

A Holistic Approach to Offline Handwritten Chinese and Japanese Text Line Recognition



Huu-Tin Hoang*, Chun-Jen Peng*, Hung Vinh Tran, Hung Le
Cinnamon AI
Email: {tin, larry, xing, toni}@cinnamon.is

Huy Hoang Nguyen
University of Oulu
Email: huy.nguyen@oulu.fi

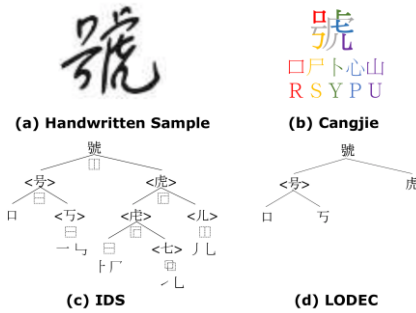


1. Introduction

Normally, we would encode the characters set of Japanese and Chinese with one-hot. However, this encoding method is costly and not represent the semantic information of characters.

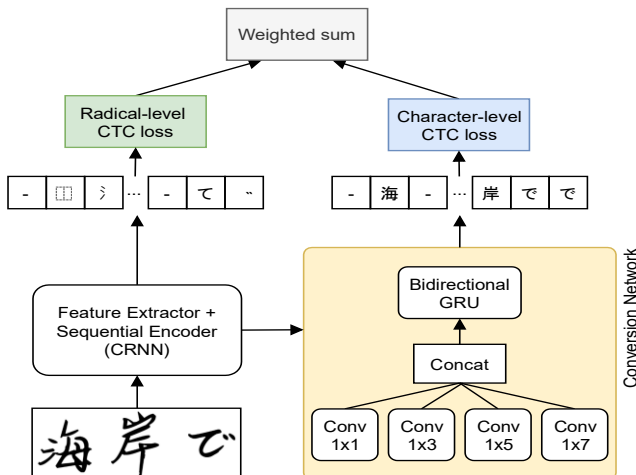
Therefore, we proposed a novel encoding method LODEC and a deep learning model LODENet that leverages auxiliary ground truths generated by LODEC or other radical-based encoding methods.

2. LODEC and Character Decomposition



Based on an observation that ones tend to write a radical, in a single stroke with a specific cursive pattern, we proposed **LODEC**, which targets to identify unique shapes of logographic characters rather than fundamental glyphs or partial shapes as in IDS or Cangjie, respectively.

3. LODENet

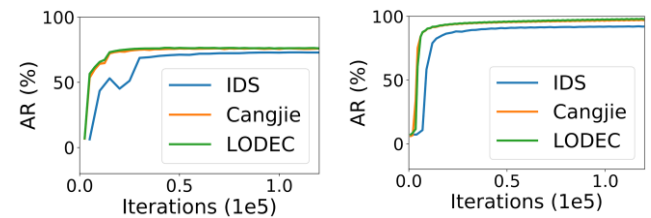


LODENet Architecture and end-to-end training scheme

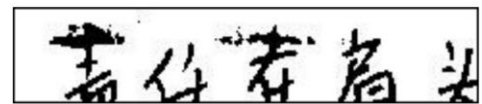
4. Quantitative Results

CASIA Dataset		3.37%
Wang et al.	88.79%	
LODENet	86.62%	
LODENet + Wikipedia Text	92.16%	
SCUT-EPT Dataset		1.64%
CNN + LSTM + CTC	75.97%	
LODENet	76.61%	
LODENet + Wikipedia Text	77.61%	

5. Qualitative Results



Learning curve on SCUT-EPT with different encodings



Ground truth	責任在肩頭
CRNN	責任"寿高头 (4)
LODENet	責任"在肩頭 (1)
Radical output	𠂇 貝 一 一 士 土 一 尸 月 头

6. Conclusions

- **LODEC encoding** that can fully represent all logograms and syllabic characters of Chinese and Japanese.
- An **end-to-end training scheme** that can be plugged in any radical-based encoding method.
- **LODENet architecture** equipped with the conversion network that learns to transcribe Japanese and Chinese contents from radical-based features.
- **SOTA results** on CASIA and SCUT-EPT, and one private Japanese dataset.