INVESTIGATING THE EFFECT OF BD-CRAFT TO TEXT DETECTION ALGORITHMS

Clarisa V. Albarillo[1] and Proceso L. Fernandez, Jr.[2]

[1]College of Computer Science, Don Mariano Marcos Memorial State University, La Union, Philippines
[2]Department of Information Systems & Computer Science, Ateneo de Manila University, Quezon City, Philippines

ABSTRACT

With the rise and development of deep learning, computer vision and document analysis has influenced the area of text detection. Despite significant efforts in improving text detection performance, it remains to be challenging, as evident by the series of the Robust Reading Competitions. This study investigates the impact of employing BD-CRAFT – a variant of CRAFT that involves automatic image classification utilizing a Laplacian operator and further preprocess the classified blurry images using blind deconvolution to the top-ranked algorithms, SenseTime and TextFuseNet. Results revealed that the proposed method significantly enhanced the detection performances of the said algorithms. TextFuseNet + BD-CRAFT achieved an outstanding h-mean result of 93.55% and shows an impressive improvement of over 4% increase to its precision yielding 95.71% while SenseTime + BD-CRAFT placed first with a very remarkable 95.22% h-mean and exhibited a huge precision improvement of over 4%.

KEYWORDS

Blind Deconvolution, Computer Vision, Image Classification, Information Retrieval, Image Processing

1. INTRODUCTION

With the rise and development of deep learning, computer vision has been tremendously transformed and reshaped. A wave of change brought about by deep learning has unavoidably influenced a specific area of research in computer vision and document analysis called text detection. Driven by the recent progress in deep learning, impressive performances have been achieved in text detection.

High level semantics embodied in scene texts are both rich and clear thus can serve as important cues in computer vision [1, 2, 3, 4, 5] due to the wide range of applications. These applications include image search [6, 7], target geolocation [8, 9], human-computer interaction [10, 11], robot navigation [12, 13], and industrial automation [14, 15], which greatly benefits from the detailed information included in text.

Despite advances in the field, text detection remains difficult and challenging due to the diversity of text patterns and complex scene structures in natural images. Text pattern diversity and scene image complexities include arbitrary font size and type, cluttered image background, and light conditioning variation. Several challenges also still exist, such as noise, blur, distortion, occlusion, and variance. Owing to the inevitable challenges and complexities, traditional text detection algorithms often entail many processing steps such as character/word candidate creation, candidate filtering, and grouping. They frequently struggle to get each module to

DOI: 10.5121/ijaia.2023.14102

perform effectively, taking significant effort in adjusting settings and designing heuristic algorithms, decreasing detection speed [16].

The launch of the Robust Reading Competition proves that the difficulty of detecting texts in scene images has piqued the interest of many researchers over time. This competition began in 2011 and was co-sponsored by the International Conference on Document Analysis and Recognition (ICDAR). The competition is structured around a number of sophisticated computer vision challenges that cover a wide range of real-world scenarios. One such challenge is the Focused Scene Text, which focuses on reading texts in real-world contexts. The scenario being examined includes "focused text"—textual images that are largely focused on the text content of interest. The Focused Scene Text Challenge has three editions: ICDAR 2011, ICDAR 2013, and ICDAR 2015. For the tasks of text localization, text segmentation, and word recognition, the ICDAR 2013 is the definitive one [17].

Additionally, the competition features three performance metrics: ICDAR 2013 evaluation, Deteval, and Intersection-over-Union (IoU), each of which has a ranking of the best algorithms. The h-means of the top-ranking algorithm, SenseTime, for these three metrics are 96.38%, 96.78% and 93.62% respectively. These shows that the third metric (IoU) appears to be the most difficult, hence this study is focused on improving the IoU results.

The algorithms having the best IoU h-mean in the above-mentioned competition include: SenseTime (the top-ranked algorithm) whose h-mean is 93.62%, TextFuseNet with 93.11% hmean, TencentAIlLab with 93.05%, VARCO with 91.71%, HIT with 91.48% h-mean and CRAFT with 91.41%. In SenseTime, a single end-to-end trainable Fast Oriented Text Spotting (FOTS) network that is designed for simultaneous detection and recognition is used. To share convolutional features across detection and identification, it specifically introduced RoIRotate [18].

Table 1. State-of-the-art results on Text Localization, based on the IoU performance metric (see Robust Reading Competition - IoU Evaluation of Task 1)

Method	Precision	Recall	H-Mean
Sensetime (2016)	**91.87%**	**95.45%**	**93.62%**
TextFuseNet (2020)	90.78%	95.58%	93.11%
TencentAILab (2017)	94.79%	91.37%	93.05%
VARCO (2020)	89.86%	93.63%	91.71%
HIT (2020)	89.22%	93.85%	91.48%
CRAFT (2018)	89.04%	93.93%	91.42%

One of the algorithms that yielded good results is the Character Region Awareness for Text detection (CRAFT) [19]. It is a text detector that localizes the individual character regions and then correlates the detected characters to a text instance and uses a convolutional neural network to generate both the affinity score, which collects all the characters into one instance, and the character region score, which is used to localize specific characters in an image. Currently, the CRAFT algorithm ranks sixth in the competition, with an IoU h-mean of 91.42%. Though CRAFT's performance is already commendable, there is still much room for improvement because it assumes that the images of ICDAR 2013 are free from any blur or image distortion. Likewise, current text detection algorithms also treats the input images to be clear and does not employ image preprocessing prior to running the text detection algorithm hence the idea of imploring image classification and image deblurring through Blind Deconvolution as image preprocessing was conceptualized.

A previous study [20] explored how blind deconvolution can be applied to improve text localization and recognition of texts in both clear and blurry datasets using a fast-bounding box algorithm. It employed manual classification of images into blurry and non-blurry classes. The findings of the study reveal that the performance results of the proposed text localization and recognition method using manual classification of images and blind deconvolution is significantly improved.

Moreover, the researchers were prompted to investigate further on the impact of using Laplacian operator in automatic image classification and blind deconvolution for deblurring. This study [21] explores on improving the detection performance of CRAFT by adding some pre-processing steps that include automatically detecting blurry images and then attempting to reduce the blur on these images prior to running the CRAFT algorithm using blind deconvolution. This proposed method is referred as BD-CRAFT, a variant of CRAFT. The resulting technique was shown to be not only significantly better than CRAFT, but it is also able to outperform the current best state-of-the-art algorithm, SenseTime, for scene text detection.

In this study, combining BD-CRAFT with two state-of-the-art algorithms to further improve their text detection performance is investigated.

2. METHODOLOGY

2.1. The Dataset

The primary focus of this research revolves on the impact of BD-CRAFT in the text detection performance of some algorithms hence the International Conference on Document Analysis and Recognition (ICDAR) 2013 Focused Scene Text Competition Challenge 2 dataset was used. The ICDAR 2013 Challenge 2 dataset is composed of images that are specifically focused on the relevant text content to depict the use case in which a user focuses a camera on a scene text for text reading and translation applications. As such, the focus text is horizontal in most cases.

It is composed of 229 training images, and 233 test images and provides “Strong”, “Weak” and “Generic” lexicons - of different sizes for the task of text detection, similar to ICDAR 2015 [22, 23]. On the other hand, recent computer vision datasets including 2017 COCO-Text [24], deTEXT [25], DOST [26], FSNS [27], MLT [28], and IEHHR [29] were not used because the focus of this study is on scene images. Moreover, the images were of various sizes and were taken in various lighting conditions and environments using different cameras. Some of the images are shown in Figure 1.

Figure 1. Sample Images from ICDAR 2013 Focused Scene Text Competition Challenge 2 Datasets

2.2. Experimentations with the BD-CRAFT Algorithm

In the recent study [21], BD-CRAFT, a variant of CRAFT, was proposed which significantly improved the performance of CRAFT. The huge improvement is due to two principle image preprocessing techniques that are executed before running the main method CRAFT. First, it utilizes the Laplacian operator with a threshold of 100 to recognize blurry input images automatically. The identified blurry images are then deblurred using Blind Deconvolution, with the point spread function (PSF) set at 1,3 since this set of PSF values produced the best results among the other investigated PSFs. Figure 2 is a flowchart describing the key processes of BDCRAFT.

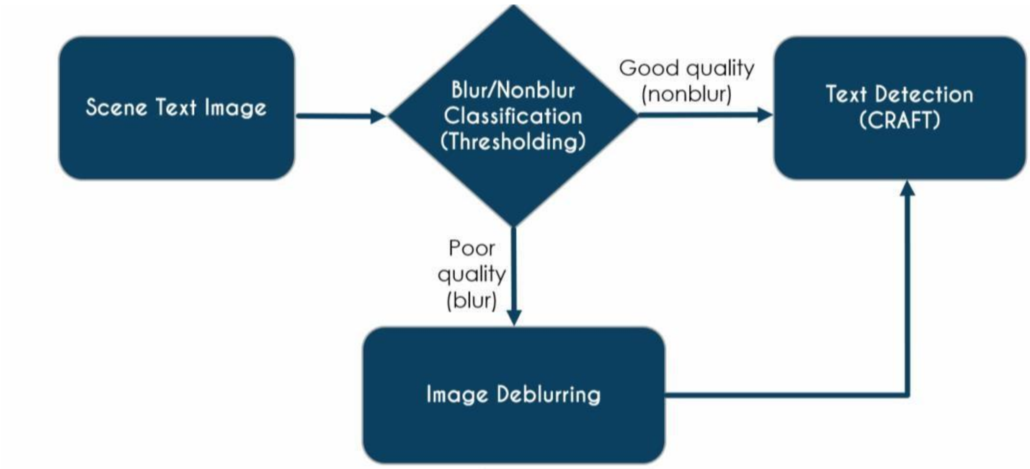

Figure 2. Flowchart of the proposed BD-CRAFT

The major focus of this research is on combining BD-CRAFT with certain state-of-the-art algorithms. The impact of combining the proposed technique is applied to the two top-performing algorithms such as SenseTime and TextFuseNet. Since their source codes are not available for download, the published results in the Robust Reading Competition, which included the actual performances (precision, recall, and h-mean) of each image in the ICDAR 2013, were utilized. The approach only uses the 63 recognized blurry images, and the target algorithm's performance results (SenseTime or TextFuseNet) are gathered and compared to the results of each image when running BD-CRAFT.

2.3. The Evaluation Metrics

To measure the accuracy of the proposed method, Intersection over Union (IoU), a well-known similarity metric that is measured as the ratio of two entities - the overlapping area and the union area - is utilized. IoU assesses how well the predicted bounding box overlaps with the ground truth box in the problem context.

The IoU scale ranges from 0 (no overlap) to the ideal value of 1 (perfect overlap). An IoU score of 0.5 or higher is regarded as a good prediction and indicates that the texts are correctly located [30]. In order to assess the IoU performance of the proposed algorithm, the precision, recall, and eventually the h-mean are computed from this prediction. The formulas for these well-known measures are provided below for completeness:

$$Precision = \frac{TP}{TP+FP} \quad \text{or} \quad Precision(G,D) = \frac{\sum_{j=1}^{|D|} Bestmatch_D(D_j)}{|D|}$$

$$Recall = \frac{TP}{TP+FN} \quad \text{or} \quad Recall(G,D) = \frac{\sum_{i=1}^{|G|} Bestmatch_G(G_i)}{|G|}$$

and

$$Hmean = 2\frac{(Recall * Precision)}{(Recall + Precision)}$$

where, *Bestmatch*$_d$ and *Bestmatch*$_g$ indicate the closest match between detection and ground truth as defined below:

$$Bestmatch_G(G_i) = \frac{2 \cdot Area(G_i \cap D_j)}{Area(G_i) + Area(D_j)}$$

$$Bestmatch_D(D_j) = \frac{2 \cdot Area(D_j \cap G_i)}{Area(D_j) + Area(G_i)}$$

Note that H-mean refers to the harmonic mean of the Precision and Recall and therefore takes into account both false positives (FP) and false negatives (FN).

2.4. Comparison with the State-of-the-Art Algorithms

The final part of this study involves providing proof of the superiority of the results of the top two algorithms against those of the modified versions (i.e. the versions that incorporate BD-CRAFT). Towards this end, the performance results on the ICDAR 2013 dataset of the target algorithms are collected and compared with the result of the modified versions (BD-CRAFT + SenseTime or BD-CRAFT + TextFuseNet). The contributory value of BD-CRAFT is established after showing that each of these three algorithms yield better IoU h-means after incorporating BDCRAFT.

3. Results and Discussion

Several studies and experiments were carried out using the ICDAR 2013 dataset in an attempt to investigate the effect of the proposed BD-CRAFT to other text detection algorithms.

3.1. The Effect of Blind Deconvolution to the Identified Blurry Images

Establishing concrete evidence that the proposed technique has an effect to the original CRAFT, the results of the 63 identified blurry images with their respective threshold of blurriness are shown in Table 1. The precision, recall and h-mean of the said images are also presented and were compared to the previous result of the original CRAFT.

As shown in the table, when blind deconvolution (BD-CRAFT) was applied to the (automatically) detected blurry images, 13 of the images had improved performance (h-mean), while 11 of the images had improved recall and precision when compared to the results when applying the original CRAFT algorithm. However, 7 images performed less when utilizing BD-CRAFT than when using standard CRAFT, as seen by lower h-mean results. There are 9 images, nevertheless, that when BD-CRAFT is used, obtained lower h-mean results. Overall, the text detection performance of CRAFT is improved when Blind Deconvolution (BD-CRAFT) is applied.

Table 1. Evaluation Performance of the 63 Blurry Images using BD-CRAFT vs CRAFT

Blurry Images	**Blurriness**	**CRAFT**			**BD-CRAFT**		
Img_no	**Threshold**	**Precision**	**Recall**	**H-Mean**	**Precision**	**Recall**	**H-Mean**
1	12.8405571	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
4	16.4762852	80.00%	57.14%	66.67%	**100.00%**	**85.71%**	**92.31%**
10	23.4311781	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
12	36.653517	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
15	28.9903253	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
16	16.1522402	0.00%	0.00%	0.00%	**100.00%**	**50.00%**	**66.67%**
20	57.4975666	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
21	81.7039225	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
23	15.9194037	100.00%	100.00%	100.00%	**94.12%**	**100.00%**	**96.87%**
25	7.97007126	33.33%	50.00%	40.00%	**100.00%**	**100.00%**	**100.00%**
26	35.1926793	90.91%	76.92%	83.33%	**100.00%**	**100.00%**	**100.00%**
29	50.7909314	100.00%	60.00%	75.00%	**75.00%**	**60.00%**	**66.67%**
34	16.1037522	100.00%	100.00%	100.00%	**45.45%**	**100.00%**	**62.50%**
38	14.3117793	100.00%	100.00%	100.00%	**100.00%**	**100.00%**	**100.00%**
39	15.4544795	88.89%	80.00%	84.21%	**83.33%**	**100.00%**	**90.91%**
44	27.4993814	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
45	21.6460108	76.47%	86.67%	81.25%	**100.00%**	**93.33%**	**96.55%**
48	17.6449657	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
49	56.3667426	100.00%	66.67%	80.00%	**100.00%**	**83.33%**	**90.90%**
52	99.1999698	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
54	39.2921302	71.43%	83.33%	76.92%	**100.00%**	**100.00%**	**100.00%**
55	83.6990885	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
58	41.8884998	100.00%	77.78%	87.50%	100.00%	77.78%	87.50%
59	97.4825748	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
61	61.5662565	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
62	64.4650431	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
63	66.767582	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
64	18.0964371	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
65	12.5359277	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
69	28.2970963	100.00%	100.00%	100.00%	**50.00%**	**100.00%**	**66.67%**
73	78.597045	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

75	62.6589596	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
77	88.4263075	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
82	56.0501649	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
84	22.8809663	100.00%	83.33%	90.91%	**71.43%**	**83.33%**	**76.92%**
85	68.5711141	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
87	49.7199117	100.00%	100.00%	100.00%	**80.00%**	**100.00%**	**88.89%**
88	81.8170171	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
89	14.206154	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
93	13.5899878	50.00%	100.00%	66.67%	**100.00%**	**100.00%**	**100.00%**
95	57.426799	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
96	11.8740355	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
98	61.3983121	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
123	54.5686799	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
125	29.4063831	83.33%	71.43%	76.92%	**100.00%**	**100.00%**	**100.00%**
131	65.3736027	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
134	45.5904438	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
138	72.2330041	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
142	73.576066	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
143	24.0986572	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
180	10.3440649	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
181	84.1780428	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
183	81.9879789	100.00%	100.00%	100.00%	**100.00%**	**50.00%**	**66.67%**
184	26.0136719	80.00%	100.00%	88.89%	**100.00%**	**100.00%**	**100.00%**
186	35.5071598	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
211	22.1367073	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
217	95.5176921	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
222	65.4059773	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
226	43.1260097	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
227	44.3401323	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
229	16.0993804	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
231	60.5148326	50.00%	33.33%	40.00%	**83.33%**	**100.00%**	**90.91%**
232	17.4316837	75.00%	100.00%	85.71%	**100.00%**	**100.00%**	**100.00%**

The performance of BD-CRAFT in comparison to SenseTime is displayed in Table 2. The table shows that when BD-CRAFT was applied, 11 images performed better in terms of h-mean than SenseTime, whereas 8 and 9 images performed better in terms of precision and recall. However, when BD-CRAFT is used, there are additional 9 images that had lower h-mean results. Overall, BD-CRAFT outperforms SenseTime in terms of performance.

Table 3 displays the assessment results for the 63 detected blurry images using TextFuseNet and BD-CRAFT. According to the ICDAR 2013 Focused Scene Text Detection Challenge results that have been published, TextFuseNet is ranked second. Nine images outperformed TextFuseNet in terms of performance when BD-CRAFT was used, while 5 images improved in terms of precision and 8 images improved in terms of recall.

However, when utilizing TextFuseNet, 8 images produce higher h-mean results. When comparing the number of images that performed better, it is clear that BD-CRAFT achieves superior outcomes to TextFuseNet. As a consequence, TextFuseNet's detection performance was 93.11%, while BD-detection CRAFT's performance was 94.47% on average.

Table 2. Evaluation Performance of the 63 Blurry Images using BD-CRAFT vs SenseTime

Blurry Images	**Blurriness**	**BD-CRAFT**			**SenseTime**		
Img_no.	**Threshold**	**Precision**	**Recall**	**H-Mean**	**Precision**	**Recall**	**H-Mean**
1	12.8405571	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
4	16.4762852	**100.00%**	**85.71%**	**92.31%**	71.43%	71.43%	71.43%
10	23.4311781	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
12	36.653517	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
15	28.9903253	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
16	16.1522402	**100.00%**	**50.00%**	**66.67%**	100.00%	100.00%	100.00%
20	57.4975666	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
21	81.7039225	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
23	15.9194037	**94.12%**	**100.00%**	**96.87%**	100.00%	100.00%	100.00%
25	7.97007126	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
26	35.1926793	**100.00%**	**100.00%**	**100.00%**	100.00%	84.62%	91.67%
29	50.7909314	**75.00%**	**60.00%**	**66.67%**	80.00%	80.00%	80.00%
34	16.1037522	**45.45%**	**100.00%**	**62.50%**	100.00%	100.00%	100.00%
38	14.3117793	**100.00%**	**100.00%**	**100.00%**	50.00%	100.00%	66.67%
39	15.4544795	**83.33%**	**100.00%**	**90.91%**	66.67%	80.00%	72.73%
44	27.4993814	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
45	21.6460108	**100.00%**	**93.33%**	**96.55%**	100.00%	100.00%	100.00%
48	17.6449657	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
49	56.3667426	**100.00%**	**83.33%**	**90.90%**	75.00%	75.00%	75.00%
52	99.1999698	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
54	39.2921302	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
55	83.6990885	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
58	41.8884998	100.00%	77.78%	87.50%	100.00%	77.78%	87.50%
59	97.4825748	**100.00%**	**100.00%**	**100.00%**	75.00%	75.00%	75.00%
61	61.5662565	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
62	64.4650431	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
63	66.767582	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
64	18.0964371	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
65	12.5359277	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
69	28.2970963	**50.00%**	**100.00%**	**66.67%**	100.00%	100.00%	100.00%
73	78.597045	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
75	62.6589596	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
77	88.4263075	**100.00%**	**100.00%**	**100.00%**	50.00%	50.00%	50.00%
82	56.0501649	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
84	22.8809663	**71.43%**	**83.33%**	**76.92%**	80.00%	66.67%	72.73%
85	68.5711141	**100.00%**	**100.00%**	**100.00%**	50.00%	100.00%	66.67%
87	49.7199117	80.00%	100.00%	88.89%	80.00%	100.00%	88.89%
88	81.8170171	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
89	14.206154	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
93	13.5899878	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
95	57.426799	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
96	11.8740355	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
98	61.3983121	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
123	54.5686799	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
125	29.4063831	**100.00%**	**100.00%**	**100.00%**	100.00%	71.43%	83.33%
131	65.3736027	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
134	45.5904438	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
138	72.2330041	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

142	73.576066	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
143	24.0986572	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
180	10.3440649	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
181	84.1780428	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
183	81.9879789	**100.00%**	**50.00%**	**66.67%**	75.00%	75.00%	75.00%
184	26.0136719	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
186	35.5071598	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
211	22.1367073	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
217	95.5176921	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
222	65.4059773	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
226	43.1260097	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
227	44.3401323	**100.00%**	**100.00%**	**100.00%**	100.00%	75.00%	85.71%
229	16.0993804	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
231	60.5148326	**83.33%**	**100.00%**	**90.91%**	100.00%	100.00%	100.00%
232	17.4316837	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Table 3. Evaluation Performance of the 63 Blurry Images using BD-CRAFT vs TextFuseNet

Blurry Images	**Blurriness**	**BD-CRAFT**			**TextFuseNet**		
Img_no.	**Threshold**	**Precision**	**Recall**	**H-Mean**	**Precision**	**Recall**	**H-Mean**
1	12.8405571	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
4	16.4762852	**100.00%**	**85.71%**	**92.31%**	100.00%	57.14%	72.73%
10	23.4311781	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
12	36.653517	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
15	28.9903253	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
16	16.1522402	**100.00%**	**50.00%**	**66.67%**	100.00%	100.00%	100.00%
20	57.4975666	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
21	81.7039225	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
23	15.9194037	**94.12%**	**100.00%**	**96.87%**	100.00%	100.00%	100.00%
25	7.97007126	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
26	35.1926793	**100.00%**	**100.00%**	**100.00%**	100.00%	84.62%	91.67%
29	50.7909314	**75.00%**	**60.00%**	**66.67%**	100.00%	60.00%	75.00%
34	16.1037522	**45.45%**	**100.00%**	**62.50%**	100.00%	100.00%	100.00%
38	14.3117793	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
39	15.4544795	**83.33%**	**100.00%**	**90.91%**	0.00%	0.00%	0.00%
44	27.4993814	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
45	21.6460108	**100.00%**	**93.33%**	**96.55%**	81.25%	86.67%	83.87%
48	17.6449657	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
49	56.3667426	**100.00%**	**83.33%**	**90.90%**	100.00%	75.00%	85.71%
52	99.1999698	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
54	39.2921302	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
55	83.6990885	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
58	41.8884998	**100.00%**	**77.78%**	**87.50%**	100.00%	88.89%	94.12%
59	97.4825748	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
61	61.5662565	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
62	64.4650431	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
63	66.767582	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
64	18.0964371	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
65	12.5359277	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
69	28.2970963	**50.00%**	**100.00%**	**66.67%**	0.00%	0.00%	0.00%
73	78.597045	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
75	62.6589596	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

77	88.4263075	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
82	56.0501649	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
84	22.8809663	**71.43%**	**83.33%**	**76.92%**	100.00%	83.33%	90.91%
85	68.5711141	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
87	49.7199117	**80.00%**	**100.00%**	**88.89%**	100.00%	100.00%	100.00%
88	81.8170171	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
89	14.206154	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
93	13.5899878	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
95	57.426799	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
96	11.8740355	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
98	61.3983121	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
123	54.5686799	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
125	29.4063831	**100.00%**	**100.00%**	**100.00%**	62.50%	71.43%	66.67%
131	65.3736027	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
134	45.5904438	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
138	72.2330041	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
142	73.576066	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
143	24.0986572	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
180	10.3440649	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
181	84.1780428	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
183	81.9879789	**100.00%**	**50.00%**	**66.67%**	100.00%	100.00%	100.00%
184	26.0136719	100.00%	100.00%	100.00%	80.00%	100.00%	88.89%
186	35.5071598	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
211	22.1367073	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
217	95.5176921	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
222	65.4059773	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
226	43.1260097	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
227	44.3401323	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
229	16.0993804	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
231	60.5148326	**83.33%**	**100.00%**	**90.91%**	80.00%	66.67%	72.73%
232	17.4316837	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

The updated ranking of the state-of-the-art algorithms for Text Detection shows BD-CRAFT ranked top (see Table 4)! Observe that BD-CRAFT is superior to SenseTime which is 0.85% (absolute) higher than SenseTime and 3.05% higher than that of the original CRAFT. The top ranking performance was achieved because of the very impressive precision of 95.24%, which tops all precision results in the table, and is significantly higher than SenseTime, by over 3%!

Table 4. Comparison of the State-of-the-art algorithms ranked by IoU H-Mean

Method	Precision	Recall	H-Mean
BD-CRAFT	**95.24%**	**93.72%**	**94.47%**
SenseTime (2016)	91.87%	95.45%	93.62%
BD-CRAFT	94.32%	92.44%	93.37%
TextFuseNet (2020)	90.78%	95.58%	93.11%
TencentAILab (2017)	94.79%	91.37%	93.05%
VARCO (2020)	89.86%	93.63%	91.71%
HIT (2020)	89.22%	93.85%	91.48%
CRAFT (2018)	89.04%	93.93%	91.42%

To demonstrate that the proposed method has an impact on improving the performance of the various scene text detection algorithms, the images that obtained better results in BD-CRAFT compared to the results obtained using other algorithms such as SenseTime and TextFuseNet are chosen and eventually used to calculate the average h-mean result of the said algorithm.

Table 5 compares the assessment findings of the 63 images produced by SenseTime + BDCRAFT to those obtained by SenseTime alone. It is apparent that 11 images generated superior results, which has a significant impact on the SenseTime's h-mean results. The application of BDCRAFT (SenseTime + BD-CRAFT) produced a very remarkable 95.22% h-mean and demonstrated a huge improvement in precision over 4%, making it the top-ranked algorithm.

Meanwhile, Table 6 reveals the evaluation results when using TextFuseNet + BD-CRAFT. It can be shown that other algorithms' h-mean results improved when BD-CRAFT was used. TextFuseNet alone produces an h-mean of 93.11%, but when combined with BD-CRAFT, it produces an h-mean of 93.55%. Additionally, the precision indicates an outstanding improvement of over 4% when BD-CRAFT (TextFuseNet + BD-CRAFT) is used, as seen by its 95.71% precision.

Table 5. Evaluation Performance of the 63 Blurry Images using SenseTime + BD-CRAFT vs SenseTime

Blurry Images	**Blurriness**	**SenseTime + BD-CRAFT**			**SenseTime**		
Img_no.	**Threshold**	**Precision**	**Recall**	**H-Mean**	**Precision**	**Recall**	**H-Mean**
1	12.8405571	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
4	16.4762852	**100.00%**	**85.71%**	**92.31%**	71.43%	71.43%	71.43%
10	23.4311781	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
12	36.653517	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
15	28.9903253	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
16	16.1522402	***100.00%***	***100.00%***	***100.00%***	100.00%	100.00%	100.00%
20	57.4975666	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
21	81.7039225	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
23	15.9194037	***100.00%***	***100.00%***	***100.00%***	100.00%	100.00%	100.00%
25	7.97007126	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
26	35.1926793	**100.00%**	**100.00%**	**100.00%**	100.00%	84.62%	91.67%
29	50.7909314	***80.00%***	***80.00%***	***80.00%***	80.00%	80.00%	80.00%
34	16.1037522	***100.00%***	***100.00%***	***100.00%***	100.00%	100.00%	100.00%
38	14.3117793	**100.00%**	**100.00%**	**100.00%**	50.00%	100.00%	66.67%
39	15.4544795	**83.33%**	**100.00%**	**90.91%**	66.67%	80.00%	72.73%
44	27.4993814	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
45	21.6460108	***100.00%***	***100.00%***	***100.00%***	100.00%	100.00%	100.00%
48	17.6449657	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
49	56.3667426	**100.00%**	**83.33%**	**90.90%**	75.00%	75.00%	75.00%
52	99.1999698	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
54	39.2921302	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
55	83.6990885	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
58	41.8884998	100.00%	77.78%	87.50%	100.00%	77.78%	87.50%
59	97.4825748	**100.00%**	**100.00%**	**100.00%**	75.00%	75.00%	75.00%
61	61.5662565	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
62	64.4650431	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
63	66.767582	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
64	18.0964371	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
65	12.5359277	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
69	28.2970963	***100.00%***	***100.00%***	***100.00%***	100.00%	100.00%	100.00%
73	78.597045	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

75	62.6589596	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
77	88.4263075	**100.00%**	**100.00%**	**100.00%**	50.00%	50.00%	50.00%
82	56.0501649	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
84	22.8809663	**71.43%**	**83.33%**	**76.92%**	80.00%	66.67%	72.73%
85	68.5711141	**100.00%**	**100.00%**	**100.00%**	50.00%	100.00%	66.67%
87	49.7199117	80.00%	100.00%	88.89%	80.00%	100.00%	88.89%
88	81.8170171	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
89	14.206154	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
93	13.5899878	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
95	57.426799	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
96	11.8740355	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
98	61.3983121	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
123	54.5686799	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
125	29.4063831	**100.00%**	**100.00%**	**100.00%**	100.00%	71.43%	83.33%
131	65.3736027	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
134	45.5904438	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
138	72.2330041	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
142	73.576066	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
143	24.0986572	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
180	10.3440649	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
181	84.1780428	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
183	81.9879789	***75.00%***	***75.00%***	***75.00%***	75.00%	75.00%	75.00%
184	26.0136719	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
186	35.5071598	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
211	22.1367073	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
217	95.5176921	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
222	65.4059773	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
226	43.1260097	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
227	44.3401323	**100.00%**	**100.00%**	**100.00%**	100.00%	75.00%	85.71%
229	16.0993804	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
231	60.5148326	***100.00%***	***100.00%***	***100.00%***	100.00%	100.00%	100.00%
232	17.4316837	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Table 6. Evaluation Performance of the 63 Blurry Images using BD-CRAFT + TextFuseNet vs TextFuseNet

Blurry Images	**Blurriness**	**TextFuseNet + BD-CRAFT**			**TextFuseNet**		
Img_no.	**Threshold**	**Precision**	**Recall**	**H-Mean**	**Precision**	**Recall**	**H-Mean**
1	12.8405571	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
4	16.4762852	**100.00%**	**85.71%**	**92.31%**	100.00%	57.14%	72.73%
10	23.4311781	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
12	36.653517	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
15	28.9903253	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
16	16.1522402	***100.00%***	***100.00%***	***100.00%***	100.00%	100.00%	100.00%
20	57.4975666	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
21	81.7039225	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
23	15.9194037	***100.00%***	***100.00%***	***100.00%***	100.00%	100.00%	100.00%
25	7.97007126	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
26	35.1926793	**100.00%**	**100.00%**	**100.00%**	100.00%	84.62%	91.67%
29	50.7909314	***100.00%***	***60.00%***	***75.00%***	100.00%	60.00%	75.00%
34	16.1037522	***100.00%***	***100.00%***	***100.00%***	100.00%	100.00%	100.00%
38	14.3117793	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

39	15.4544795	**83.33%**	**100.00%**	**90.91%**	0.00%	0.00%	0.00%
44	27.4993814	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
45	21.6460108	**100.00%**	**93.33%**	**96.55%**	81.25%	86.67%	83.87%
48	17.6449657	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
49	56.3667426	**100.00%**	**83.33%**	**90.90%**	100.00%	75.00%	85.71%
52	99.1999698	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
54	39.2921302	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
55	83.6990885	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
58	41.8884998	***100.00%***	***88.89%***	***94.12%***	100.00%	88.89%	94.12%
59	97.4825748	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
61	61.5662565	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
62	64.4650431	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
63	66.767582	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
64	18.0964371	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
65	12.5359277	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
69	28.2970963	**50.00%**	**100.00%**	**66.67%**	0.00%	0.00%	0.00%
73	78.597045	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
75	62.6589596	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
77	88.4263075	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
82	56.0501649	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
84	22.8809663	***100.00%***	***83.33%***	***90.91%***	100.00%	83.33%	90.91%
85	68.5711141	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
87	49.7199117	***100.00%***	***100.00%***	***100.00%***	100.00%	100.00%	100.00%
88	81.8170171	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
89	14.206154	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
93	13.5899878	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
95	57.426799	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
96	11.8740355	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
98	61.3983121	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
123	54.5686799	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
125	29.4063831	**100.00%**	**100.00%**	**100.00%**	62.50%	71.43%	66.67%
131	65.3736027	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
134	45.5904438	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
138	72.2330041	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
142	73.576066	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
143	24.0986572	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
180	10.3440649	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
181	84.1780428	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
183	81.9879789	***100.00%***	***100.00%***	***100.00%***	100.00%	100.00%	100.00%
184	26.0136719	100.00%	100.00%	100.00%	80.00%	100.00%	88.89%
186	35.5071598	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
211	22.1367073	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
217	95.5176921	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
222	65.4059773	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
226	43.1260097	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
227	44.3401323	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
229	16.0993804	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
231	60.5148326	**83.33%**	**100.00%**	**90.91%**	80.00%	66.67%	72.73%
232	17.4316837	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

An updated ranking of the top performing algorithms would show an impressive ranking of this proposed algorithm when integrated to the original algorithm (see Table 7). It can be observed that when BD-CRAFT is applied to other algorithms, their h-mean results were improved. When using TextFuseNet alone, the h-mean result is 93.11% but when BD-CRAFT is (TextFuseNet + BD-CRAFT) applied, it yields 93.55% h-mean. Additionally, the precision indicates an outstanding improvement of over 4% when BD-CRAFT (TextFuseNet + BD-CRAFT) is used, as seen by its 95.71% precision.

Moreover, the state-of-the-art algorithm SenseTime is also explored. When applying BD-CRAFT (SenseTime + BD-CRAFT), it resulted in a very impressive 95.22% h-mean and showed huge precision improvement over 4% which made it to retain its place to be the top-ranked algorithm.

Table 7. Comparison of the State-of-the-art algorithms ranked by IoU h-mean

Method	**Precision**	**Recall**	**H-Mean**
SenseTime + BD-CRAFT	**96.79 %**	**94.65 %**	**95.22 %**
SenseTime (2016)	91.87%	95.45%	93.62%
TextFuseNet + BD-CRAFT	***95.71%***	***94.45%***	***93.55%***
BD-CRAFT	94.32%	92.44%	93.37%
TextFuseNet (2020)	90.78%	95.58%	93.11%
TencentAILab (2017)	94.79%	91.37%	93.05%
VARCO (2020)	89.86%	93.63%	91.71%
HIT (2020)	89.22%	93.85%	91.48%
CRAFT (2018)	89.04%	93.93%	91.42%

4. CONCLUSION AND RECOMMENDATION

In this study, we incorporate the BD-CRAFT, a variation of the CRAFT algorithm that involves pre-processing steps where images were automatically classified as blurry or non-blurry using a Laplacian operator, then applying a deblurring method known as Blind Deconvolution. This study improves two (2) state-of-the-art algorithms for text detection, SenseTime and TextFuseNet. The increase in the overall h-mean as well as some of the precision and recall values demonstrate the considerable improvements in each of the ensuing algorithm versions. TextFuseNet + BDCRAFT produces an h-mean of 93.55%, while the precision shows a remarkable improvement of over 4%, as seen by its precision of 95.71%. Additionally, SenseTime + BD-CRAFT was the topranked algorithm with a very remarkable 95.22% h-mean and a significant precision improvement of over 4%. Evidence suggests that when BD-CRAFT is integrated with other algorithms, their performances are improved; as a result, BD-CRAFT significantly affects the performance of text detection algorithms.

It would be interesting to investigate other state-of-the-art scene text detection algorithms as potential areas for future research because they could gain from the discussed technique as well. Additionally, it could be beneficial to explore other pre-processing techniques that can be applied to text detection algorithms in general.

ACKNOWLEDGEMENTS

The completion of this undertaking could not have been possible without the participation and assistance of many individuals whose names may not all be enumerated, their contributions are sincerely appreciated and gratefully acknowledged. Through the guidance, blessings and graces of the Lord Almighty, the preparation of this undertaking was made possible. It is with deep appreciation that we offer this piece of work to CHED, Ateneo de Manila University and DMMMSU for their support towards the fulfillment of this scholarly endeavor. It is the kindness of these acknowledged that this research sees the light of the day. We submit this work with great humility and utmost regard.

REFERENCES

[1] Tian, Z., Huang, W., He, T., He, P., & Qiao, Y. (2016, October). Detecting text in natural image with connectionist text proposal network. In European conference on computer vision (pp. 56-72). Springer, Cham.

[2] Shi, B., Bai, X., & Belongie, S. (2017). Detecting oriented text in natural images by linking segments. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 25502558).

[3] Shi, B., Bai, X., & Yao, C. (2016). An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. *IEEE transactions on pattern analysis and machine intelligence*, *39*(11), 2298-2304.

[4] Zhou, X., Yao, C., Wen, H., Wang, Y., Zhou, S., He, W., & Liang, J. (2017). East: an efficient and accurate scene text detector. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 5551-5560).

[5] He, P., Huang, W., Qiao, Y., Loy, C. C., & Tang, X. (2016). Reading scene text in deep convolutional sequences. In *Thirtieth AAAI conference on artificial intelligence*.

[6] Zhu, Y., Yao, C., & Bai, X. (2016). Scene text detection and recognition: Recent advances and future trends. *Frontiers of Computer Science*, *10*(1), 19-36.

[7] Gordo, A., Almazán, J., Revaud, J., & Larlus, D. (2016, October). Deep image retrieval: Learning global representations for image search. In *European conference on computer vision* (pp. 241257). Springer, Cham.

[8] Utomo, M. N. Y., Adji, T. B., & Ardiyanto, I. (2018, March). Geolocation prediction in social media data using text analysis: A review. In *2018 International Conference on Information and Communications Technology (ICOIACT)* (pp. 84-89). IEEE.

[9] Renshaw, C. K., Hewitt, J., Grimming, R., & Driggers, R. G. (2022, July). Sensor Optimization for Image-Based Geolocation. In *Computational Optical Sensing and Imaging* (pp. JF2D-6). Optica Publishing Group.

[10] Akhtiamov, O., Sidorov, M., Karpov, A. A., & Minker, W. (2017, August). Speech and Text Analysis for Multimodal Addressee Detection in Human-Human-Computer Interaction. In *INTERSPEECH* (pp. 2521-2525).

[11] Brodić, D., & Amelio, A. (2016, September). Analysis of the human-computer interaction on the example of image-based CAPTCHA by association rule mining. In *International Workshop on Symbiotic Interaction* (pp. 38-51). Springer, Cham.

[12] Nguyen, D., Shijian, L., Ouarti, N., & Mokhtari, M. (2017, March). Text-Edge-Box: An Object Proposal Approach for Scene Texts Localization. In *Applications.*

[13] Panchpor, A. A., Shue, S., & Conrad, J. M. (2018, January). A survey of methods for mobile robot localization and mapping in dynamic indoor environments. In *2018 Conference on Signal Processing And Communication Engineering Systems (SPACES)* (pp. 138-144). IEEE.

[14] Cheng, Q. (2021, June). The application of computer vision technology in the field of industrial automation. In *International Conference on Applications and Techniques in Cyber Security and Intelligence* (pp. 274-279). Springer, Cham.

[15] Gao, Y., Li, X., Wang, X. V., Wang, L., & Gao, L. (2021). A review on recent advances in visionbased defect recognition towards industrial intelligence. *Journal of Manufacturing Systems*.

[16] Liao, M., Shi, B., Bai, X., Wang, X., & Liu, W. (2017, February). Textboxes: A fast text detector with a single deep neural network. In *Thirty-first AAAI conference on artificial intelligence*.

[17] Robust Reading Competition - Focused Scene Text Overview. 2020. https://rrc.cvc.uab.es/?ch=2&com=introduction, Online. accessed October 3, 2021.

[18] Kim, S., & Park, T. (2020). Learning Robust Feature Representations for Scene Text Detection. arXiv preprint arXiv:2005.12466.

[19] Baek, Y., Lee, B., Han, D., Yun, S., & Lee, H. (2019). Character region awareness for text detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 9365-9374).

[20] Albarillo C., & Fernandez, P. (2019). Blind Deconvolution Model for Scene Text Localization and Recognition. In *19th Philippine Computing Society Congress* (pp. 38-45).

[21] Albarillo, C. V., & Fernandez Jr, P. L. (2022). Impact of Automatic Image Classification and Blind Deconvolution in Improving Text Detection Performance of the CRAFT Algorithm. *arXiv preprint arXiv:2211.15999*.

[22] Karatzas, D., Gomez-Bigorda, L., Nicolaou, A., Ghosh, S., Bagdanov, A., Iwamura, M., ... & Valveny, E. (2015, August). ICDAR 2015 competition on robust reading. In 2015 13th International Conference on Document Analysis and Recognition (ICDAR) (pp. 1156-1160). IEEE.

[23] Ma, J., Shao, W., Ye, H., Wang, L., Wang, H., Zheng, Y., & Xue, X. (2018). Arbitraryoriented scene text detection via rotation proposals. IEEE Transactions on Multimedia, 20(11), 3111-3122.

[24] Veit, A., Matera, T., Neumann, L., Matas, J., & Belongie, S. (2016). Coco-text: Dataset and benchmark for text detection and recognition in natural images. *arXiv preprint arXiv:1601.07140*.

[25] Yang, C., Yin, X. C., Yu, H., Karatzas, D., & Cao, Y. (2017). ICDAR2017 robust reading challenge on text extraction from biomedical literature figures (DeTEXT). In 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR) (Vol. 1, pp. 1444-1447). IEEE.

[26] Iwamura, M., Matsuda, T., Morimoto, N., Sato, H., Ikeda, Y., & Kise, K. (2016). Downtown osaka scene text dataset. In European Conference on Computer Vision (pp. 440-455). Springer, Cham.

[27] Smith, R., Gu, C., Lee, D. S., Hu, H., Unnikrishnan, R., Ibarz, J., & Lin, S. (2016). Endto-end interpretation of the french street name signs dataset. In European Conference on Computer Vision (pp. 411-426). Springer, Cham.

[28] Nayef, N., Yin, F., Bizid, I., Choi, H., Feng, Y., Karatzas, D., & Ogier, J. M. (2017). Icdar2017 robust reading challenge on multi-lingual scene text detection and script identification-rrc-mlt. In 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR) (Vol. 1, pp. 14541459). IEEE.

[29] Fornés, A., Romero, V., Baró, A., Toledo, J. I., Sánchez, J. A., Vidal, E., & Lladós, J. (2017). ICDAR2017 competition on information extraction in historical handwritten records. In 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR) (Vol. 1, pp. 1389-1394). IEEE.

[30] Robust Reading Competition - IoU Evaluation of Task 1. 2020. https://rrc.cvc.uab.es/?ch=2&com=evaluation&task=1&f=1&e=3. Online. accessed October 3, 2021.

AUTHORS

Clarisa V. Albarillo is an Assistant Professor III of the Don Mariano Marcos Memorial State University and currently taking up PhD in Computer Science at the Ateneo de Manila University. She is a research enthusiast in the field of artificial intelligence, computer vision, image processing, and social research.

Proceso L. Fernandez Jr. is a Professor of Computer Science at the Ateneo de Manila University. His main research interests include those in the areas of Computer Vision, Machine Learning, and Algorithms and Complexity Theory. He currently also serves as the Director of the Ateneo Intellectual Property Office of said University.