LIST OF FIGURES

1.1 Schematic diagram of the scene text detection 4

2.1 Example illustration of Stroke Width Transform. Taken from [4]. 6
2.2 Example illustration of Maximally Stable Extremal Regions. Below each image is the extracted MSERs, green regions represent character regions and red regions represent non-character regions. Taken from [25]. 7
2.3 Example illustration of CLAHE. 10
2.4 Example illustration of Dilation. 11
2.5 Example illustration of quadtree in image representation. 11
2.6 Schematic diagram of EAST [39] 13

3.1 Flowchart of the proposed scene text detection method. Process with bold text is the main contribution of this thesis. 15
3.2 Flowchart of the preprocessing 16
3.3 Example illustration of the contrast enhancement. 17
3.4 Example illustration of the CTR. 18
3.5 Schematic diagram of detection network [39]. 19
3.6 Example result of detection network. The green box is called bounding box which represent the text location. 20
3.7 Example illustration of Bounding Boxes Filtering 21

4.1 MSRA-TD500 ground truth format [34] 23
4.2 IC13 ground truth [28] 24
4.3 IC15 ground truth [28] 25
4.4 Example of SynthText dataset [1] 25
4.5 Example result of CTR with homogeneous threshold value set to 1. The green boxes indicate the candidate text. 29
4.6 Example result of the CTR with homogeneous threshold value set to 5 and 10. The green boxes indicate the candidate text. 30
4.7 Example result of CTR with sets of minimum region size threshold. The green boxes indicate the candidate text. 30
4.8 Example result of CTR with sets of dilation kernel size. The green boxes indicate the candidate text. 31
4.9 Example result of CTR with and without contrast enhancement. The green boxes indicate the candidate text. 32
4.10 Example of images that contain small text. 33
4.11 Example of image that contains small text. (a) Detection network result. (b) Detection network + PCTR result. (c) Ground truth. (d) CTR map. . 33

4.12 Result of with and without preprocessing and CTR on low contrast data. First row is the result of without preprocessing and CTR. Second row is the result of with preprocessing and CTR. Third row is the ground truth. . . . 34
4.13 Example illustration of proposed method with and without merging process on MSRA-TD500 dataset. Green boxes is the ground truth and red boxes is the prediction. . . . 36
4.14 Example illustration of the proposed bounding boxes merging limitation on curved text. . . . 37
4.15 Example of text that was not labeled as text in the ground truth detected as text. (a) Proposed method result. (b) Ground truth. . . . 38
4.16 Example of text like (pattern) object detected as text. (a) Proposed method result. (b) Ground truth. . . . 38
4.17 Example result of false positive reduction. (a) Detection network result on MSRA-TD500. (b) Detection network result on IC13. (c) Detection network result on IC15. (d) Detection result on MSRA-TD500. (e) Detection result on IC13. (f) Detection result on IC15. . . . 39
4.18 Example of proposed method detection result. (a) MSRA-TD500. (b) IC13. (c) IC15. . . . 40