DUET: Detection Utilizing Enhancement for Text in Scanned or Captured Documents

Eun-Soo Jung, HyeongGwan Son, Kyusam Oh,
Yongkeun Yun, Soonhwan Kwon, and Min Soo Kim

Background and Motivation

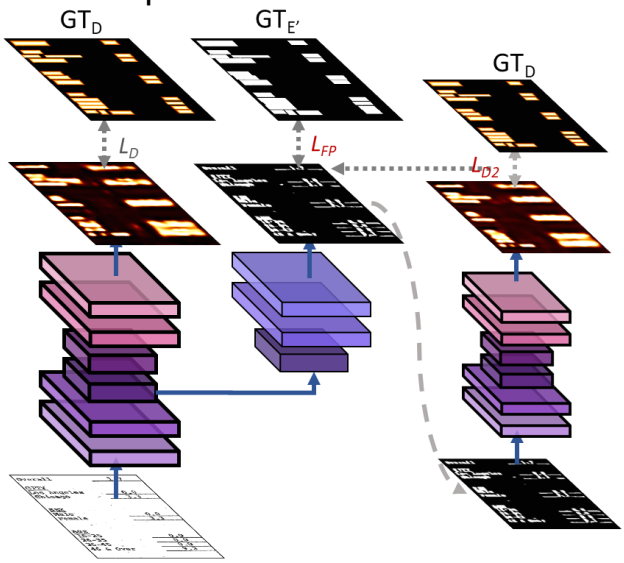
- Most of the previous studies on text detection focus on text in the wild[1-6]
- **Text in the wild** (or scene text) :(Relatively) More labeled data
- **Text in documents**
  : Very few labeled data
  → insufficient to train DNN
- Text detection models trained with scene text
  : *Limited* to cover features of document images

Contributions

- Text detector with improved accuracies for document images
  - Synthesizing training data
  - Multi-task learning
  - Overcome various types of noise in document images

Training

- **Main task**: text detection
- **Auxiliary task**: text enhancement
- **Weak-supervision** to train enhancement task for real data
  - Binaraized detection GT (ground truth)
    → false positive loss
  - Using interim trained detector
    → detection loss for enhanced output

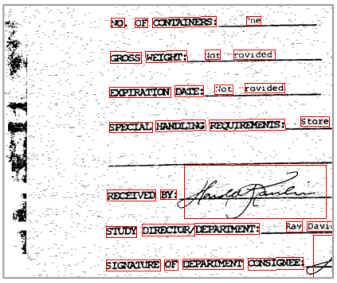


Results

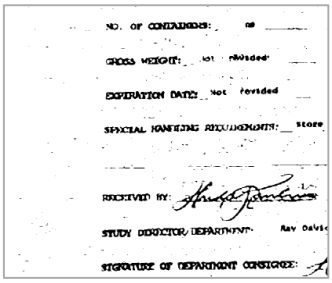
- Test data: **FUNSD** dataset

Method	Precision	Recall	F-score
Faster R-CNN	70.4	84.8	76.9
EAST	51.6	84.0	63.9
CRAFT	91.2	84.2	87.6
CharNet	**95.1**	57.4	71.6
DUET *(proposed)*	93.1	**92.2**	**92.6**

- Output examples



Detection



Enhancement

ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

SAMSUNG SDS