

# Supplementary Material

## BSUV-Net: A Fully-Convolutional Neural Network for Background Subtraction of Unseen Videos

M. Ozan Tezcan, Janusz Konrad, Prakash Ishwar  
Department of Electrical and Computer Engineering  
Boston University  
Boston, MA  
[mtezcan, jkonrad, pi]@bu.edu

### Selection of training and test sets for evaluating performance on unseen videos

In this paper, we introduced a supervised background subtraction (BGS) algorithm for unseen videos. As for all supervised learning algorithms, the size and diversity of the training data are crucially important for the learning process. Generally speaking, for most of the state-of-the-art deep neural networks, the best approach is to use all of the available training data. Unfortunately, CDNet 2014 [1] does not provide different videos for training and testing. Instead, it provides some frames from each video as training data and the remaining ones – as test data. However, this type of division is not useful for evaluating the performance on *unseen* videos.

For comparing the performance of different models on *unseen* videos, we split the dataset into 18 different sets of training/testing videos as shown in Tables 1 and 2. When training a supervised algorithm, the main assumption is that the training set is diverse enough to cover a wide range of test scenarios. For example, if there are no examples that include shadow in the training set, then it is impossible for the network to learn how to classify shadow regions. Therefore, we designed the splits so that the training set for each split contains some videos from the same category as the test videos. We did not perform a full “leave- $k$ -videos-out” cross-validation due to prohibitive time need to train BSUV-Net. In all of the tests, we used videos from “baseline”, “bad weather”, “intermittent object motion”, “low frame rate” and “shadow” categories during training since they span most of the common scenarios. For videos from more difficult scenarios, we progressively added additional categories into the training set. In particular, we considered “PTZ”, “thermal” and “turbulence” categories as the most difficult ones since they have substantially different data characteristics from other categories. “PTZ” is the only category with significant

camera movement and zoom in/out, while “thermal” and “turbulence” categories capture different scene properties than the remaining categories (far- and near-infrared spectrum instead of RGB, respectively). For these 3 categories, we used more videos in the training set, than in the other categories. Please note that a “leave- $k$ -videos-out” approach would have more videos in the training set compared to our splits and is therefore likely to yield better results.

### References

- [1] N. Goyette, P.-M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar. A novel video dataset for change detection benchmarking. *IEEE Trans. Image Process.*, 23(11):4663–4679, 2014.



Table 2. Training and test splits  $S_{13}$  to  $S_{18}$  used for evaluation.

category	video	$S_{13}$	$S_{14}$	$S_{15}$	$S_{16}$	$S_{17}$	$S_{18}$
baseline	highway	Tr	Tr	Tr	Tr	Tr	Tr
	pedestrians	Tr	Tr	Tr	Tr	Tr	Tr
	office	Tr	Tr	Tr	Tr	Tr	Tr
	PETS2006	Tr	Tr	Tr	Tr	Tr	Tr
bad weather	skating	Tr	Tr	Tr	Tr	Tr	Tr
	blizzard	Tr	Tr	Tr	Tr	Tr	Tr
	snowFall	Tr	Tr	Tr	Tr	Tr	Tr
	wetSnow	Tr	Tr	Tr	Tr	Tr	Tr
intermittent object motion	abandonedBox	Tr	Tr	Tr	Tr	Tr	Tr
	parking	Tr	Tr	Tr	Tr	Tr	Tr
	sofa	Tr	Tr	Tr	Tr	Tr	Tr
	streetLight	Tr	Tr	Tr	Tr	Tr	Tr
	tramstop	Tr	Tr	Tr	Tr	Tr	Tr
	winterDriveway	Tr	Tr	Tr	Tr	Tr	Tr
low framerate	port 0.17fps	Tr	Tr	Tr	Tr	Tr	Tr
	tramCrossroad 1fps	Tr	Tr	Tr	Tr	Tr	Tr
	tunnelExit 0.35fps	Tr	Tr	Tr	Tr	Tr	Tr
	turnpike 0.5fps	Tr	Tr	Tr	Tr	Tr	Tr
shadow	backdoor	Tr	Tr	Tr	Tr	Tr	Tr
	bungalows	Tr	Tr	Tr	Tr	Tr	Tr
	busStation	Tr	Tr	Tr	Tr	Tr	Tr
	copyMachine	Tr	Tr	Tr	Tr	Tr	Tr
	cubicle	Tr	Tr	Tr	Tr	Tr	Tr
	peopleInShade	Tr	Tr	Tr	Tr	Tr	Tr
camera jitter	badminton	Tr	Tr	Tr	Tr	Tr	Tr
	traffic	Tr	Tr	Tr	Tr	Tr	Tr
	boulevard	Tr	Tr	Tr	Tr	Tr	Tr
	sidewalk	Tr	Tr	Tr	Tr	Tr	Tr
dynamic background	boats	Tr	Tr	Tr	Tr	Tr	Tr
	canoe	Tr	Tr	Tr	Tr	Tr	Tr
	fall	Tr	Tr	Tr	Tr	Tr	Tr
	fountain01	Tr	Tr	Tr	Tr	Tr	Tr
	fountain02	Tr	Tr	Tr	Tr	Tr	Tr
	overpass	Tr	Tr	Tr	Tr	Tr	Tr
night videos	bridgeEntry						
	busyBoulevard						
	fluidHighway						
	streetCornerAtNight						
	tramStation						
	winterStreet						
PTZ	continuousPan	Tr	Test				
	intermittentPan	Tr	Test				
	twoPositionPTZCam	Test	Tr				
	zoomInZoomOut	Test	Tr				
thermal	corridor			Tr	Test	Tr	Tr
	diningRoom			Tr	Test	Tr	Tr
	lakeSide			Test	Tr	Tr	Tr
	library			Test	Tr	Tr	Tr
	park			Test	Tr	Tr	Tr
turbulence	turbulence0			Tr	Tr	Test	Tr
	turbulence1			Tr	Tr	Test	Tr
	turbulence2			Tr	Tr	Tr	Test
	turbulence3			Tr	Tr	Tr	Test