Early wildfire smoke detection in videos

Taanya Gupta, Hengyue Liu ,Bir Bhanu
Department of Electrical and Computer Engineering, University of California, Riverside, CA 92521, USA

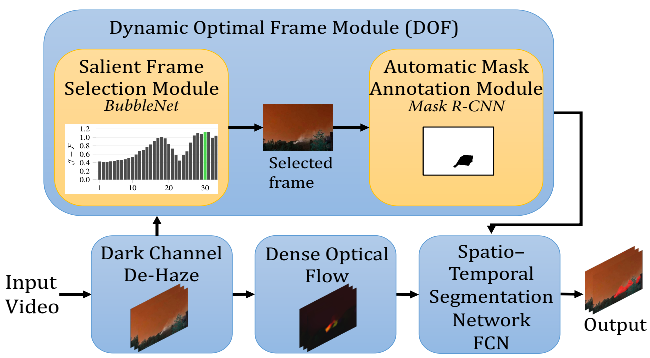
Introduction

- A novel approach that detects wildfire smoke at the pixel level from videos by integrating spatial and temporal features into a semi-supervised deep learning-based video object detection technique.
- A means of mitigating the paucity of data by adopting an online training method that focuses on a specific smoke video and transfers generic features to specific ones.
- An accounting of the physical parameter of haze, which impairs the visibility of smoke in the sky during a wildfire.

Related Work

- In addition to using classical convolutional neural networks, Hu et al. also considered the temporal aspect.
- A faster region-based convolutional neural network was used by Kim et al. to find possible wildfire regions.

Architecture



- Dark Channel De-Haze Pre-Processor: To get rid of the haze from the video frames, state-of-the-art technique is adopted: single image haze removal using dark channel prior.
- Dynamic Optimal Frame (DOF) Module: Provides two annotated frames that help the network fine-tune the already trained model for the specific video using the salient frame selection module and the automatic mask annotation tool.
- Dense Optical Flow Module: Optical flow provides the gradient of the vertical and horizontal axis to compute the temporal features.
- Spatio-Temporal Segmentation Network: A fully convolutional network that takes in a total of five inputs and outputs smoke masks generated for the whole video.

Algorithm 1 Wildfire smoke detection in video

$WildfireSmokeDetectionModel(video):$

$V_d \leftarrow \phi_D(video)$

where, $V_d =$ Dehazed video

$I_t \leftarrow \phi_S(V_d)$

where, $I_t =$ Optimal frame image

$I_{t-1_{mask}}, I_{t_{mask}} \leftarrow \phi_A(I_{t-1}, I_t)$

where, $I_{t_{mask}} =$ Mask of frame I_t

$IF_t \leftarrow \phi_{DO}(I_{t-1}, I_t)$

where, $IF_t =$ Flow vectors of I_{t-1} and I_t

return $\phi_{ST}(I_{t-1}, I_t, I_{t-1_{mask}}, I_{t_{mask}}, IF_t)$

Algorithm 1 - Takes a video as an input and outputs the smoke segmentation in the video frames. Each frame is de-hazed then ranked by salient frame selection module, and the mask of the selected frame is generated by the automatic mask annotation tool. The mask is used for online training the segmentation network. Finally, the spatio-temporal segmentation network returns the segmentation of the smoke.

Quantitative Results

Video	True Positives	False Positives	IOU
Video 1	118±0.128	4±0.257	0.974±0.015
Video 2	100±0	0±0	0.991±0.004
Video 3	142±0.302	21±0.591	0.942±0.023
Video 4	131±0.496	25±0.777	0.867±0.025
Video 5	114±0.093	2±0.186	0.990±0.004

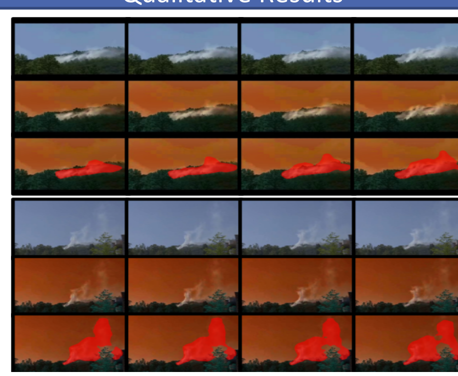
Video	Pixel accuracy	Precision	Recall	F1 score
Video 1	0.973±0.014	0.872±0.045	0.831±0.044	0.850±0.036
Video 2	0.991±0.004	0.913±0.034	0.908±0.040	0.910±0.032
Video 3	0.945±0.025	0.743±0.051	0.765±0.045	0.753±0.039
Video 4	0.870±0.023	0.698±0.034	0.706±0.046	0.701±0.024
Video 5	0.990±0.004	0.869±0.043	0.949±0.018	0.907±0.027

Quantitative metrics (mean standard deviation) defined for videos from the smoke video test dataset

Model	Mean IOU	Mean F1 score
Our Model	0.95	0.82
OSVOS	0.88	0.74
Unet	0.55	0.73
Fpn	0.79	0.65
LinkNet	0.39	0.55

Comparisons of results over different segmentation frameworks.

Qualitative Results



Visualizations of de-hazed and segmentation results of consecutive frames from 3 videos respectively. The first row shows the original frames; the second row is the de-hazed frame after pre-processing; the third row is the segmentation result (smoke regions are marked in red) for each video.

Conclusion

- Detects wildfire at an early stage using a novel semi-supervised spatio- temporal approach.
- Takes into account the physical parameter of haze surrounding the smoke.
- The proposed work detects smoke pixels of multiple sizes and mitigates the problem of annotated data by using online training.