Cut and Compare: End-to-end Offline Signature Verification Network

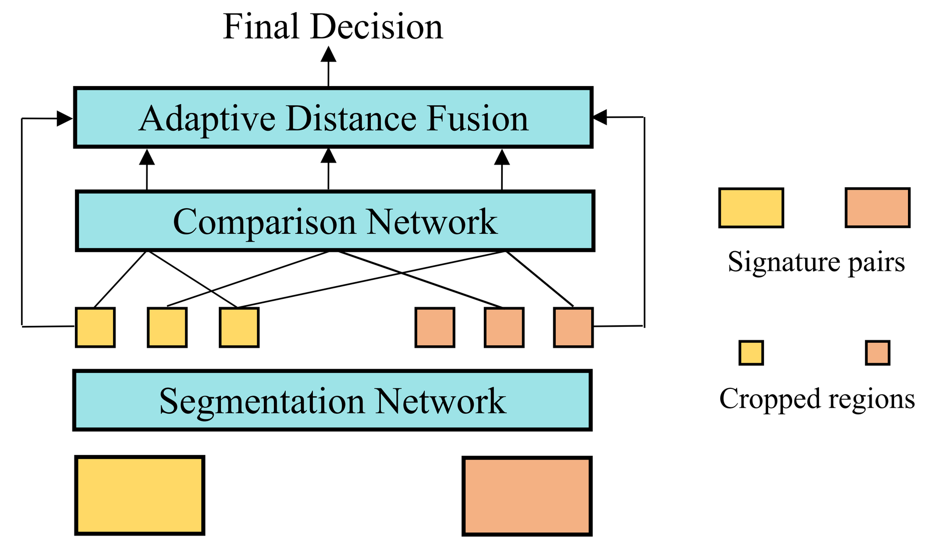
Xi Lu[1], Linlin Huang[1], Fei Yin[2]
[1] *Beijing Jiaotong University, Beijing, China*
[2] *Institute of Automation, Chinese Academy of Sciences, Beijing, China*

Introduction

- Offline signature verification is to verify whether a handwritten signature image is produced by a claimed person or not. Extensive researches have been conducted in this field due to its potential applications in finance, archeology, criminal investigation, and so forth. Despite the constant improvements made by the previous methods, the detailed features of signatures, which are usually embedded in local regions, have not been exploited sufficiently.
- We propose a novel end-to-end network, named *Cut and-Compare*, which can learn discriminative and informative regions automatically by a modified STN model and the regions are compared by an ARC model. To address the intrapersonal variability problem, we design a smoothed double-margin loss to train the network.
- The proposed cut-and-compare network and smoothed double-margin loss are shown to be effective in experiments and state-of-the-art performance is achieved on several datasets, including CEDAR, GPDS Synthetic, BHSig-H and BHSig-B, which are of different languages.

Architecture

The network consists of a segmentation network, a comparison network and an adaptive fusion module:



⇒ **Segmentation Network** Based on the Spatial Transformer Network (STN), the handwritten signature image is fed into a multi-head localization network (see right side) to get a quantity of transformation parameters. And the sampler generates a segment of the input image for each θ. Assuming A to be a 2D affine transformation matrix, the point-wise transformation can be written as

$$\begin{pmatrix} x_i^s \\ y_i^s \end{pmatrix} = \mathrm{A}\begin{pmatrix} x_i^t \\ y_i^t \\ 1 \end{pmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}\begin{pmatrix} x_i^t \\ y_i^t \\ 1 \end{pmatrix}$$

For the only purpose of cropping transformation, let

$$\mathrm{A} = \begin{bmatrix} a & 0 & c \\ 0 & e & f \end{bmatrix}$$

$$a = \frac{W'}{W}, e = \frac{H'}{H}, \theta = \begin{bmatrix} c \\ f \end{bmatrix}$$

Using learned θ, the sampling operation can generate output image regions with help of bilinear interpolation.

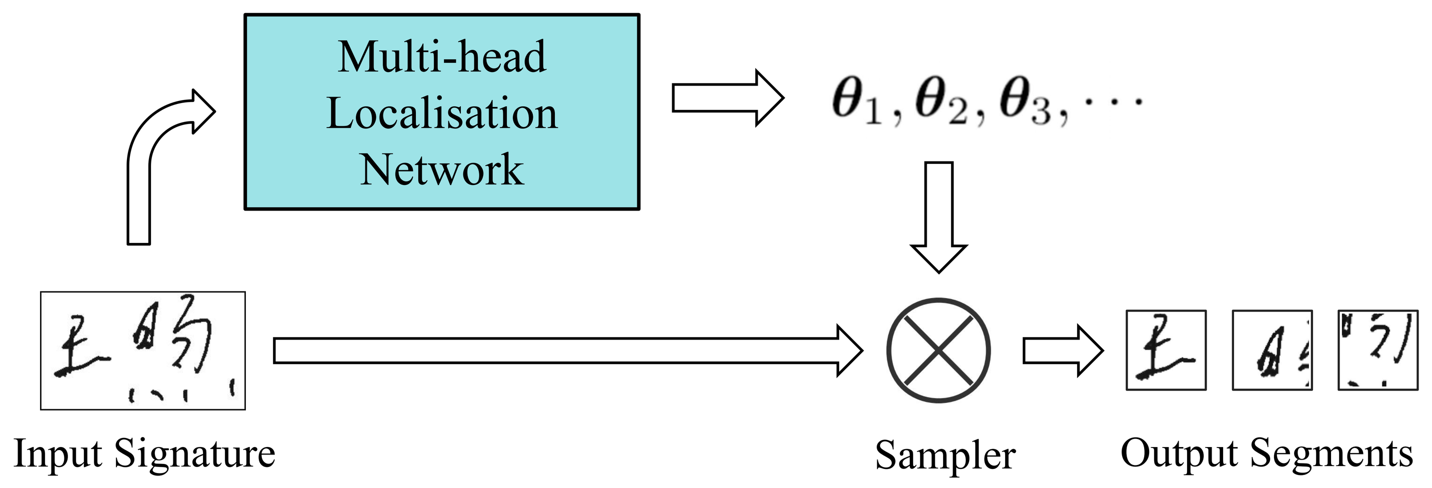
⇒ **Comparison Network** Our comparison network is based on the Attentive Recurrent Comparator (ARC), which can compare two images alternately in an attentive way. For more details, wo refer to the original paper.

⇒ **Adaptive Distance Fusion** For a pair of input handwritten signature image, the comparison network outputs several distances between each corresponding region pair. We utilize the ratio of foreground pixel values to weight the distances.

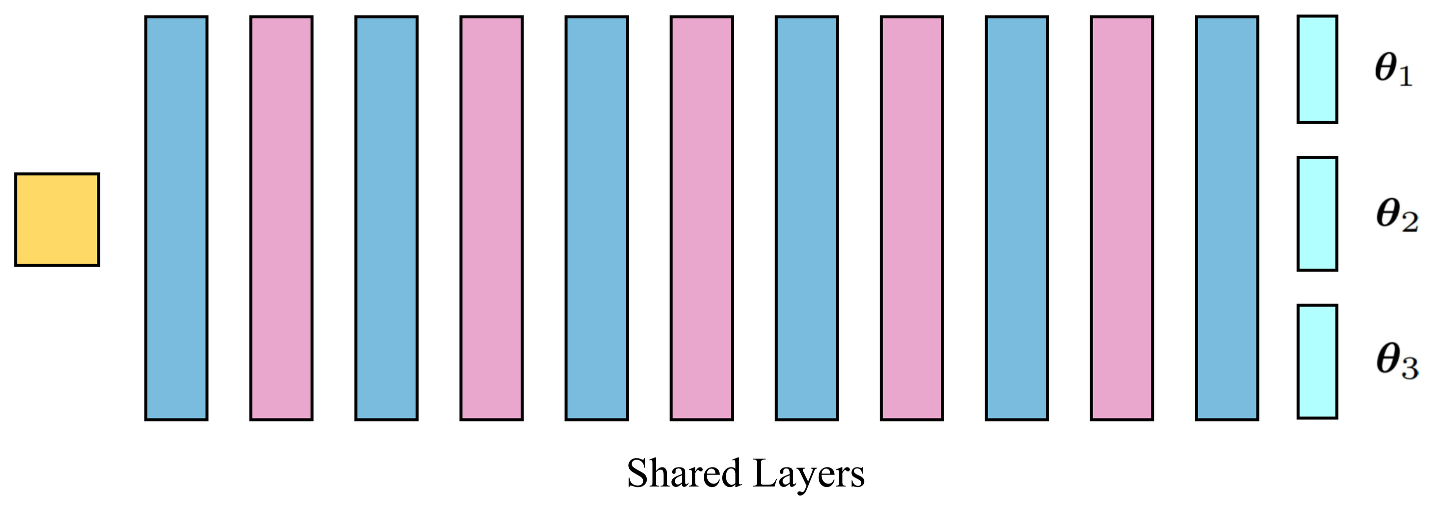
⇒ **Smoothed Double-margin Loss** Suppose the output of the signature verification model is dis, namely, the distance between the two input signature images. The proposed smooth double-margin loss is

$$L = y\left\{\frac{\ln(e^{\alpha(dis-m)}+1)}{\alpha}\right\} + (1-y)\left\{\frac{\ln(e^{\beta(n-dis)}+1)}{\beta}\right\}$$

Segmentation Network



Multi-head Localisation Network



Color code used: yellow = input signature image, blue = Conv + BN + ReLU, pink = max-pooling layer, cyan = fully-connected layer.

Comparison with the state-of-the-arts

Databases	Methods	Training		Testing		FAR	FRR	EER
		#signers	sampling	#signers	sampling			
CEDAR	Surroundedness Feature [2]	50	✓	5	✓	8.33	8.33	8.33
	Compact Correlated Feature [1]	50	✓	5	✓	0.00	0.00	**0.00**
	Signet [9]	50	✓	5	✓	0.00	0.00	**0.00**
	DeepHSV [12]	50	✓	5	×	0.00	0.00	**0.00**
	IDN [13]	50	✓	5	✓	5.87	2.17	3.62
	Region Based Metric Learning [11]	50	×	5	×	-	-	4.55
	Convolution Siamese Network [10]	50	✓	5	✓	-	-	8.50
	Ours	50	×	5	×	4.34	4.34	**4.34 / 0.00**[a]
GPDS Synthetic	Signet [9]	3200	✓	800	✓	22.24	22.24	22.24
	DeepHSV [12]	3500	✓	500	×	-	-	9.95
	Region Based Metric Learning [11]	2000	×	2000	×	-	-	8.89
	Convolution Siamese Network [10]	2000	✓	2000	✓	-	-	10.37
	Ours	2000	×	2000	×	7.87	7.87	**7.87**
BHSig-H	Compact Correlated Feature [1]	100	✓	60	✓	13.10	15.09	14.10
	Signet [9]	100	✓	60	✓	15.36	15.36	15.36
	DeepHSV [12]	100	✓	60	×	-	-	13.34
	IDN [13]	100	✓	60	✓	8.99	4.93	6.96
	Ours	100	×	60	×	5.97	5.97	**5.97**
BHSig-B	Compact Correlated Feature [1]	50	✓	50	✓	15.78	14.43	15.10
	Signet [9]	50	✓	50	✓	13.89	13.89	13.89
	DeepHSV [12]	80	✓	20	×	-	-	11.92
	IDN [13]	50	✓	50	✓	4.12	5.24	4.68
	Ours	50	×	50	×	3.96	3.96	**3.96**
ChnSig	Region Based Metric Learning [11]	1000	×	243	×	-	-	**9.91**
	Ours	1000	×	243	×	10.21	10.21	10.21

Comparison of different loss functions