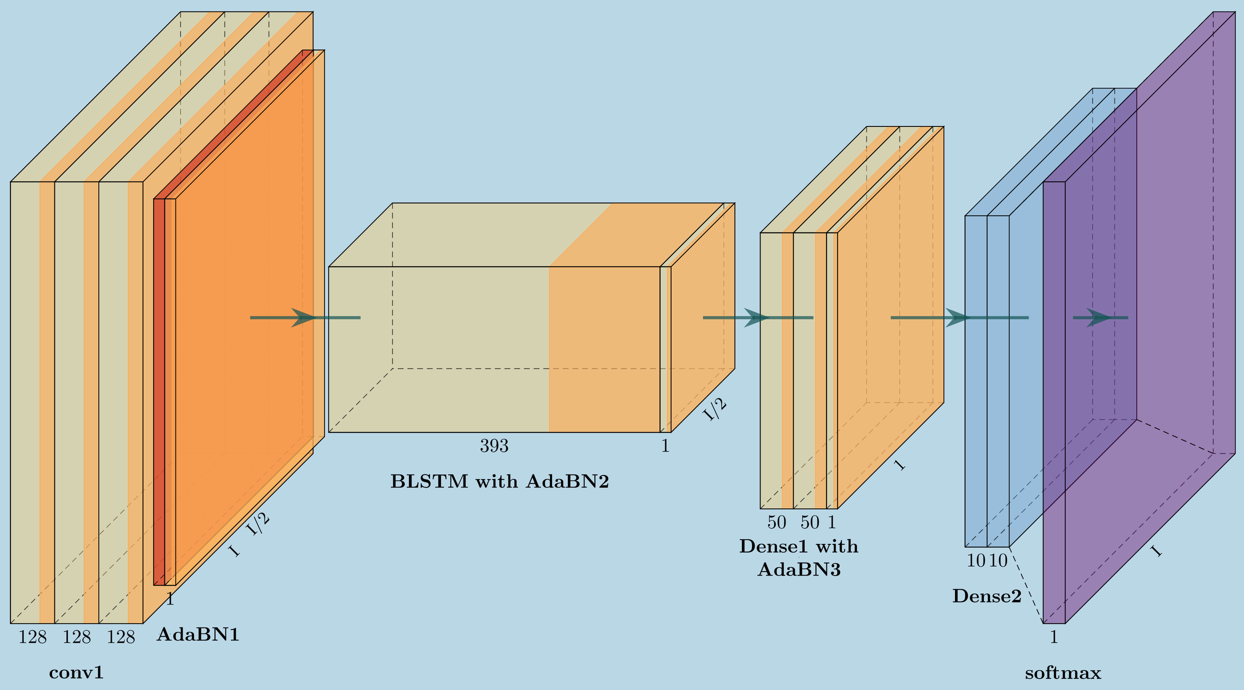
$$\mu_j \leftarrow \mu_j + \frac{dk}{n_j} \tag{2}$$

$$\sigma_j^2 \leftarrow \frac{\sigma_j^2 n_j}{n_j + k} + \frac{\sigma^2 k}{n_j + k} + \frac{d^2 n_j k}{(n_j + k)^2} \tag{3}$$

$$n_j \leftarrow n_j + k \tag{4}$$

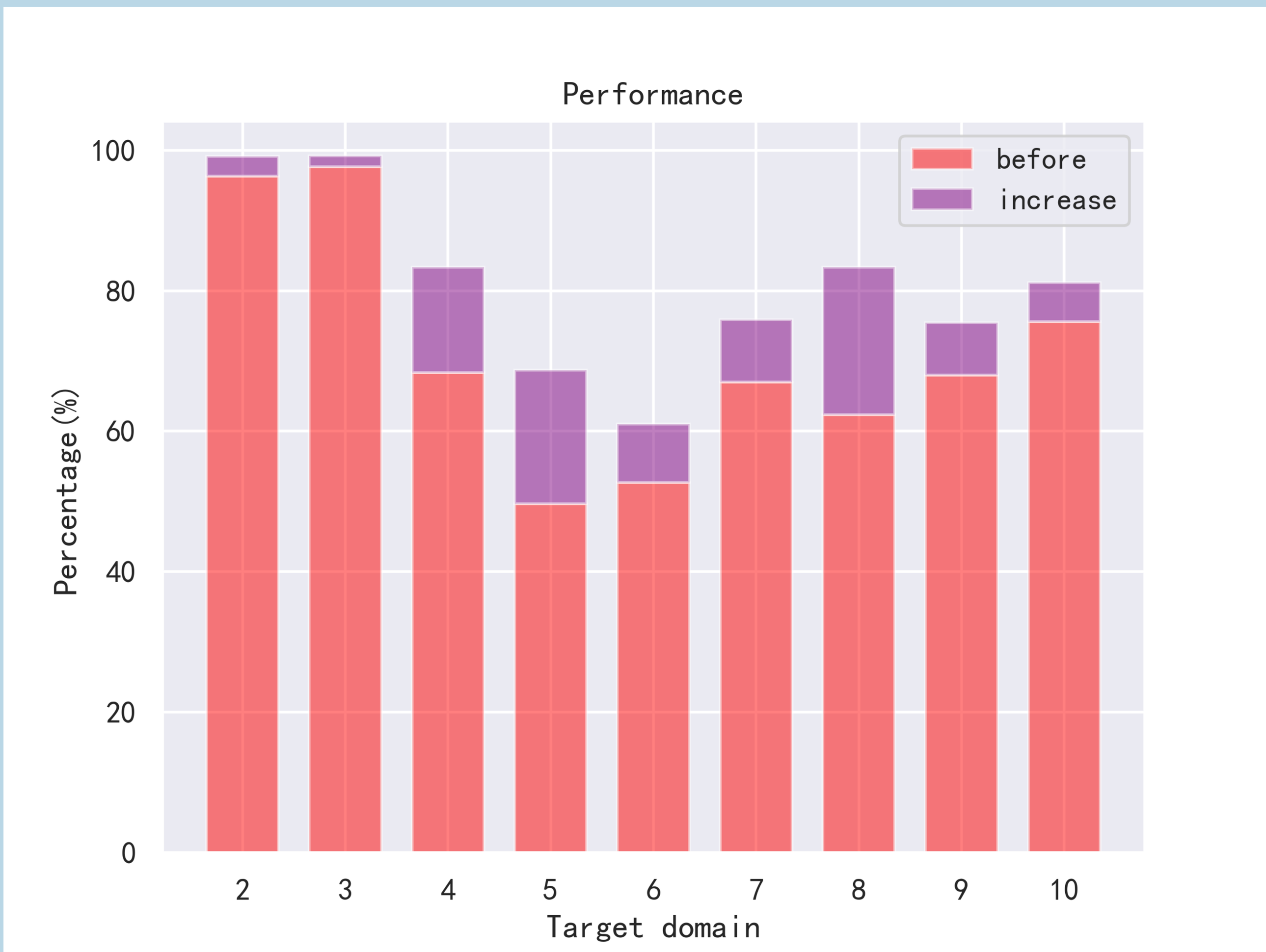
Model

To address the cross-people problem, we propose a deep neural network(DNN) that combines convolutional neural network(CNN) and bilateral long short-term memory(BLSTM). In each layer of the network, we also add an AdaBN layer which is able to increase the generalization ability of the DNN. Different from the original AdaBN method, we explore the feasibility for semi-supervised learning.



Evaluation

We have performed experiments to verify the performance of our approach across people, devices, and poses. For experiments on the dataset, one type(namely T_i) of samples was randomly selected. These samples are also divided into a training set and a testing set at a 7:3 ratio. We use samples in the training set to train the network combining AdaBN layers unsupervised to update model trained before. Then, 10 % of the samples in the training set is used to fine-tune the model. Lastly, the testing set was used to test. This step was repeated 5 times. The baseline of our experiment is the accuracy of the naive CNN-BLSTM network: in this experiment, we used some data that is different from the training data set for testing for comparison. For example, the base model is trained using the data from Person 1, but the testing data may come from Person 2. The data are shown below. All results are the average of five experiments.



Target domain	Before Transfer	After Transfer	Percentage increase
2	96.35%	99.14%	2.79%
3	97.64%	99.23%	1.59%
4	68.33%	83.33%	15%
5	49.67%	68.67%	19%
6	52.67%	61.02%	8.35%
7	67%	75.92%	8.92%
8	62.33%	83.33%	21%
9	68%	75.45%	7.45%
10	75.59%	81.2%	5.61%

Contributions

The main contributions of this work are summarized as follows:

1. We provide a data set on mobile-phone based airwriting recognition with a number of over 50000, taking into account many factors such as people, devices, and poses.
2. We propose an adaptive model that can be used in mobile phones for airwriting. It can reach more than 99% accuracy when considering multiple writing methods.
3. To address the problem that the general model cannot adapt well to personalized data, we propose a transfer learning method that can significantly improve the performance of the model on personalized data by more than 10%.