Handwritten Character Recognition by Alternately Trained Relaxation Convolutional Neural Network

Chunpeng Wu, Wei Fan, Yuan He, Jun Sun, Satoshi Naoi

Fujitsu R&D Center, Co., Ltd.

Sep 1st, 2014

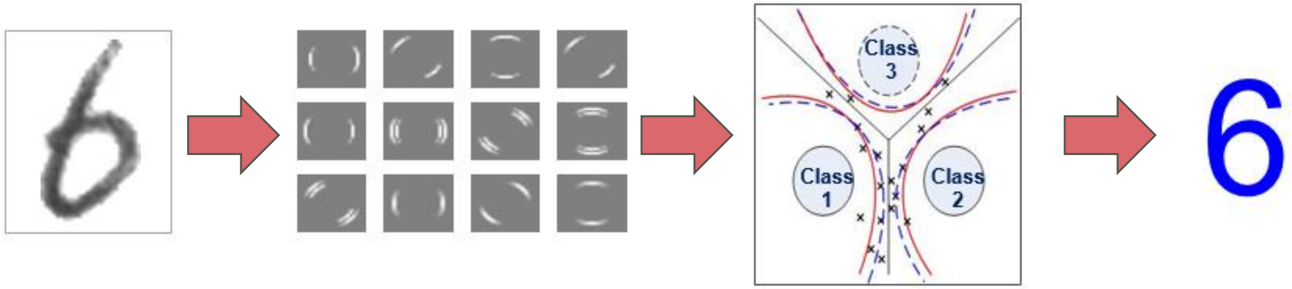
Outline

- Introduction to Convolutional Neural Network (CNN)
- Proposed Method
  - R-CNN: Relaxation CNN
  - ATR-CNN: Alternately Trained R-CNN
- Experiments
  - Handwriting Digits - MNIST
  - Handwriting Chinese - ICDAR’13 Competition Dataset
- Conclusions

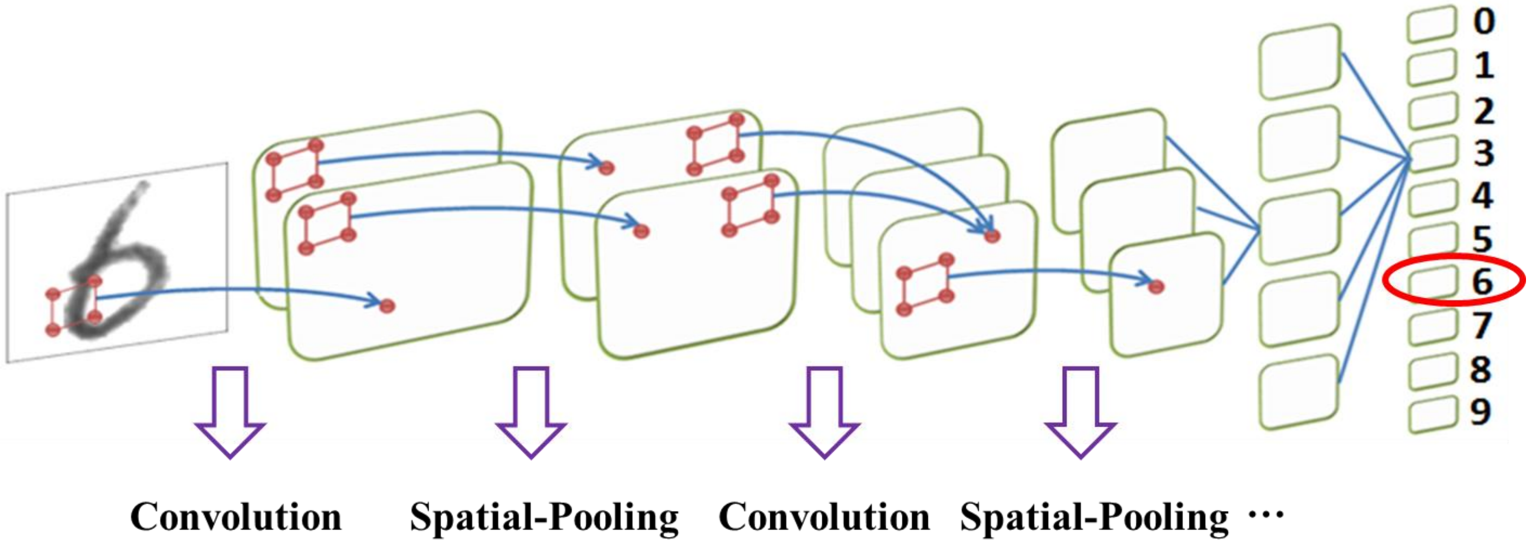
Introduction

FUJITSU

- Traditional Handwriting Recognition Methods
  - Handcrafted features + Classifiers



- Recent Deep Convolutional Neural Networks (CNN)
  - Learned features + Classifiers



Introduction

FUJITSU

- Success of CNN relies on
  - High performance computing (GPUs)
  - Flexible structure of neural networks
  - Availability of larger datasets
  - Effective learning algorithms

- Challenges of CNN Based Methods
  - Slow convergence
    - *CNN structure* vs *the scale of training dataset*
  - Over-fitting
    - Typical stochastic regularizing techniques
      - Dropout
      - Drop-connect
      - Make spatial-pooling a stochastic process

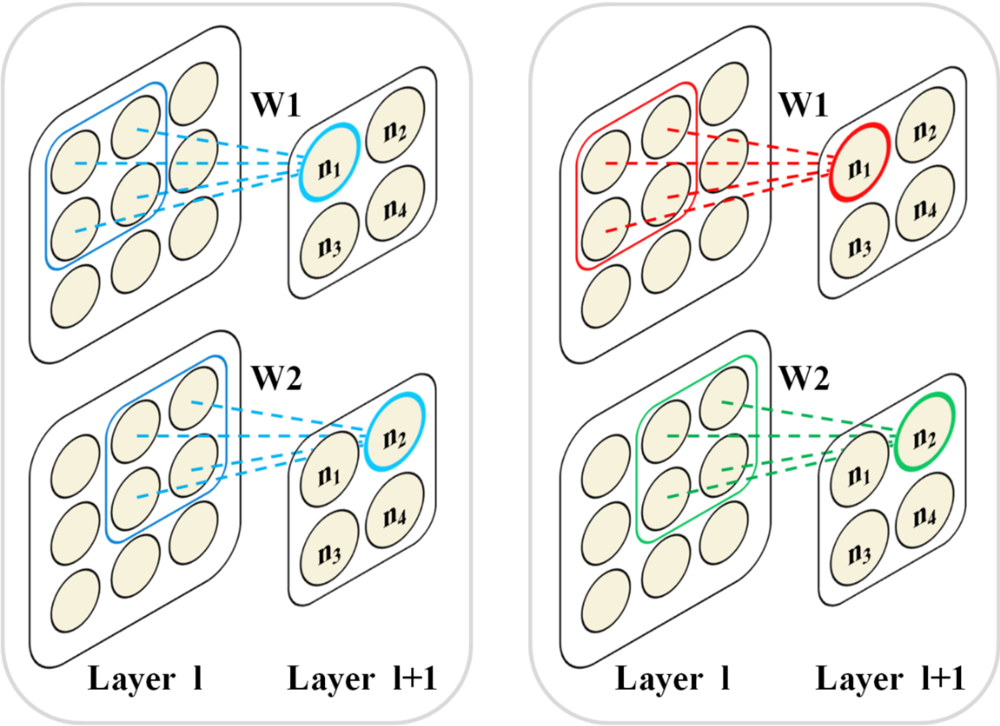
Proposed Method

- R-CNN: Relaxation CNN
  - Neurons within a feature map ***do not*** share the same kernel
  - Endow CNN with more expressive power

- ATR-CNN: Alternately Trained R-CNN
  - ***Randomly stop*** one layer from learning at one epoch
  - Regularize R-CNN

Proposed Method

- R-CNN
  - Enhance the learning ability of CNN



CNN: Convolution
$W_1 = W_2$

R-CNN: Relaxation Convolution
$W_1 \neq W_2$

CNN:

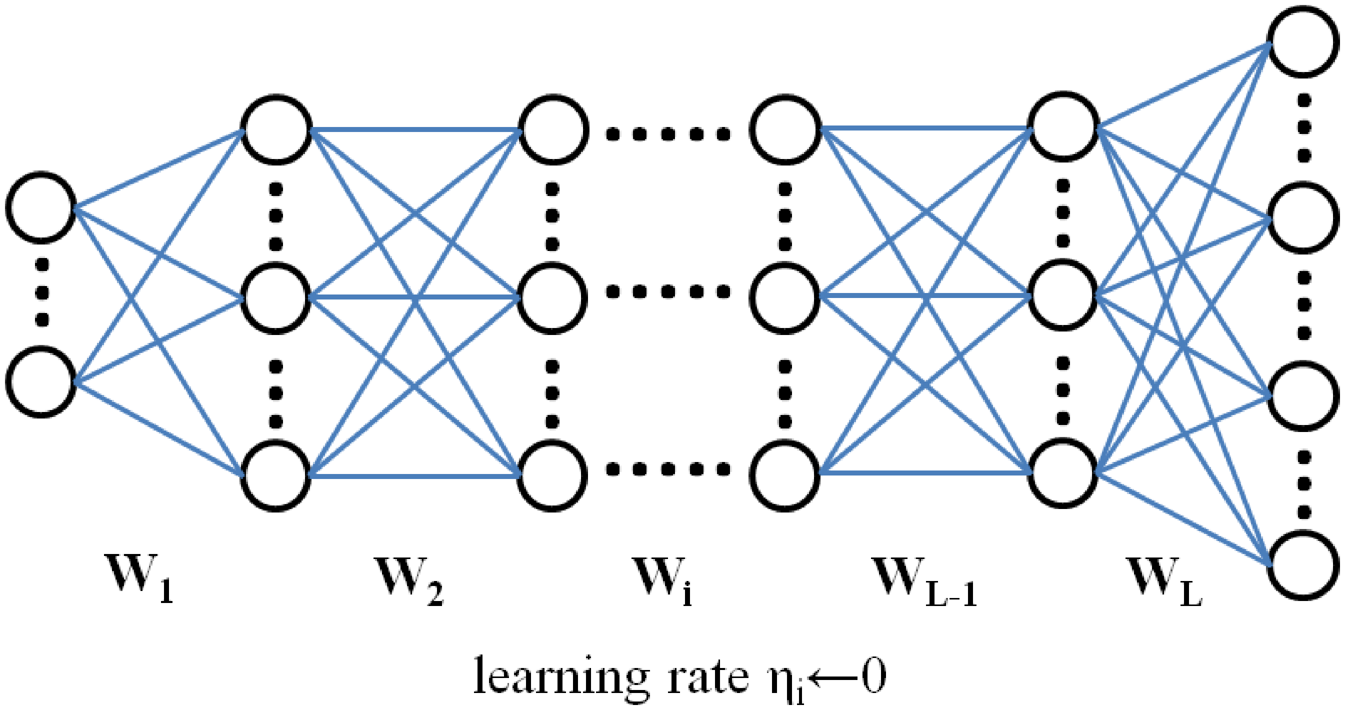
Neurons n_1 and n_2 share the ***same*** weight matrix W_1 (or W_2)

R-CNN:

Neurons n_1 and n_2 use ***different*** weight matrices W_1 and W_2

Proposed Method

- ATR-CNN
  - Randomly fix a learning rate to zero at one epoch
  - Regularization



Proposed Method

FUJITSU

- ATR-CNN

Each layer has a learning rate η_i

Randomly fix a η_i to zero at one epoch

Revert η_i to its original value after this epoch

Table I

ALGORITHM: ALTERNATELY TRAINING R-CNN

Initialization:

set $\boldsymbol{W} = \{W_1, W_2, ..., W_L\}$, $W_i \sim Uni(0,1)$

set $\boldsymbol{\eta} = \{\eta_1, \eta_2, ..., \eta_L\}$, $\eta_i \leftarrow C_i$

Output:

$\boldsymbol{W} = \{W_1, W_2, ..., W_L\}$

Iteration:

while not converging **do**

randomly choose η_k from $\boldsymbol{\eta}$

$\eta_k^{'} \leftarrow \eta_k$

$\eta_k \leftarrow 0$

for *sub-epoch* $\leftarrow$ 1 *to SE* **do**

for *all mini-batches of samples* **do**

forward propagate a mini-batch of samples

for $layer \leftarrow L$ *to* 1 **do**

compute ∇W_{layer} by back-propagation

$W_{layer} \leftarrow W_{layer} - \eta_{layer} * \nabla W_{layer}$

end for

end for

end for

$\eta_k \leftarrow \eta_k^{'}$

end while

Experiments – Handwriting Digits

- MNIST (Training: 60000 Testing: 10000)
- Our ATR-CNN
  - In-32Conv5-32MaxP2-64Conv3-64MaxP2-64RX3-64RX3-Out
- NVIDIA GTX 690, 64GB RAM

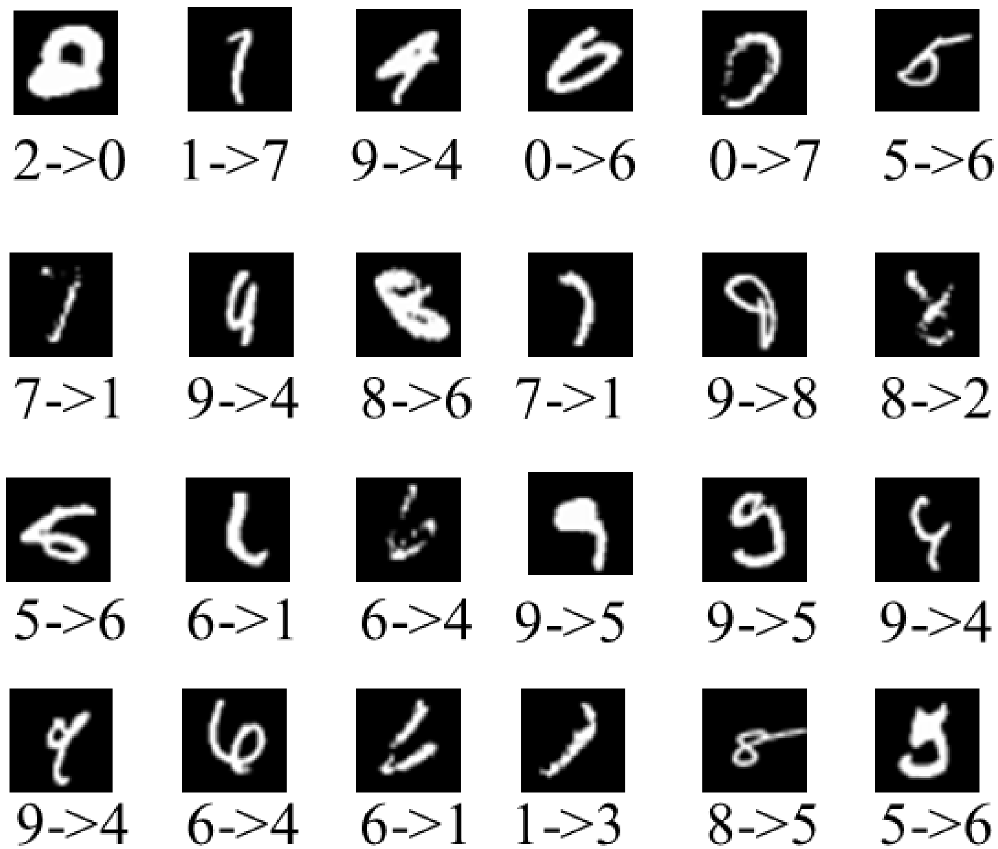
Method	Model	Error Rate (%)
Lecun et al. [33]	Boosted Letnet-4	0.70
Mizukami et al [2]	KNN	0.57
Lauer et al. [3]	TFE-SVM	0.54
Keysers et al. [1]	KNN	0.52
Simard et al. [31]	CNN	0.40
Wan et al. [11]	CNN+DropConnect	0.280 ± 0.032
Hinton et al. [18]	CNN+DropOut	0.280 ± 0.016
Our Method	R-CNN	0.274 ± 0.021
Our Method	**ATR-CNN**	$\mathbf{0.254 \pm 0.014}$

Experiments – Handwriting Digits

FUJITSU

MNIST

- Misclassified samples（ground-truth -> prediction）



Experiments – Handwriting Chinese

FUJITSU

- Testing Set
  - ICDAR’13 Competition Dataset (224,419 samples, 3755 classes)
- Our ATR-CNN
  - In-64Conv5-64MaxP2-128Conv3-128MaxP2-**128RX3**-128MaxP2-**256RX3**-256Full1-Out
- Narrow the gap between machine and human

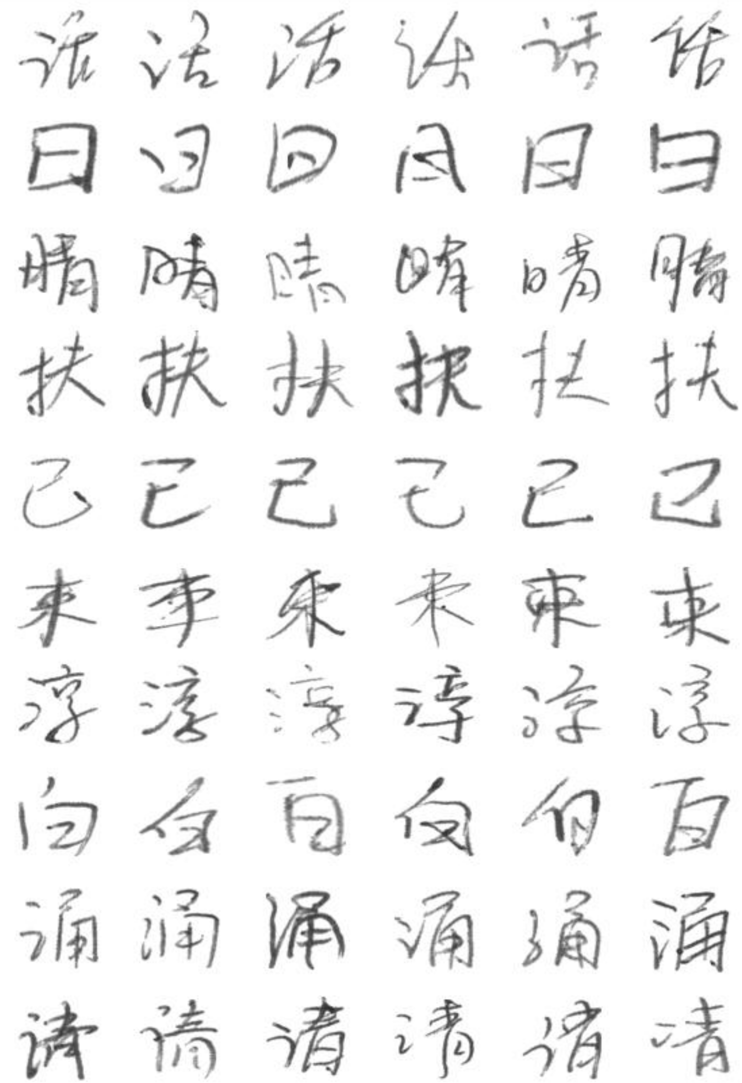
Method	Training Set	Error Rate (%)
THU [6]	CASIA-HWDB 1.0-1.1, 2.0-2.2	7.44
HIT [7]	CASIA-HWDB 1.0-1.1	7.38
Liu et al. [5]	CASIA-HWDB 1.0-1.1	7.28
Our ICDAR’13 [12]	CASIA-HWDB 1.1	5.23
MCDNN [22]	CASIA-HWDB 1.1	5.53
MCDNNs Voting [22]	CASIA-HWDB 1.1	4.35
R-CNN	CASIA-HWDB 1.1	5.32±0.09
R-CNNs Voting	CASIA-HWDB 1.1	4.45
ATR-CNN	CASIA-HWDB 1.1	4.96±0.08
ATR-CNNs Voting	**CASIA-HWDB 1.1**	**3.94**
Human [12]	-	3.87

Experiments – Handwriting Chinese

FUJITSU

- Misclassified Samples
  - Top 10 errors
  - Ground-truth -> Prediction
- Difficulties
  - Cursive writing
  - Touching strokes
  - Confusion in shapes

Error	Examples
话→活	
曰→日	
晴→睛	
扶→快	
己→已	
束→柬	
淳→谆	
白→自	
涌→诵	
请→清	

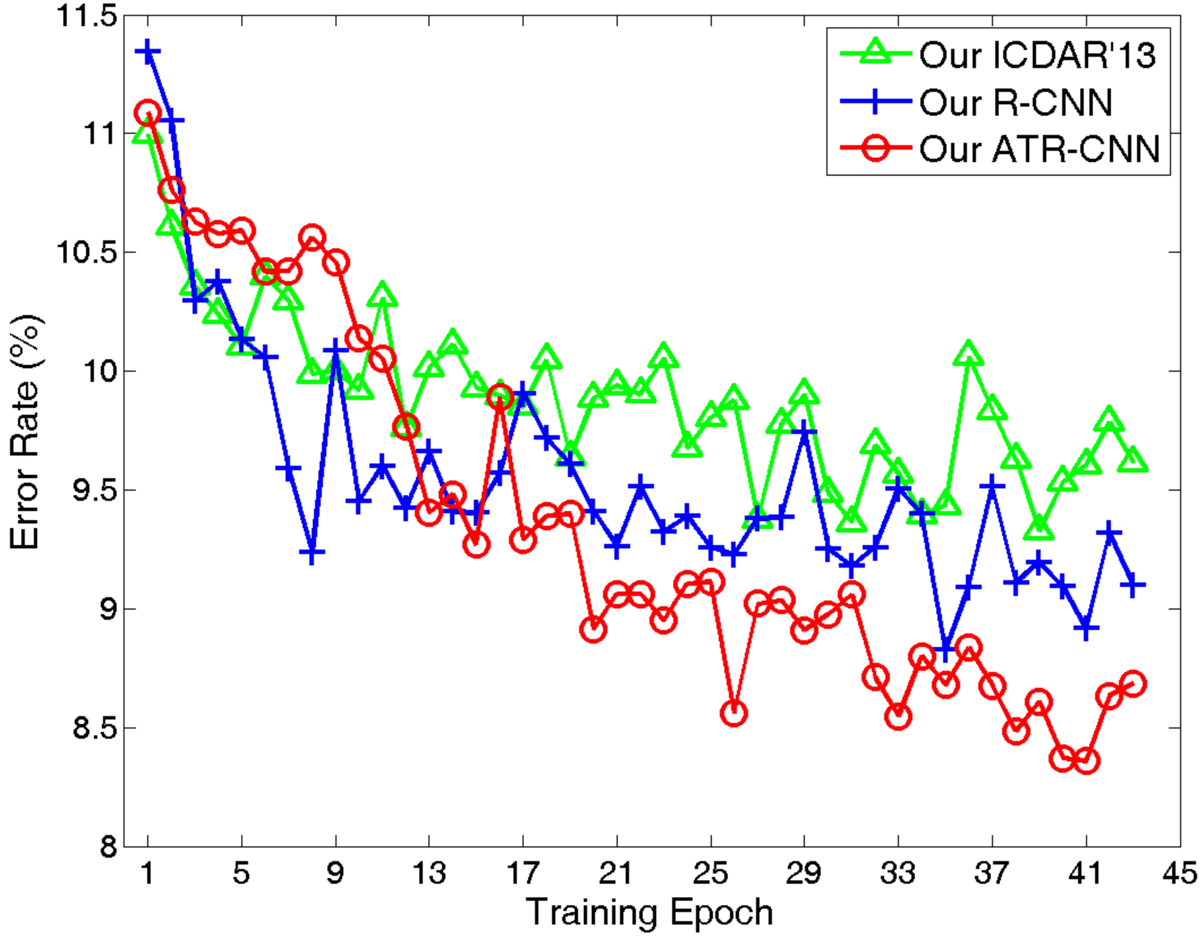


Experiments – Handwriting Chinese

FUJITSU

Contributions

- Relaxation (Blue curve), Alternate Training (Red curve)
- Both contribute to the improvement of recognition accuracy



Conclusions

- R-CNN
  - Neurons within a feature map ***do not*** share the same kernel
  - Endow CNN with more expressive power
- ATR-CNN
  - ***Randomly stop*** one layer from learning at one epoch
  - Regularize R-CNN
- Experiments
  - Both contribute to the improvement of recognition accuracy

Questions?

shaping tomorrow with you