Region-based annotation data of fire images for intelligent surveillance system

By Wahyono

WORD COUNT 2082
TIME SUBMITTED 04-OCT-2023 01:19PM
PAPER ID 103214704

Data in Brief 41 (2022) 107925

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Region-based annotation data of fire images for intelligent surveillance system

Wahyono[a,*], Andi Dharmawan[a], Agus Harjoko[a], Chrystian[a], Faisal Dharma Adhinata[b]

[a] Department of Computer Science and Electronics, Universitas Gadjah Mada, Indonesia
[b] Institut Teknologi Telkom Purwokerto, Indonesia

ARTICLE INFO

Article history:
Received 21 October 2021
Revised 25 January 2022
Accepted 2 February 2022
Available online 4 February 2022

Keywords:
Fire
Intelligent surveillance system
Computer vision
Segmentation

ABSTRACT

This paper presents fire segmentation annotation data on 12 commonly used and publicly available "VisiFire Dataset" videos from http://signal.ee.bilkent.edu.tr/VisiFire/. This annotations dataset was obtained by per-frame, manual hand annotation over the fire region with 2684 total annotated frames. Since this annotation provides per-frame segmentation data, it offers a new and unique fire motion feature to the existing video, unlike other fire segmentation data that are collected from different still images. The annotations dataset also provides ground truth for segmentation task on videos. With segmentation task, it offers better insight on how well a machine learning model understood, not only detecting whether a fire is present, but also its exact location by calculating metrics such as Intersection over Union (IoU) with this annotations data. This annotations data is a tremendously useful addition to train, develop, and create a much better smart surveillance system for early detection in high-risk fire hotspots area.

* Corresponding author.
E-mail address: wahyo@ugm.ac.id.

https://doi.org/10.1016/j.dib.2022.107925
2352-34

Specifications Table

Subject	Computer Vision and Pattern Recognition
Specific subject area	Fire segmentation, image annotation, intelligent surveillance system, intersection over union
Type of data	Annotations Binary mask images Video
How the data were acquired	Each raw video converted into frames image format. Per-frame fire segmentation annotations then acquired by manual hand annotation using VIA (VGG Image Annotator) tools [1]. Binary mask images were created from previous fire region annotations data by python script [4]. Fire segmentation videos were created by converting previous binary mask images to videos.
Data format	Fire segmentation videos: AVI Fire segmentation annotations: CSV and JSON Binary mask images: JPG
Description of data collection	Our dataset consists of 12 fire segmentation videos, 12 fire segmentation annotations (in JSON project files or CSV annotations format), and 2684 total binary mask images collected from all fire videos.
Data source location	Source video 1. http://signal.ee.bilkent.edu.tr/VisiFire/Demo/FireClips/controlled1.avi 2. http://signal.ee.bilkent.edu.tr/VisiFire/Demo/FireClips/controlled2.avi 3. http://signal.ee.bilkent.edu.tr/VisiFire/Demo/FireClips/controlled3.avi 4. http://signal.ee.bilkent.edu.tr/VisiFire/Demo/FireClips/forest1.avi 5. http://signal.ee.bilkent.edu.tr/VisiFire/Demo/FireClips/forest2.avi 6. http://signal.ee.bilkent.edu.tr/VisiFire/Demo/FireClips/forest3.avi 7. http://signal.ee.bilkent.edu.tr/VisiFire/Demo/FireClips/forest4.avi 8. http://signal.ee.bilkent.edu.tr/VisiFire/Demo/FireClips/forest5.avi 9. http://signal.ee.bilkent.edu.tr/VisiFire/Demo/FireClips/fBackYardFire.avi 10. http://signal.ee.bilkent.edu.tr/VisiFire/Demo/FireClips/ForestFire1.avi 11. http://signal.ee.bilkent.edu.tr/VisiFire/Demo/FireClips/fire1.avi 12. http://signal.ee.bilkent.edu.tr/VisiFire/Demo/FireClips/40m_PanFire_20060824.avi Raw videos data are available publicly online on [2]: • Institution: Bilkent University • City: Ankara • Country: Turkey
Data accessibility	Repository name: Region-based Annotation Data of Fire Images for Intelligent Surveillance System Data identification number: https://doi.org/10.5281/zenodo.5893854 Direct URL to data: https://zenodo.org/record/5893854

Value of the Data

- This dataset offers new insight on fire motion feature. Current other existing fire segmentation data was only consisted of independent, different still images [3]. With this new continuous video annotations data, this will open and further develop better new methods of fire detection and segmentation.
- This dataset added significant value of the fire location data in each video frame with semantic segmentation to the existing dataset. This data is essential for better fire detection model to not only able to detect if there was a fire, but also shows the precise location.
- In practical implementation, this dataset is valuable for computer vision researcher to experiment, develop, and create an intelligent surveillance system that can be used for early fire detection on high-risk fire hotspots area, preventing injuries, and other major losses.

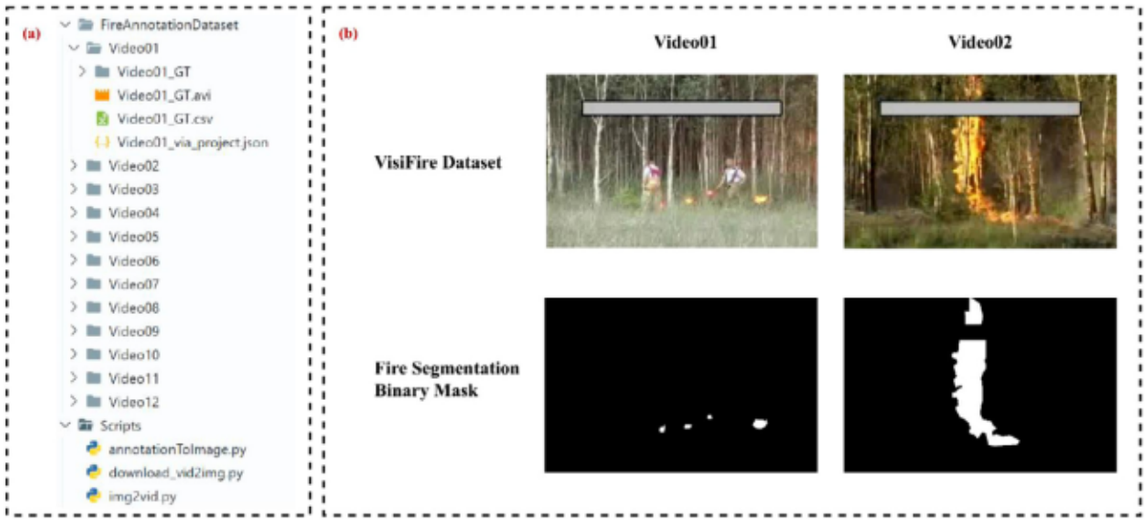

Fig. 1. Files within Region-based Annotation Data of Fire Images. (a) dataset structure; (b) Sample of Segmentation binary masks Ground Truth on Video01, and Video02. We highly recommend visiting dataset overview at reference [4], for better video viewing experience.

1. Data Description

This annotation dataset [4] provides fire segmentation data derived from 12 commonly used video for fire detection tasks. The fire videos were based on publicly available, online VisiFire Dataset [2], which used on numerous research such as on [5,6] to name a few. For each fire video, this dataset provides per-frame segmentation data with three data formats (annotation, image, and video).

For better reproducibility and convenience, we also included in our annotation dataset [4] additional scripts (/Scripts/download_vid2img.py, /Scripts/annotationToImage.py, /Scripts/img2vid.py), where these scripts correspond to the I, III, IV of our video annotation process (See Fig. 2 for details). We hope this script will be helpful, and able to assist other for future video annotation process.

The main dataset structure is shown in Fig. 1 (a). For each video folder "VideoN" (Video01, Video02, …, Video12) in the FireAnnotationDataset, it includes annotation from VIA (VGG Image Annotator) project file named "VideoN_via_project.json", and "VideoN_GT.csv" file. Then for images, "VideoN_GT" folder stores all video frame binary masks, the ground truth or GT for short ("VideoN_GT_Frame_001.jpg", …, "VideoN_GT_Frame_lastframe.jpg"). Lastly, each folder also provides previous image binary masks in video form named "VideoN_GT.avi".

In details, for annotation format, the data we provide are CSV and VIA project files. We choose to add the project files because of the export flexibility it provides. It can export into any common annotation format such as csv (the current data type we used), json, and COCO (Common Objects in Context), which commonly used on deep learning model such as Faster R-CNN [7]. For image format, we use previous csv annotation file to draw the segmentation shown in Fig. 1(b) binary masks. For video format, segmentation videos were acquired by converting previous binary masks using the same FPS as the original video conversion. Table 1 provides these fire video details and specification that were used to annotate the segmentation data.

2. Experimental Design, Materials and Methods

The first process (I) of annotating the fi[10]ideo, shown in Fig. 2, was to convert the downloaded original fire video into images. We use Python OpenCV (Open-Source Computer Vision) library VideoCapture method. In specific, all videos FPS we sample were native, apart from

Table 1
Fire Video Segmentation VisiFire origin and specification used on the annotation process.

Video Data	VisiFire Dataset [2]	Resolution	FPS	Total Annotated Frame	Video Content
Video01	controlled1.avi	400 × 256	15	260	A controlled forest fire.
Video02	controlled2.avi	400 × 256	15	246	A controlled forest fire.
Video03	controlled3.avi	400 × 256	15	208	A controlled forest fire.
Video04	forest1.avi	400 × 256	15	200	A controlled forest fire.
Video05	forest2.avi	400 × 256	15	245	A controlled forest fire.
Video06	forest3.avi	400 × 256	15	255	A controlled forest fire.
Video07	forest4.avi	400 × 256	15	219	A controlled forest fire.
Video08	forest5.avi	400 × 256	15	216	A controlled forest fire.
Video09	fBackYardFire.avi	320 × 240	2	241	A fire generated with a red ground.
Video10	ForestFire1.avi	400 × 256	15	218	A controlled forest fire.
Video11	fire1.avi	320 × 240	5	236	A fire in a pot and a person walking by.
Video12	40m_PanFire_20,060,824.avi	320 × 240	30	140	A fire in a bucket very far away from the camera.

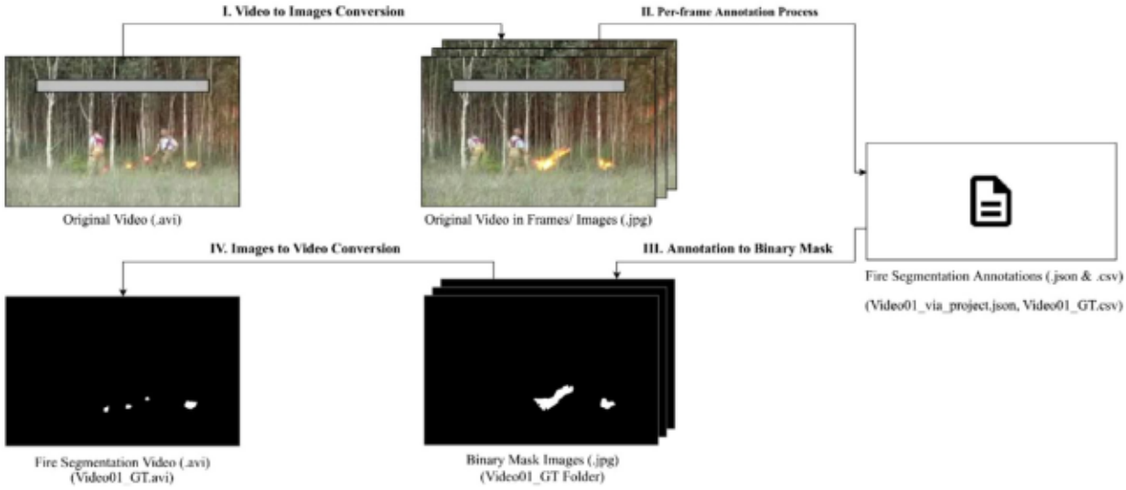

Fig. 2. Illustration of the fire images annotation process.

Video09 to Video11, where we need to truncate it down for more balanced dataset, and due to resource limitation of per-frame annotation process. With this script, we provide parameters of what video to convert and how many FPS to sample, then the script will output video frames.

The second (II) and most extensive process was three months of manual per-frame annotation. For this we use VIA (VGG Image Annotator) Tools. For each frame in the fire video, we annotate fire area with polygon region tool. We assigned each video with its own VIA project. Result of this process was a VIA project json file, in which it contains detailed information and annotation. In this dataset, we also choose to export the annotation as csv file (comma separated value), however VIA is not limited to, and can, export to other format such as json or json COCO format.

The third process (III), from previous csv file, we then use Python script [4] to read the annotation data and fed this data to OpenCV fillPoly method to create our binary masks. This script also needs manual parameter of original video resolution shown in Table 1. Result of this third process was fire segmentation binary masks, an important ground truth data in machine learning semantic segmentation tasks.

The fourth final process (IV) then was to convert the binary masks back into videos. In here, we implement OpenCV VideoWriter, and using the same FPS on previous original video shown

in Table 1. We input the binary mask images folder address to the script, and it will output the final fire segmentation video.

Ethics Statements

The data included is anonymized or includes indefinable information e.g. faces. The primary dataset were collected by Prof. A. Enis Cetin and are available publicly online as Sample Video Clips on http://signal.ee.bilkent.edu.tr/VisiFire/ under public domain license.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which could influence the work reported in this article

CRediT Author Statement

Wahyono: Conceptualization, Formal analysis, Methodology, Writing – review & editing, Supervision, Validation; **Andi Dharmawan:** Writing – review & editing; **Agus Harjoko:** Writing – review & editing; **Chrystian:** Data curation, Investigation, Writing – original draft; **Faisal Dharma Adhinata:** Investigation, Validation.

Acknowledgments

This work was supported by the Ministry of Education Culture, Research and Technology, Republic of Indonesia under World Class Research (WCR) Grant (Grant Number 111/E4.1/AK.04.PT/2021 and 4506/UN1/DITLIT/DIT-LIT/PT/2021).

References

[1] A. Dutta, A. Zisserman, The VIA annotation software for images, audio and video, in: Proceedings of the 27th ACM International Conference on Multimedia MM '19, 2019, doi:10.1145/3343031.3350535.
[2] A. Enis Cetin, Computer vision based fire detection dataset., (2014). http://signal.ee.bilkent.edu.tr/VisiFire/. Accessed May 15, 2021.
[3] J. Mlích, K. Koplík, M. Hradiš, P. Zemčík, Fire segmentation in still images, Lecture Notes in Computer Science, Springer, 2020 (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), doi:10.1007/978-3-030-40605-9_3.
[4] Wahyono, A. Dharmawan, A. Harjoko, Chrystian, F.D. Adhinata, Region-Based Annotation Data of Fire Images for Intelligent Surveillance System, Zenodo, 2022, doi:10.5281/zenodo.5893854.
[5] B.U. Töreyin, Y. Dedeoğlu, U. Güdükbay, A.E. Çetin, Computer vision based method for real-time fire and flame detection, Pattern Recognit. Lett. (2006) 27, doi:10.1016/j.patrec.2005.06.015.
[6] M. Torabian, H. Pourghassem, H. Mahdavi-Nasab, Fire detection based on fractal analysis and spatio-temporal features, Fire Technol. 57 (2021), doi:10.1007/s10694-021-01129-7.
[7] S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN: towards real-time object detection with region proposal networks, Adv. Neural Inf. Process. Syst. (2015).

Region-based annotation data of fire images for intelligent surveillance system

ORIGINALITY REPORT

12%
SIMILARITY INDEX

PRIMARY SOURCES

1	www.biorxiv.org Internet	39 words — 2%
2	portalrevistas.uct.cl Internet	34 words — 2%
3	Agus Eko Minarno, Indah Soesanti, Hanung Adi Nugroho. "Batik Nitik 960 Dataset for Classification, Retrieval, and Generator", Data, 2023 Crossref	26 words — 2%
4	www.researchgate.net Internet	24 words — 1%
5	Xubing Yang, Zhichun Hua, Li Zhang, Xijian Fan, Fuquan Zhang, Qiaolin Ye, Liyong Fu. "Preferred vector machine for forest fire detection", Pattern Recognition, 2023 Crossref	18 words — 1%
6	Weixing Zhang, Anna K. Liljedahl, Mikhail Kanevskiy, Howard E. Epstein, Benjamin M. Jones, M. Torre Jorgenson, Kelcy Kent. "Transferability of the Deep Learning Mask R-CNN Model for Automated Mapping of Ice-Wedge Polygons in High-Resolution Satellite and UAV Images", Remote Sensing, 2020 Crossref	14 words — 1%
7	vbn.aau.dk Internet	13 words — 1%
8	mafiadoc.com Internet	10 words — 1%
9	etheses.whiterose.ac.uk Internet	8 words — < 1%
10	www.codeproject.com	

Internet

8 words — < 1%

11 www.akademiabaru.com
Internet

7 words — < 1%

EXCLUDE QUOTES	OFF	EXCLUDE SOURCES	OFF
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	OFF