

000 001 002 003 004 Rebuttal Supplementary Evaluations and Results 005 006 007 008 009 010 011 012

013 Anonymous CVPR submission
014
015
016
017 Dear AC and reviewers, We show additional evaluations and results in this document. Sec. 1 shows the rain removal
018 output frames' PSNR/SSIM comparison with the two additional video based rain removal methods, i.e., *VST-ICCV17'* [4]
019 and *TCL-TIP15'* [2]. Sec. 2 shows additional visual comparisons.
020
021
022

023 1. Derain Frame PSNR/SSIM Comparison 024

025 Table. 1 shows additional PSNR/SSIM evaluations for two more video based derain methods, i.e., *VST-ICCV17'* [4] and
026 *TCL-TIP15'* [2]. The best performance for each video sequence has been highlighted in red, and the second best in blue. It
027 can be seen from the table, the proposed *SPAC-CNN* algorithm's advantage holds among all additional competing methods.
028 We also copy the results of *DDN-CVPR17'* [1] and *VMD-CVPR17'* from the manuscript for your reference.
029
030
031
032

033 Table 1: Rain removal performance comparison between different methods in terms of scene reconstruction PSNR/SSIM.
034
035
036
037

038 Camera Motion	Clip No.	Rain		DDN-CVPR17 [1]		VMD-CVPR17 [3]		VST-ICCV17' [4]		TCL-TIP15' [2]		SPAC-CNN	
		-		Image-Based		Video-Based		Video-Based		Video-Based		Video-Based	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
039 panning unstable camera	a1	28.46	0.94	28.02	0.95	26.96	0.92	26.14	0.94	29.87	0.96	29.78	0.97
	a2	28.09	0.95	27.38	0.95	24.80	0.93	24.03	0.83	29.01	0.96	30.09	0.96
	a3	27.84	0.93	27.41	0.94	26.45	0.90	20.50	0.70	28.82	0.95	29.75	0.96
	a4	31.48	0.95	32.47	0.97	29.55	0.94	33.41	0.96	34.12	0.98	34.82	0.98
	avg. a	28.97	0.94	28.82	0.95	26.94	0.92	26.02	0.86	30.46	0.96	31.11	0.97
040 speed 20-30 km/h	b1	28.72	0.92	29.48	0.96	24.09	0.84	22.25	0.76	28.07	0.94	31.19	0.96
	b2	29.49	0.90	30.23	0.95	25.81	0.89	25.13	0.79	32.41	0.97	34.05	0.98
	b3	31.04	0.95	31.39	0.97	26.12	0.90	22.08	0.84	28.29	0.94	33.73	0.98
	b4	27.99	0.92	29.83	0.96	25.90	0.88	25.63	0.80	30.38	0.95	33.79	0.97
	avg. b	29.31	0.92	30.23	0.96	25.48	0.88	23.77	0.80	29.79	0.95	33.19	0.97

041 In the table, the results are calculated as average of all frames in each video sequence. Video-based methods generally
042 perform better for Group *a* data with panning unstable camera. When camera motion becomes larger, video-based methods
043 face greater challenge than image-based methods, which explains why *TCL-TIP15'* performs better for Group *a* whilst worse
044 for Group *b* as compared with *DDN-CVPR17'*. However, our proposed *SPAC-CNN* proves to be able to produce best results
045 among all competing methods for both Group *a* and Group *b* data.
046
047
048
049
050
051
052
053

108

2. Additional Visual Comparison

109
110
111
112
113

We show selected frames of each test video to highlight the performance of our method. Comparisons are carried out between *DDN-CVPR17'* [1] (single image-based), *TCL-TIP15'* [2] (video-based), *VST-ICCV17'* [4] (video-based), *VMD-CVPR17'* [3] (video-based), and the proposed *SPAC-CNN*.

114



Figure 1: Derain output comparison for testing data *a1*: synthetic rain, panning unstable camera.

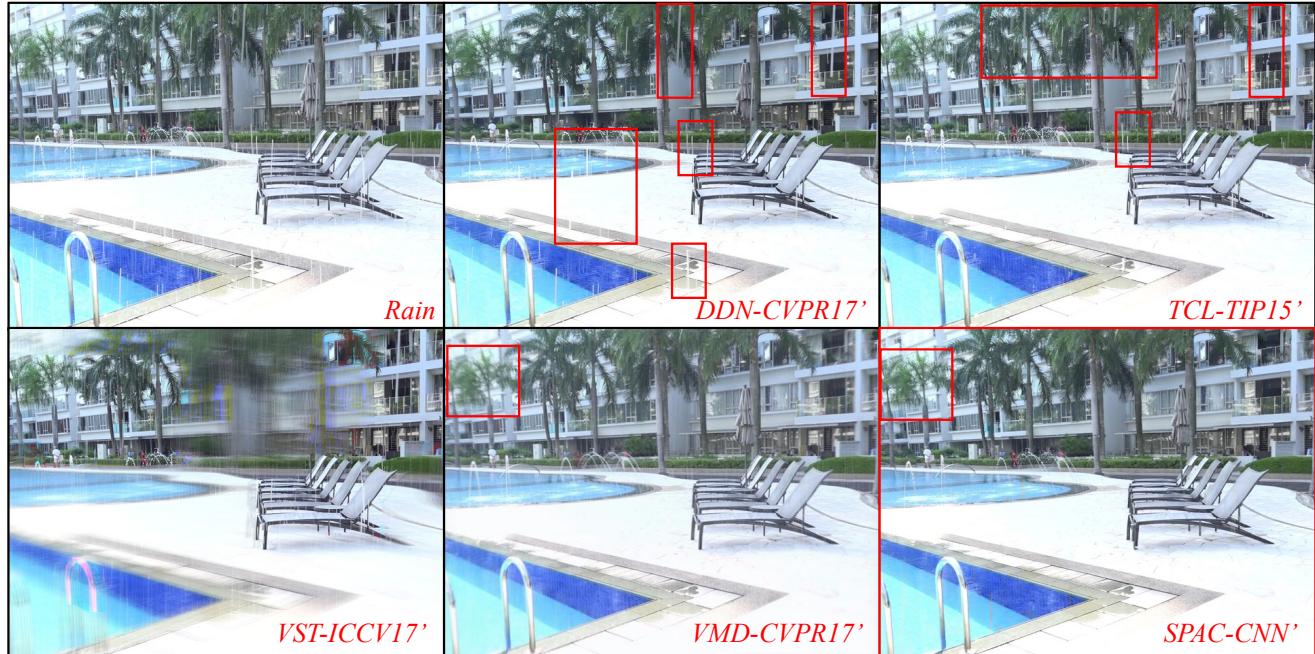
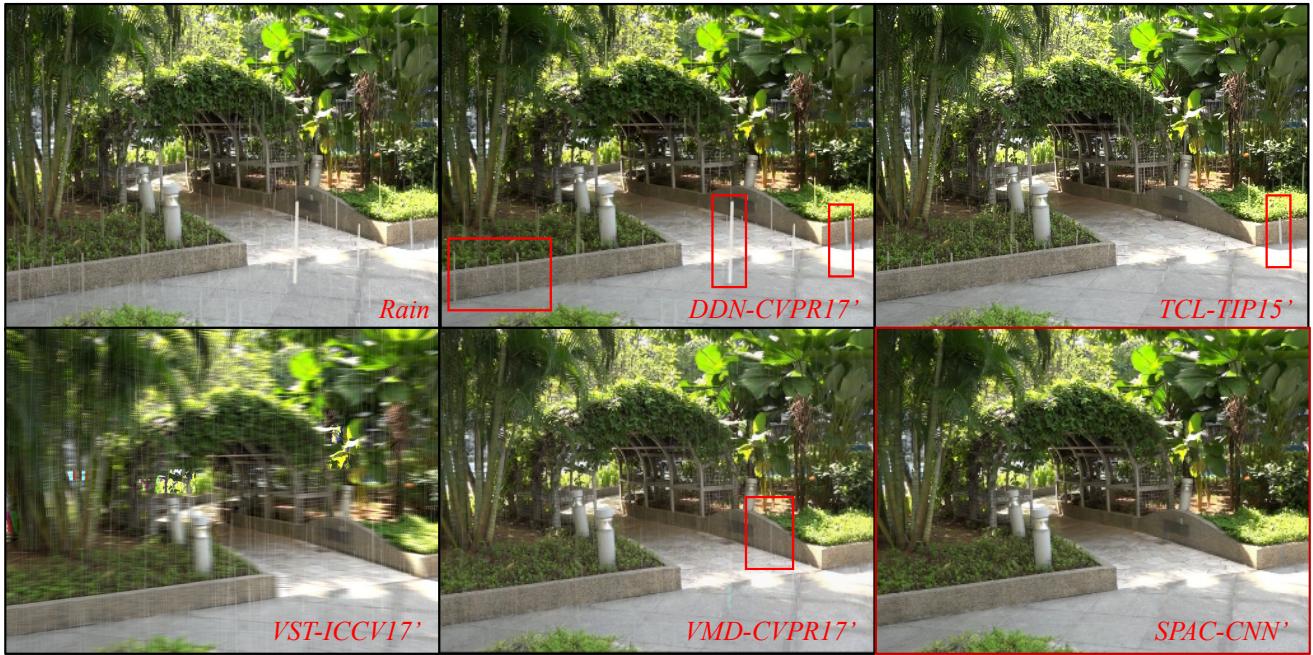
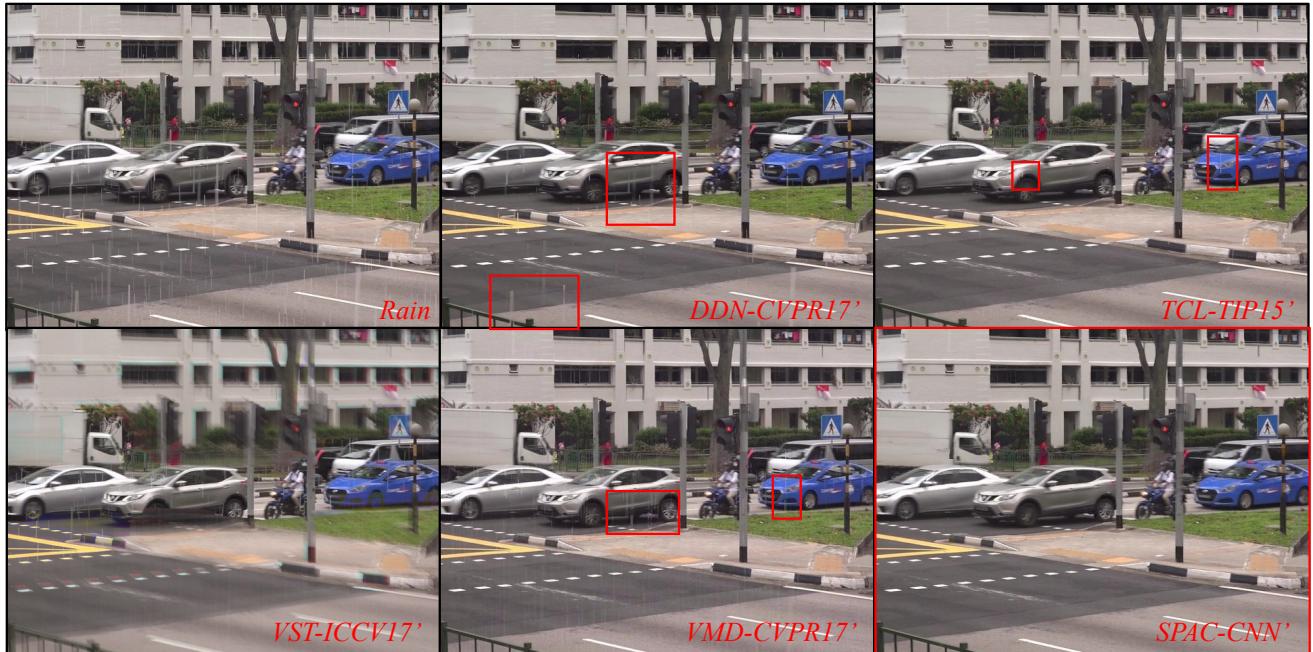
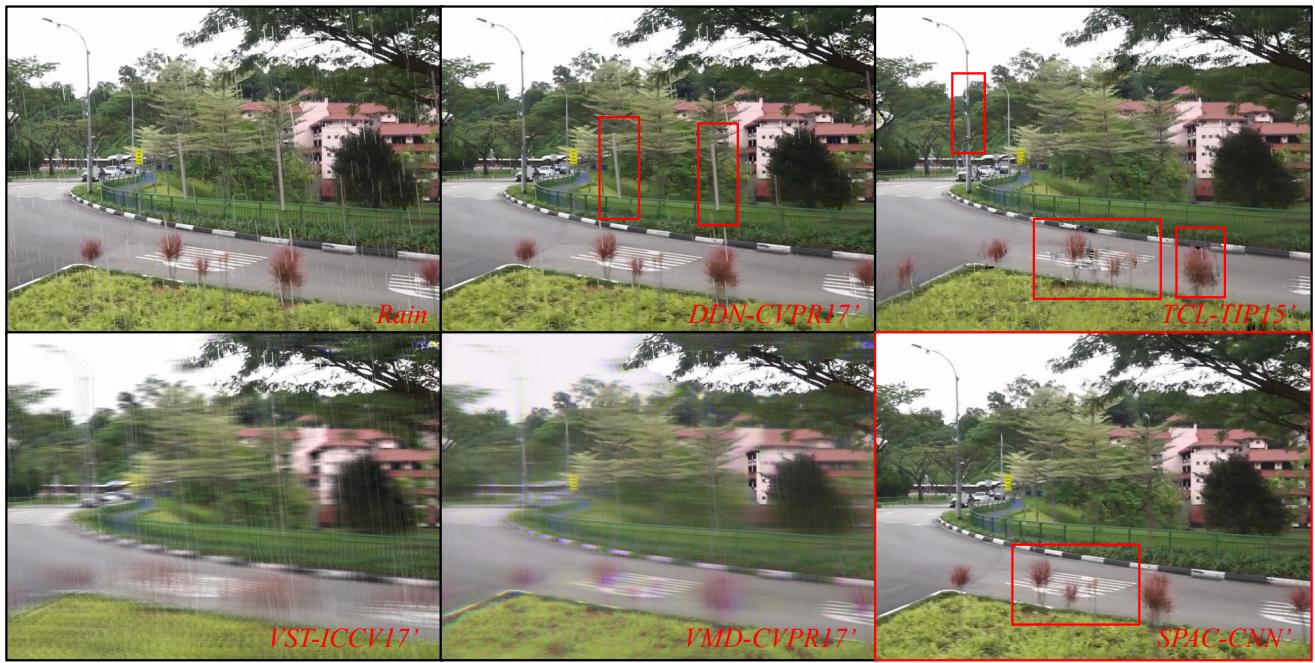
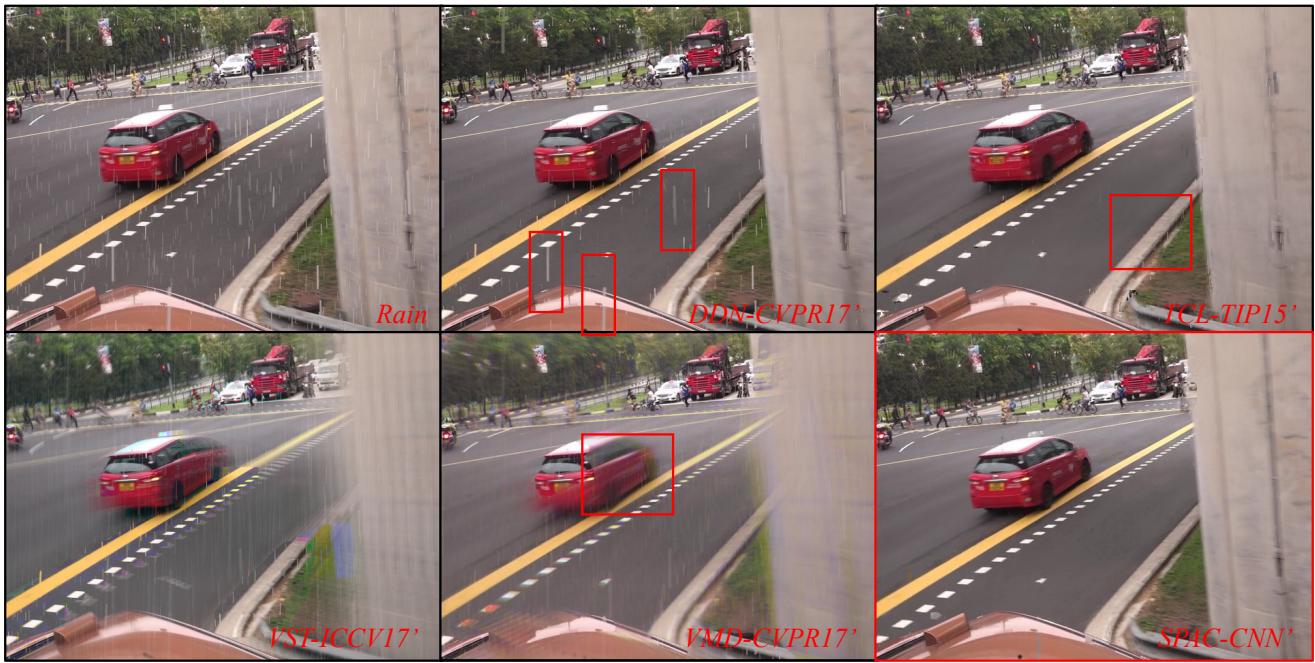
134
135
136
137
138

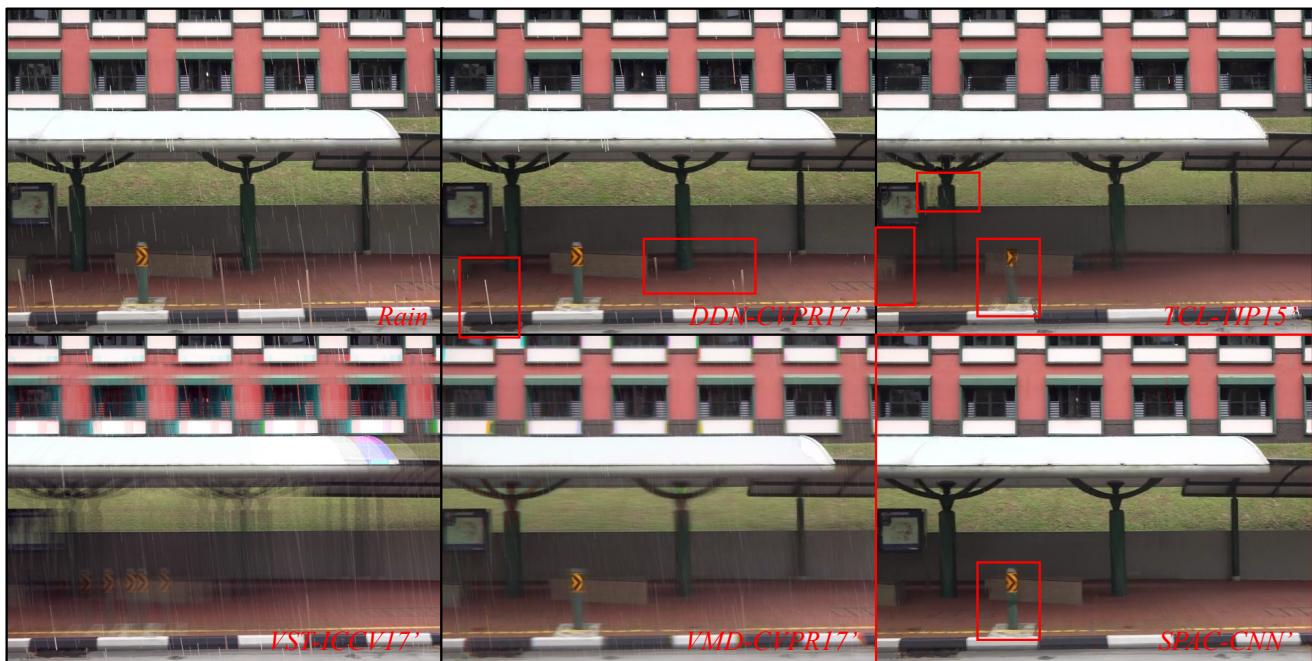
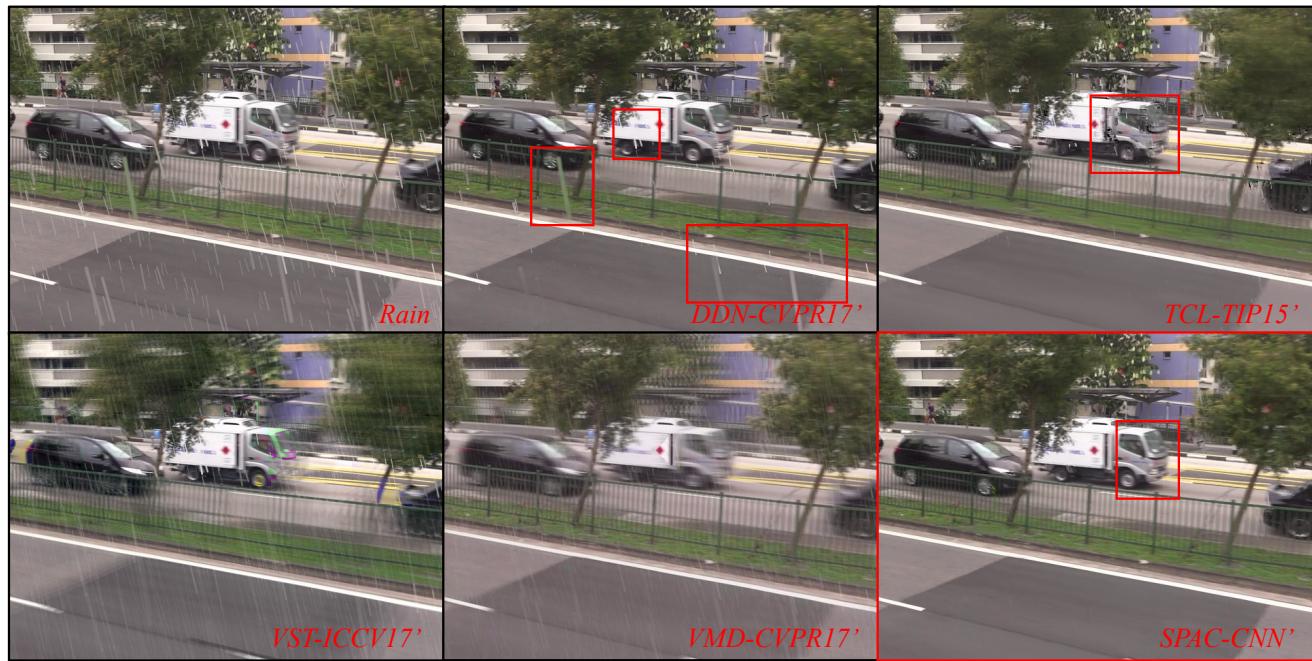
Figure 2: Derain output comparison for testing data *a2*: synthetic rain, panning unstable camera.

159
160
161

162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215

216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237Figure 3: Derain output comparison for testing data $a3$: synthetic rain, panning unstable camera.270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269Figure 4: Derain output comparison for testing data $a4$: synthetic rain, stable camera, very dynamic scene.292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323

Figure 5: Derain output comparison for testing data $b1$: synthetic rain, fast moving camera.Figure 6: Derain output comparison for testing data $b2$: synthetic rain, fast moving camera.

Figure 7: Derain output comparison for testing data $b3$: synthetic rain, fast moving camera.Figure 8: Derain output comparison for testing data $b4$: synthetic rain, fast moving camera.

540	References	594
541		595
542	[1] X. Fu, J. Huang, D. Z. Y. Huang, X. Ding, and J. Paisley. Removing rain from single images via a deep detail network. In <i>IEEE Conference on Computer Vision and Pattern Recognition</i> , 2017. 1, 2	596
543		597
544	[2] J. H. Kim, J. Y. Sim, and C. S. Kim. Video deraining and desnowing using temporal correlation and low-rank matrix completion. <i>IEEE Transactions on Image Processing</i> , 24(9):2658–2670, Sept 2015. 1, 2	598
545		599
546	[3] W. Ren, J. Tian, Z. Han, A. Chan, and Y. Tang. Video desnowing and deraining based on matrix decomposition. In <i>IEEE Conference on Computer Vision and Pattern Recognition</i> , pages 4210–4219, 2017. 1, 2	600
547		601
548	[4] W. Wei, L. Yi, Q. Xie, Q. Zhao, D. Meng, and Z. Xu. Should we encode rain streaks in video as deterministic or stochastic? In <i>International Conference on Computer Vision</i> , pages 2535–2544, Oct 2017. 1, 2	602
549		603
550		604
551		605
552		606
553		607
554		608
555		609
556		610
557		611
558		612
559		613
560		614
561		615
562		616
563		617
564		618
565		619
566		620
567		621
568		622
569		623
570		624
571		625
572		626
573		627
574		628
575		629
576		630
577		631
578		632
579		633
580		634
581		635
582		636
583		637
584		638
585		639
586		640
587		641
588		642
589		643
590		644
591		645
592		646
593		647