Radical Counter Network for Robust Chinese Character Recognition

Yun-qing Li, Yi-xing Zhu, Jun Du, Chang-jie Wu, Jian-shu Zhang

Introduction

➢ Background

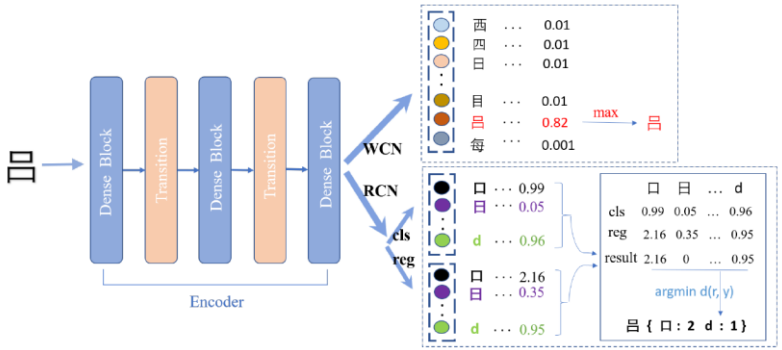
Chinese character recognition remains a challenging work due to its large character categories, high similarity between characters and different application scenarios.

➢ Motivation

Mainly researches have the following shortcomings:

- The character-based methods can recognize common characters well but unable to handle unseen situation.
- Some radical-based methods can recognize unseen characters but need RNN-based decoder for sequence decoding.

Model



RCN

➢ Architecture

- CNN-based encoder
- RCM: judges whether the radicals exist.
- RRM : estimates the number of existing radicals

➢ Training

We use BCE loss for classification and MSE loss for regression.

$$L_{\mathrm{RCN}} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=0}^{1} y_{ij}^{\mathrm{cls}} \log\left(o_{ij}^{\mathrm{cls}}\right) + \frac{\lambda}{n}\sum_{i=1}^{n} d(o_i^{\mathrm{m}}, y_i^{\mathrm{reg}})$$

➢ Inference

We choose the category with the closest distance in the candidate set as the prediction result.

$$o^{\mathrm{mcls}} = arg\max_{y^{\mathrm{cls}} \epsilon Y} d_1(o^{\mathrm{cls}}, y^{\mathrm{cls}})$$

$$o^{\mathrm{mreg}} = o^{\mathrm{mcls}} \cdot o^{\mathrm{reg}}$$

$$c = arg\min_{y^{\mathrm{reg}} \epsilon Y} d_2(o^{\mathrm{mreg}}, y^{\mathrm{reg}})$$

M-RCN

We combine these two decoders as a new multi-task network to get both global and local information from the shared encoder.

➢ Training

$$L = L_{\mathrm{RCN}} + \mu \sum_{i=1}^{K} y_i^{\mathrm{WCN}} \log(o_i^{\mathrm{WCN}})$$

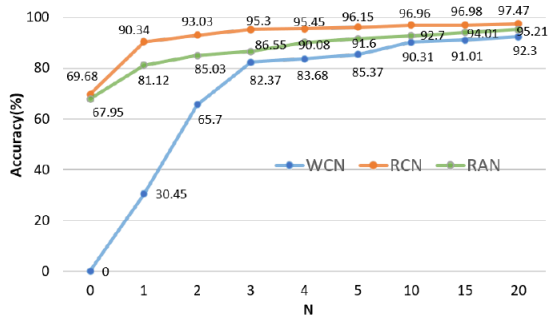
➢ Inference

We only utilize the output parts of RCN to produce the predicted character.

Experiment

➢ Experiment on Printed Character Dataset

In order to show the great superiority of RCN in zero-shot or few-shot conditions, we first design a few experiments on printed characters.



➢ Experiment on Scene Character Dataset

Method	Accuracy	GFLOPs
AlexNet [2]	73.0	-
Overfeat [2]	76.0	-
ResNet50 [2]	78.2	-
ResNet152 [2]	79.0	-
Google Inception [2]	80.5	-
DenseNet	**79.45**	1.0751
RCN	**84.12**	1.0745
M-RCN	**84.93**	1.0779
RAN [6]	**85.22**	1.0926

➢ Ablation Studies

1. The effect of different modules

Model	Accuracy	Acc ↓
w/o RRM	80.50%	**3.62%**
w/o RCM	78.61%	**4.73%**
RCN	84.12%	-

2. The effect of different distance metrics

d_1	Ed	Cd	Ed	Cd
d_2	Cd	Cd	Ed	Ed
accuracy	83.02%	83.27%	83.87%	**84.12%**

Future work

1. investigate the ability of RCN in other language recognition tasks. 2. expand the model to text line recognition.