



Asian Conference on Pattern Recognition

ACPR 2021: **Pattern Recognition** pp 218–233

DCT-DWT-FFT Based Method for Text Detection in Underwater Images

[Ayan Banerjee](#), [Palaiahnakote Shivakumara](#) , [Soumyajit Pal](#),
[Umapada Pal](#) & [Cheng-Lin Liu](#)

Conference paper | [First Online: 10 May 2022](#)

66 Accesses

Part of the [Lecture Notes in Computer Science](#) book series
(LNCS, volume 13189)

Abstract

Text detection in underwater images is an open challenge because of the distortions caused by refraction, absorption of light, particles, and variations depending on depth, color, and nature of water. Unlike existing methods aimed at text detection in natural scene images, in this paper, we have proposed a novel method for text detection in

This website sets only cookies which are necessary for it to function. They are used to enable core functionality such as security, network management and accessibility. These cookies cannot be switched off in our systems. You may disable these by changing your browser settings, but this may affect how the website functions. Please view our privacy policy for further details on how we process your information. Dismiss

Transform (DCT), Discrete Wavelet Transform (DWT), and Fast Fourier Transform (FFT) for image enhancement to highlight the text features. The enhanced image is fed to a modified Character Region Awareness for Text Detection (CRAFT) model to detect text in underwater images. To explore enhancement methods, we evaluate six combinations of image enhancement techniques, namely, DCT-DWT-FFT, DCT-FFT-DWT, DWT-DCT-FFT, DWT-FFT-DCT, FFT-DCT-DWT, FFT-DWT-DCT. Experimental results on our dataset of underwater images and benchmark datasets of natural scene text detection, namely, MSRA-TD500, ICDAR 2019 MLT, ICDAR 2019 ArT, Total-Text, CTW1500, and COCO Text show that the proposed method performs well for both underwater and natural scene images and outperforms the existing methods on all the datasets.

Keywords

Under water images

Text detection in underwater images

Image enhancement Discrete cosine transform

Wavelet transform Fourier transform

Modified Character Region Awareness for Text

This website sets only cookies which are necessary for it to function. They are used to enable core functionality such as security, network management and accessibility. These cookies cannot be switched off in our systems. You may disable these by changing your browser settings, but this may affect how the website functions. Please view our privacy policy for further details on how we process your information. Dismiss

▼ Chapter

EUR 24.95

Price excludes VAT (Malaysia)

- DOI: 10.1007/978-3-031-02444-3_16
- Chapter length: 16 pages
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy Chapter

▼ eBook

EUR 128.39

Price includes VAT (Malaysia)

- ISBN: 978-3-031-02444-3
- Instant PDF download
- Readable on all devices
- Own it forever
- Exclusive offer for individuals only
- Tax calculation will be finalised during checkout

Buy eBook

▼ Softcover Book

EUR 159.99

Price excludes VAT (Malaysia)

- ISBN: 978-3-031-02443-6
- Dispatched in 3 to 5 business days
- Exclusive offer for individuals only
- Free shipping worldwide
[Shipping restrictions may apply, check to see if you are impacted.](#)
- Tax calculation will be finalised during checkout

Buy Softcover Book

This website sets only cookies which are necessary for it to function. They are used to enable core functionality such as security, network management and accessibility. These cookies cannot be switched off in our systems. You may disable these by changing your browser settings, but this may affect how the website functions. Please view our privacy policy for further details on how we process your information. Dismiss

1. Xue, M., et al.: Deep invariant texture features for water image classification. SN Appl. Sci. **2**(12), 1–19 (2020). <https://doi-org.ezproxy.um.edu.my/10.1007/s42452-020-03882-w>

2. Kezebou, L., Oludare, V., Panetta, K., Againa, S.S.: Underwater object tracking benchmark and dataset. In: Proceedings of the HST (2019). <https://doi-org.ezproxy.um.edu.my/10.1109/HST47167.2019.9032954>

3. Wang, Y., Xie, H., Zha, Z.J., Xing, M., Fu, Z., Zhang, Y.: ContourNet: Taking a further step toward accurate arbitrary-shaped scene text detection. In: Proceedings of the CVPR, pp. 11753–11762 (2020)

4. Cao, Y., Ma, S., Pan, H.: FDTA: Fully convolutional scene text detection with text attention. IEEE Access 155441–155449 (2020)

5. Zhang, W., Xiang, S.: Face anti-spoofing detection based on DWT-LBP-DCT features. Signal Process. Image Commun. (2020). <https://doi-org.ezproxy.um.edu.my/10.1016/j.sipr.2020.101587>

This website sets only cookies which are necessary for it to function. They are used to enable core functionality such as security, network management and accessibility. These cookies cannot be switched off in our systems. You may disable these by changing your browser settings, but this may affect how the website functions. Please view our privacy policy for further details on how we process your information. Dismiss

region awareness for text detection. In:
Proceedings of the CVPR, pp. 9365–9374 (2019)

7. Liu, C., Yang, C., Hou, J.B., Wu, L.H., Zhu, X.B., Xiao, L.: GCCNet: Grouped channel composition network for scene text detection. *Neurocomputing* **454**, 135–151 (2021)

8. Shi, J., Chen, L., Su, F.: Accurate arbitrary-shaped scene text detection via iterative polynomial parameter regression. In: Ishikawa, H., Liu, C.-L., Pajdla, T., Shi, J. (eds.) ACCV 2020. LNCS, vol. 12624, pp. 241–256. Springer, Cham (2021).
https://doi-org.ezproxy.um.edu.my/10.1007/978-3-030-69535-4_15

9. Qin, X., Jiang, J., Yuan, C.A., Qiao, S., Fan, W.: Arbitrary shape natural scene text detection method based on soft attention mechanism and dilated convolution. *IEEE Access* 122685–122694 (2020)

10. Dai, P., Li, Y., Zhang, H., Li, J., Cao, X.: Accurate scene text detection via scale-aware data augmentation and shape similarity constraint

This website sets only cookies which are necessary for it to function. They are used to enable core functionality such as security, network management and accessibility. These cookies cannot be switched off in our systems. You may disable these by changing your browser settings, but this may affect how the website functions. Please view our privacy policy for further details on how we process your information. Dismiss

11. Hu, Z., Wu, X., Wang, J.: TCATD: text contour attention for scene text detection. In: Proceedings of the ICPR, pp. 1083–1088 (2021)

12. Liao, M., Lyu, P., He, M., Yao, C., Wu, W., Bai, X.: Mask TextSpotter: an end-to-end trainable neural network for spotting text with arbitrary shapes. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) Computer Vision – ECCV 2018. LNCS, vol. 11218, pp. 71–88. Springer, Cham (2018). https://doi-org.ezproxy.um.edu.my/10.1007/978-3-030-01264-9_5

13. Deng, G., Ming, Y., Xue, J.-H.: RFRN: A recurrent feature refinement network for accurate and efficient scene text detection. Neurocomputing **453**, 465–481 (2021)

14. Liu, J., Zhong, Q., Yuan, Y., Su, H., Du, B.: SemiText: scene text detection with semi-supervised learning. Neurocomputing **407**, 343–353 (2020)

15. You, M. et al. Arbitrarily-oriented text detection

This website sets only cookies which are necessary for it to function. They are used to enable core functionality such as security, network management and accessibility. These cookies cannot be switched off in our systems. You may disable these by changing your browser settings, but this may affect how the website functions. Please view our privacy policy for further details on how we process your information. Dismiss

16. Chowdhury, P.N., et al.: A new episodic learning-based network for text detection on human body in sports images. IEEE Trans Circuits Syst. Video Technol. (2021). <https://doi-org.ezproxy.um.edu.my/10.1109/TCSVT.2021.3092713>

17. Chowdhury, T., Shivakumara, P., Pal, U., Tong, L., Raghavendra, R., Chanda, S.: DCINN: deformable convolution and inception based neural network for tattoo text detection through skin region. In: Lladós, J., Lopresti, D., Uchida, S. (eds.) Document Analysis and Recognition – ICDAR 2021: 16th International Conference, Lausanne, Switzerland, September 5–10, 2021, Proceedings, Part II, pp. 335–350. Springer International Publishing, Cham (2021). https://doi-org.ezproxy.um.edu.my/10.1007/978-3-030-86331-9_22

18. Zhou, X., et al.: East: an efficient and accurate scene text detector. In: Proceedings of the CVPR, pp. 2642–2651 (2017)

This website sets only cookies which are necessary for it to function. They are used to enable core functionality such as security, network management and accessibility. These cookies cannot be switched off in our systems. You may disable these by changing your browser settings, but this may affect how the website functions. Please view our privacy policy for further details on how we process your information. Dismiss

-
20. Chng, C.K., Liu, Y., Sun, Y., Ng, C.C., Luo, C., Ni, Z.: ICDAR2019 robust reading challenge on arbitrarily-shaped text-RRC-ArT. In: Proceedings of the ICDAR, pp. 1571–1576 (2019)
-

Acknowledgements

The work of Cheng-Lin Liu was supported by the National Key Research and Development Program under Grant No. 2018AAA0100400 and the National Natural Science Foundation of China under Grant No. 61721004. This work also partially supported by TIH, ISI, Kolkata.

Author information

Authors and Affiliations

**Computer Vision and Pattern Recognition Unit,
Indian Statistical Institute, Kolkata, India**

Ayan Banerjee, Soumyajit Pal & Umapada Pal

**Department of Computer System and Technology,
Faculty of Computer Science and Information
Technology, University of Malaya, Kuala Lumpur,
Malaysia**

Palaiahnakote Shivakumara

This website sets only cookies which are necessary for it to function. They are used to enable core functionality such as security, network management and accessibility. These cookies cannot be switched off in our systems. You may disable these by changing your browser settings, but this may affect how the website functions. Please view our privacy policy for further details on how we process your information. Dismiss

**School of Artificial Intelligence, University of
Chinese Academy of Sciences, Beijing, China**

Cheng-Lin Liu

Corresponding author

Correspondence to [Palaiahnakote Shivakumara](#).

Editor information

Editors and Affiliations

Korea University, Seoul, Korea (Republic of)

Christian Wallraven

Nanjing University, Nanjing, China

Qingshan Liu

Osaka University, Osaka, Japan

Hajime Nagahara

Rights and permissions

[Reprints and Permissions](#)

Copyright information

© 2022 Springer Nature Switzerland AG

About this paper

Cite this paper

This website sets only cookies which are necessary for it to function. They are used to enable core functionality such as security, network management and accessibility. These cookies cannot be switched off in our systems. You may disable these by changing your browser settings, but this may affect how the website functions. Please view our privacy policy for further details on how we process your information. Dismiss

Computer Science, vol 13189. Springer, Cham. https://doi-org.ezproxy.um.edu.my/10.1007/978-3-031-02444-3_16

[.RIS](#)  [.ENW](#)  [.BIB](#) 

DOI

https://doi-org.ezproxy.um.edu.my/10.1007/978-3-031-02444-3_16

Published	Publisher Name	Print ISBN
10 May 2022	Springer, Cham	978-3-031-02443-6

Online ISBN	eBook Packages
978-3-031-02444-3	Computer Science
	Computer Science
	(R0)

Not logged in - 103.18.3.232

10122 SpringerLink Malaysia eJourna Consortium - Higher Education (3000716851) - University of Malaya UM (2000000861) - 13008 SpringerLink Malaysia eJournal Consortium – Higher Education (3001044276) - 8354 Springerlink Malaysia consortium (3000519906) - 6816 SpringerLink Malaysia eJournal Consortium - Higher Education (3000155375)

SPRINGER NATURE

© 2022 Springer Nature Switzerland AG. Part of [Springer Nature](#).

This website sets only cookies which are necessary for it to function. They are used to enable core functionality such as security, network management and accessibility. These cookies cannot be switched off in our systems. You may disable these by changing your browser settings, but this may affect how the website functions. Please view our privacy policy for further details on how we process your information. [Dismiss](#)