Is it Raining Outside?
Detection of Rainfall using General-Purpose Surveillance Cameras

Joakim Bruslund Haurum, Chris H. Bahnsen, Thomas B. Moeslund
Visual Analysis of People Laboratory
Aalborg University, Denmark
{joha, cb, tbm}@create.aau.dk

2019 IEEE/CVF Conference Computer Vision and Pattern Recognition (CVPR) Workshop

Tian Liu

Sep 21, 2020

Outline

- Background and objectives
- Methodology
  - Bossu Rain Detection Algorithm (Statistical method)
  - 3D CNN (Deep Learning based)
- Experiments and Results
  - AAU Visual Rain Dataset
- Conclusions

Background and Objectives

- Rain detection is critical for rain removal algorithm
- Build a system to automatically detect rain using surveillance cameras (traffic or home cameras)
- Compared the two methods based on AAU Visual Rain Dataset (215 hrs surveillance video)
  - Statistical method (Bossu’s algorithm)
  - 3D CNN

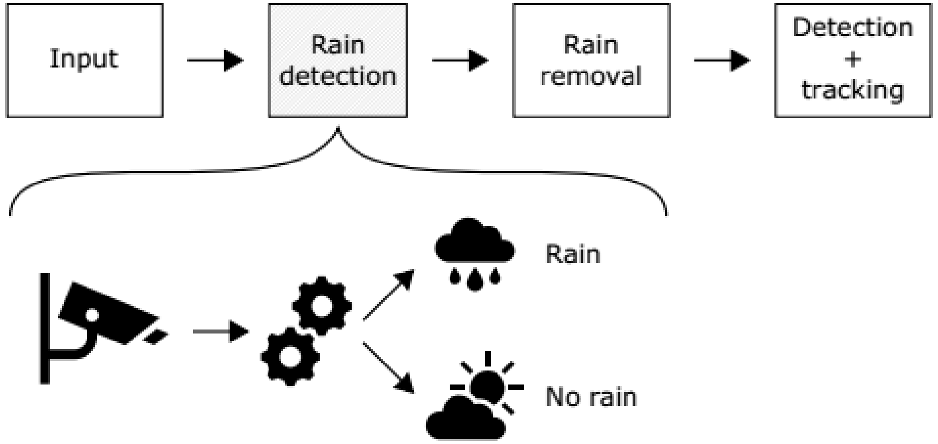
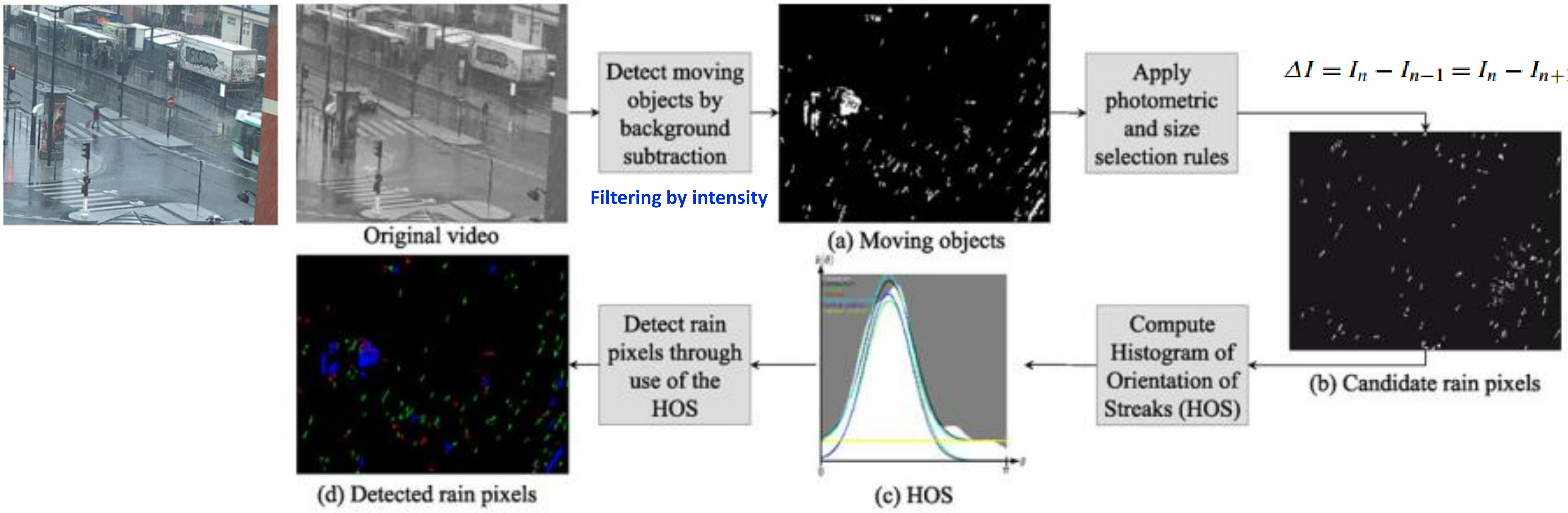


Figure 1. The proposed system at-a-glace. For rain removal algorithms to be effective in an integrated surveillance framework, the presence or absence of rain must be detected in a pre-processing step.

Bossu Rain Detection Algorithm (2011)



$$\Delta I = I_n - I_{n-1} = I_n - I_{n+1} \geq c$$

Filtering by intensity

Rain: green
Noise: red
Pixel filtered by size: blue

- Fit the histogram with Gaussian using Expectation Maximum algorithm
- Detect rain based on the surface of Gaussian Distribution

Bossu Rain Detection Workflow

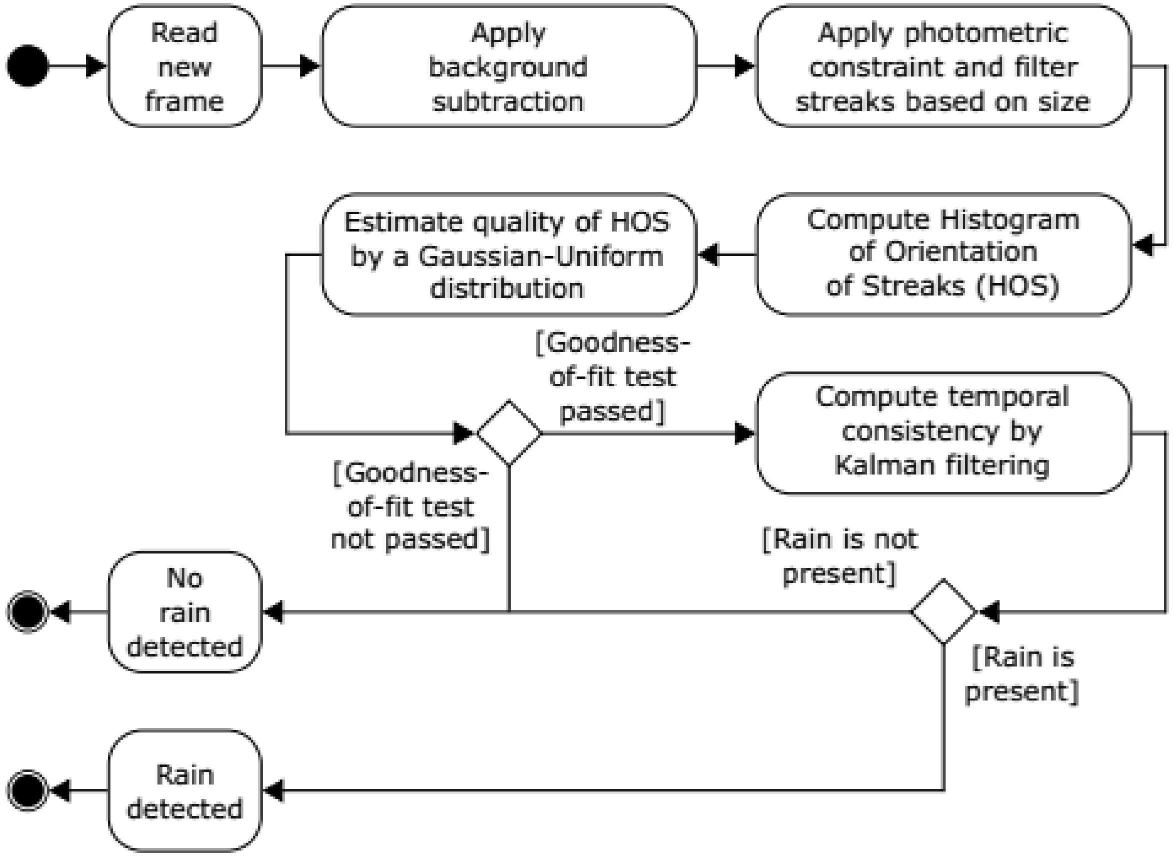


Figure 3. Activity diagram of the rain detection algorithm by Bossu *et al.* [6].

C3D CNN

- FCN modified from C3D CNN by Tran (2015) to avoid cropping or resizing input images
- Network trained as a binary classification problem

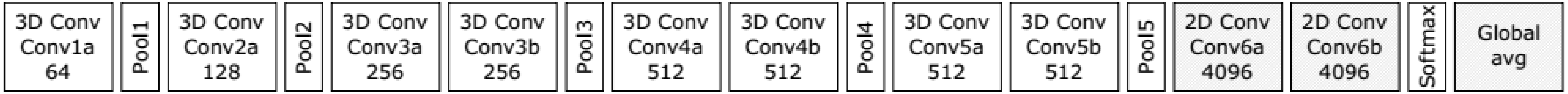


Figure 4. Overview of the modified C3D CNN architecture. 3D conv and 2D conv denotes 3D and conventional 2D convolutions, respectively. Pool denotes max pooling and the number of filters are denoted in the bottom of each box. Our modifications of the network are marked in grey. Figure adapted from [33].

AAU Visual Rain Dataset

Dataset	Duration (hh:mm)	Frames	Frame rate	Native resolution	Cropped resolution	Camera model	Distance to gauge Mech.	Laser
Crossing1-trn	87:38	9,276,654	30	1024 × 640	700 × 612	AXIS Q1615-E	580 m	1230 m
Crossing1-val	20:37	2,184,499	30	1024 × 640	700 × 612	AXIS Q1615-E	580 m	1230 m
Crossing2-tst	106:59	9,463,287	25	640 × 480	276 × 338, 276 × 112	AXIS M114-E	1820 m	970 m

Table 1. Overview of the AAU VIRADA dataset. Mech. denotes the mechanical tipping-scale rain gauge whereas Laser denotes the laser disdrometer. The two noted cropped resolutions for the Crossing2-tst dataset are for the asphalt and brick crops, respectively.

	Detected rain %	
Measurement device	Laser	Mech.
Crossing1-trn	19.67	17.97
Crossing1-val	20.86	14.63
Crossing2-tst	8.68	1.65

Table 2. Overview of the ratio of detected rain for the AAU VIRADA dataset, per measurement device. Mech. denotes the mechanical tipping-scale rain gauge whereas Laser denotes the laser disdrometer.

(a) Crossing1

(b) Crossing2

Figure 2. Sample views of the traffic crossings from the AAU VIRADA dataset. Discarded regions are shown with a semi-transparent overlay. We denote the upper region of Crossing2 as Crossing2-brick whereas the lower region is denoted as Crossing2-asphalt.

Bossu's Algorithm Implementation

- Tuned the hyperparameters based on grid search
- Two evaluation methods:
  - Bossu-EM: per-frame EM estimated HOS parameters
  - Bossu-Kalman: Kalman smoothed HOS parameters

Parameter	Search space	Selected value
MoG warm-up frames	[500]	500
c	[3, 5]	3
Minimum BLOB size	[4]	4
Maximum BLOB size	[50:50:200]	200
dm	[0.5:0.5:2.0]	0.50
EM max iterations	100	100
D_c	[0.01:0.01:0.20]	0.19
Π_{rain}	[0.20:0.02:0.50]	0.40

Table 3. Values and search space for the Bossu parameter search. The ranges in the search space follow the python convention, with [3,5] being a list of parameters and [0.5:0.5:2.0] referring to values in the range from 0.5 to 2.0 with an interval of 0.5

- For rain sequences, the Bossu algorithm should detect rain for at least 60 % of the frames, preferably more.
- For no-rain sequences, the Bossu algorithm should detect rain for maximum 40 % of the frames, preferably less.

C3D Training Results

- C3D-FCN: evaluated with entire frame
- C3D-Center: evaluated with 112x112 center patch

(a) Crossing1

Hyperparameter	Value
Batch size	128
Sequence stride	8
Learning rate	0.01
Momentum	0.9
Weight decay	0.0001
γ	0.1
s	5

Table 4. C3D hyperparameters.

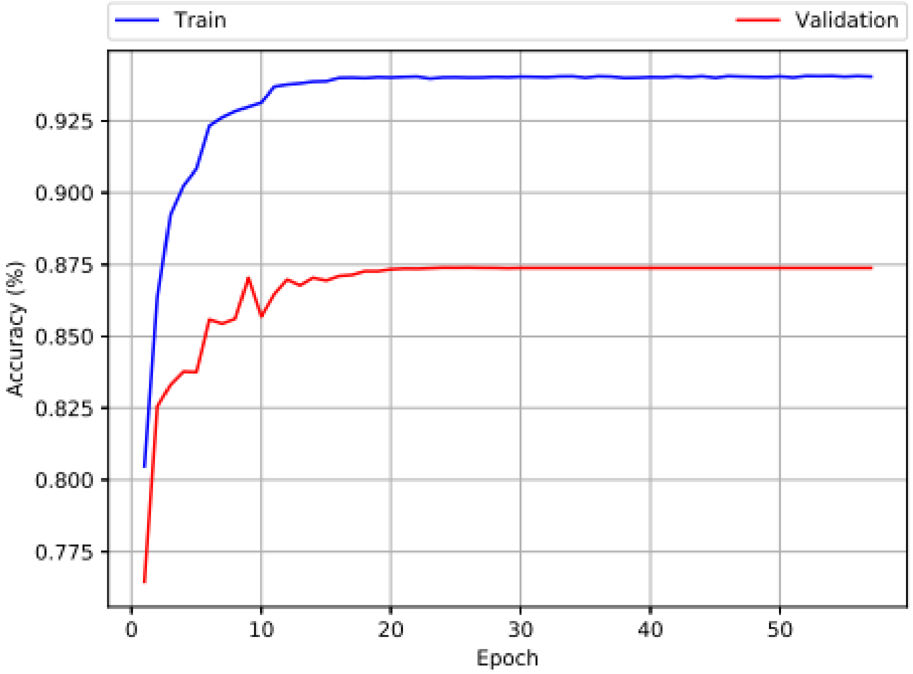


Figure 5. Average accuracy per epoch for the trained C3D CNN.

Results

Sequence	Method	TP	TN	FP	FN	Acc	F1	MCC
Crossing1-trn	C3D-FCN	215244	923596	7766	12802	**0.9823**	**0.9544**	**0.9435**
	C3D-Center	202932	920692	10672	25114	0.9691	0.9190	0.9007
	Bossu-EM	1096369	3498967	3915541	719777	0.4978	0.3211	0.0603
	Bossu-Kalman	1119361	3460093	3954415	696785	0.4961	0.3249	0.0663
Crossing1-val	C3D-FCN	30126	208447	7405	27042	**0.8738**	**0.6362**	**0.5821**
	C3D-Center	25320	199555	16298	31847	0.8237	0.5126	0.4159
	Bossu-EM	263008	912983	805113	192395	0.5411	0.3453	0.0887
	Bossu-Kalman	267253	909237	808859	188150	0.5413	0.3490	0.0945
Crossing2-asphalt	C3D-FCN	0	1069231	0	102717	**0.9124**	0.0000	0.0000
	C3D-Center	245	1039578	40409	102474	0.8792	0.0034	-0.0837
	Bossu-EM	224853	6335804	2257136	591994	0.6972	0.1363	**0.0080**
	Bossu-Kalman	234181	6264711	2328229	582666	0.6907	**0.1386**	0.0010
Crossing2-brick	C3D-FCN	72619	729561	350381	30095	**0.6783**	0.2763	0.2248
	C3D-Center	75690	720369	359557	27024	0.6731	**0.2814**	**0.2359**
	Bossu-EM	281084	5837499	2755441	535763	0.6502	0.1459	0.0141
	Bossu-Kalman	290583	5762519	2830421	526264	0.6433	0.1476	0.0158

Table 5. Rain detection results on the AAU VIRADA dataset, using labels from the laser disdrometer.

$$\mathrm{Acc} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}} \tag{9}$$

$$\mathrm{F1} = \frac{2\,\mathrm{TP}}{2\,\mathrm{TP}+\mathrm{FP}+\mathrm{FN}} \tag{10}$$

$$\mathrm{MCC} = \frac{\mathrm{TP}\cdot\mathrm{TN}-\mathrm{FP}\cdot\mathrm{FN}}{\sqrt{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}} \tag{11}$$

MCC = -1: total disagreement
MCC = 0 : pure guesswork
MCC = 1 : perfect predications

CNN Outperforms Bossu (Perfect Acc vs. Random guess)

Sequence	Method	TP	TN	FP	FN	Acc	F1	MCC
Crossing1-trn	C3D-FCN	215244	923596	7766	12802	**0.9823**	**0.9544**	**0.9435**
	C3D-Center	202932	920692	10672	25114	0.9691	0.9190	0.9007
	Bossu-EM	1096369	3498967	3915541	719777	0.4978	0.3211	0.0603
	Bossu-Kalman	1119361	3460093	3954415	696785	0.4961	0.3249	0.0663
Crossing1-val	C3D-FCN	30126	208447	7405	27042	**0.8738**	**0.6362**	**0.5821**
	C3D-Center	25320	199555	16298	31847	0.8237	0.5126	0.4159
	Bossu-EM	263008	912983	805113	192395	0.5411	0.3453	0.0887
	Bossu-Kalman	267253	909237	808859	188150	0.5413	0.3490	0.0945
Crossing2-asphalt	C3D-FCN	0	1069231	0	102717	**0.9124**	0.0000	0.0000
	C3D-Center	245	1039578	40409	102474	0.8792	0.0034	-0.0837
	Bossu-EM	224853	6335804	2257136	591994	0.6972	0.1363	**0.0080**
	Bossu-Kalman	234181	6264711	2328229	582666	0.6907	**0.1386**	0.0010
Crossing2-brick	C3D-FCN	72619	729561	350381	30095	**0.6783**	0.2763	0.2248
	C3D-Center	75690	720369	359557	27024	0.6731	**0.2814**	**0.2359**
	Bossu-EM	281084	5837499	2755441	535763	0.6502	0.1459	0.0141
	Bossu-Kalman	290583	5762519	2830421	526264	0.6433	0.1476	0.0158

Table 5. Rain detection results on the AAU VIRADA dataset, using labels from the laser disdrometer.

Performance can be improved

Entire Frame Outperforms Center Patch

Sequence	Method	TP	TN	FP	FN	Acc	F1	MCC
Crossing1-trn	C3D-FCN	215244	923596	7766	12802	**0.9823**	**0.9544**	**0.9435**
	C3D-Center	202932	920692	10672	25114	0.9691	0.9190	0.9007
	Bossu-EM	1096369	3498967	3915541	719777	0.4978	0.3211	0.0603
	Bossu-Kalman	1119361	3460093	3954415	696785	0.4961	0.3249	0.0663
Crossing1-val	C3D-FCN	30126	208447	7405	27042	**0.8738**	**0.6362**	**0.5821**
	C3D-Center	25320	199555	16298	31847	0.8237	0.5126	0.4159
	Bossu-EM	263008	912983	805113	192395	0.5411	0.3453	0.0887
	Bossu-Kalman	267253	909237	808859	188150	0.5413	0.3490	0.0945
Crossing2-asphalt	C3D-FCN	0	1069231	0	102717	**0.9124**	0.0000	0.0000
	C3D-Center	245	1039578	40409	102474	0.8792	0.0034	-0.0837
	Bossu-EM	224853	6335804	2257136	591994	0.6972	0.1363	**0.0080**
	Bossu-Kalman	234181	6264711	2328229	582666	0.6907	**0.1386**	0.0010
Crossing2-brick	C3D-FCN	72619	729561	350381	30095	**0.6783**	0.2763	0.2248
	C3D-Center	75690	720369	359557	27024	0.6731	**0.2814**	**0.2359**
	Bossu-EM	281084	5837499	2755441	535763	0.6502	0.1459	0.0141
	Bossu-Kalman	290583	5762519	2830421	526264	0.6433	0.1476	0.0158

Table 5. Rain detection results on the AAU VIRADA dataset, using labels from the laser disdrometer.

C3D Generalizes with Similar Training Surfaces

Sequence	Method	TP	TN	FP	FN	Acc	F1	MCC
Crossing1-trn	C3D-FCN	[illegible]	923596	7766	12802	**0.9823**	**0.9544**	**0.9435**
					25114	0.9691	0.9190	0.9007
					719777	0.4978	0.3211	0.0603
					696785	0.4961	0.3249	0.0663
Crossing1-val					27042	**0.8738**	**0.6362**	**0.5821**
					31847	0.8237	0.5126	0.4159
					192395	0.5411	0.3453	0.0887
			[illegible]	[illegible]	188150	0.5413	0.3490	0.0945
Crossing2-asphalt	C3D-FCN	0	1069231	0	102717	**0.9124**	0.0000	0.0000
	C3D-Center	245	1039578	40409	102474	0.8792	0.0034	-0.0837
	Bossu-EM	224853	6335804	2257136	591994	0.6972	0.1363	**0.0080**
	Bossu-Kalman	234181	6264711	2328229	582666	0.6907	**0.1386**	0.0010
Crossing2-brick	C3D-FCN	72619	729561	350381	30095	**0.6783**	0.2763	0.2248
	C3D-Center	75690	720369	359557	27024	0.6731	**0.2814**	**0.2359**
	Bossu-EM	281084	5837499	2755441	535763	0.6502	0.1459	0.0141
	Bossu-Kalman	290583	5762519	2830421	526264	0.6433	0.1476	0.0158

Table 5. Rain detection results on the AAU VIRADA dataset, using labels from the laser disdrometer.

(a) Crossing1

(b) Crossing2

Dynamic Effects Filtered by Center Patch

Sequence	Method	TP	TN	FP	FN	Acc	F1	MCC
Crossing1-trn	C3D-FCN	[illegible]	923596	7766	12802	**0.9823**	**0.9544**	**0.9435**
					25114	0.9691	0.9190	0.9007
					719777	0.4978	0.3211	0.0603
					696785	0.4961	0.3249	0.0663
Crossing1-val					27042	**0.8738**	**0.6362**	**0.5821**
					31847	0.8237	0.5126	0.4159
					192395	0.5411	0.3453	0.0887
			[illegible]	[illegible]	188150	0.5413	0.3490	0.0945
Crossing2-asphalt	C3D-FCN	0	1069231	0	102717	**0.9124**	0.0000	0.0000
	C3D-Center	245	1039578	40409	102474	0.8792	0.0034	-0.0837
	Bossu-EM	224853	6335804	2257136	591994	0.6972	0.1363	**0.0080**
	Bossu-Kalman	234181	6264711	2328229	582666	0.6907	**0.1386**	0.0010
Crossing2-brick	C3D-FCN	72619	729561	350381	30095	**0.6783**	0.2763	0.2248
	C3D-Center	75690	720369	359557	27024	0.6731	**0.2814**	**0.2359**
	Bossu-EM	281084	5837499	2755441	535763	0.6502	0.1459	0.0141
	Bossu-Kalman	290583	5762519	2830421	526264	0.6433	0.1476	0.0158

Table 5. Rain detection results on the AAU VIRADA dataset, using labels from the laser disdrometer.

(a) Crossing1

(b) Crossing2

Conclusions

- Provided by far the biggest rain dataset captured by general purpose surveillance cameras
- C3D outperforms Bossu's method
- Performance depends on the similarity of investigated region-of-interest and the training data
- Task of rain detection for general-purpose surveillance cameras not solved yet

Future Work / Ideas

- Optimize the C3D CNN with hyperparameter optimization for higher accuracy
- CNN to infer rain amount
- In realistic application of CNN, where to find training data set?
- Try other more updated statistical methods to infer rain amount

Back Up