YOLOv3-SPP의 layer와 input&output 크기

Sequential 단위로 layer index 취급하겠음. 초기 img_size = 416 설정함.

idx	Layer	Input (C, H, W)	Output (C, H, W)	동작 방식
0	Conv-BN-LeakyReLU	(3, 416, 416)	(32, 416, 416)	
1	Conv-BN-LeakyReLU	(32, 416, 416)	(64, 208, 208)	
2	Conv-BN-LeakyReLU	(64, 208, 208)	(32, 208, 208)	
3	Conv-BN-LeakyReLU	(32, 208, 208)	(64, 208, 208)	
4	weightedFeatureFusion	(64, 208, 208)	(64, 208, 208)	Layer4 input + Layer1 output
5	Conv-BN-LeakyReLU	(64, 208, 208)	(128, 104, 104)	
6	Conv-BN-LeakyReLU	(128, 104, 104)	(64, 104, 104)	
7	Conv-BN-LeakyReLU	(64, 104, 104)	(128, 104, 104)	
8	weightedFeatureFusion	(128, 104, 104)	(128, 104, 104)	Layer8 input + Layer5 output
9	Conv-BN-LeakyReLU	(128, 104, 104)	(128, 104, 104)	
10	Conv-BN-LeakyReLU	(128, 104, 104)	(128, 104, 104)	
11	weightedFeatureFusion	(128, 104, 104)	(128, 104, 104)	Layer11 input + Layer8 output
12	Conv-BN-LeakyReLU	(128, 104, 104)	(256, 52, 52),	
13	Conv-BN-LeakyReLU	(256, 52, 52)	(128, 52, 52)	
14	Conv-BN-LeakyReLU	(128, 52, 52)	(256, 52, 52)	
15	weightedFeatureFusion	(256, 52, 52)	(256, 52, 52)	
16	Conv-BN-LeakyReLU	(256, 52, 52)	(128, 52, 52)	
17	Conv-BN-LeakyReLU	(128, 52, 52)	(256, 52, 52)	
18	weightedFeatureFusion	(256, 52, 52)	(256, 52, 52)	
19	Conv-BN-LeakyReLU	(256, 52, 52)	(128, 52, 52)	
20	Conv-BN-LeakyReLU	(128, 52, 52)	(256, 52, 52)	
21	weightedFeatureFusion	(256, 52, 52)	(256, 52, 52)	
22	Conv-BN-LeakyReLU	(256, 52, 52)	(128, 52, 52)	
23	Conv-BN-LeakyReLU	(128, 52, 52)	(256, 52, 52)	
24	weightedFeatureFusion	(256, 52, 52)	(256, 52, 52)	
25	Conv-BN-LeakyReLU	(256, 52, 52)	(128, 52, 52)	
26	Conv-BN-LeakyReLU	(128, 52, 52)	(256, 52, 52)	
27	weightedFeatureFusion	(256, 52, 52)	(256, 52, 52)	
28	Conv-BN-LeakyReLU	(256, 52, 52)	(128, 52, 52)	
29	Conv-BN-LeakyReLU	(128, 52, 52)	(256, 52, 52)	
30	weightedFeatureFusion	(256, 52, 52)	(256, 52, 52)	
31	Conv-BN-LeakyReLU	(256, 52, 52)	(128, 52, 52)	
32	Conv-BN-LeakyReLU	(128, 52, 52)	(256, 52, 52)	
33	weightedFeatureFusion	(256, 52, 52)	(256, 52, 52)	
34	Conv-BN-LeakyReLU	(256, 52, 52)	(128, 52, 52)	
35	Conv-BN-LeakyReLU	(128, 52, 52)	(256, 52, 52)	
36	weightedFeatureFusion	(256, 52, 52)	(256, 52, 52)	
37	Conv-BN-LeakyReLU	(256, 52, 52)	(512, 26, 26)	
38	Conv-BN-LeakyReLU	(512, 26, 26)	(256, 26, 26)	

39	Conv-BN-LeakyReLU	(256, 26, 26)	(512, 26, 26)	
40	weightedFeatureFusion	(512, 26, 26)	(512, 26, 26)	
41	Conv-BN-LeakyReLU	(512, 26, 26)	(256, 26, 26)	
42	Conv-BN-LeakyReLU	(256, 26, 26)	(512, 26, 26)	
43	weightedFeatureFusion	(512, 26, 26)	(512, 26, 26)	
44	Conv-BN-LeakyReLU	(512, 26, 26)	(256, 26, 26)	
45	Conv-BN-LeakyReLU	(256, 26, 26)	(512, 26, 26)	
46	weightedFeatureFusion	(512, 26, 26)	(512, 26, 26)	
47	Conv-BN-LeakyReLU	(512, 26, 26)	(256, 26, 26)	
48	Conv-BN-LeakyReLU	(256, 26, 26)	(512, 26, 26)	
49	weightedFeatureFusion	(512, 26, 26)	(512, 26, 26)	
50	Conv-BN-LeakyReLU	(512, 26, 26)	(256, 26, 26)	
51	Conv-BN-LeakyReLU	(256, 26, 26)	(512, 26, 26)	
52	weightedFeatureFusion	(512, 26, 26)	(512, 26, 26)	
53	Conv-BN-LeakyReLU	(512, 26, 26)	(256, 26, 26)	
54	Conv-BN-LeakyReLU	(256, 26, 26)	(512, 26, 26)	
55	weightedFeatureFusion	(512, 26, 26)	(512, 26, 26)	
56	Conv-BN-LeakyReLU	(512, 26, 26)	(256, 26, 26)	
57	Conv-BN-LeakyReLU	(256, 26, 26)	(512, 26, 26)	
58	weightedFeatureFusion	(512, 26, 26)	(512, 26, 26)	
59	Conv-BN-LeakyReLU	(512, 26, 26)	(256, 26, 26)	
60	Conv-BN-LeakyReLU	(256, 26, 26)	(512, 26, 26)	
61	weightedFeatureFusion	(512, 26, 26)	(512, 26, 26)	
62	Conv-BN-LeakyReLU	(512, 26, 26)	(1024, 13, 13)	
63	Conv-BN-LeakyReLU	(1024, 13, 13)	(512, 13, 13)	
64	Conv-BN-LeakyReLU	(512, 13, 13)	(1024, 13, 13)	
65	weightedFeatureFusion	(1024, 13, 13)	(1024, 13, 13)	
66	Conv-BN-LeakyReLU	(1024, 13, 13)	(512, 13, 13)	
67	Conv-BN-LeakyReLU	(512, 13, 13)	(1024, 13, 13)	
68	weightedFeatureFusion	(1024, 13, 13)	(1024, 13, 13)	
69	Conv-BN-LeakyReLU	(1024, 13, 13)	(512, 13, 13)	
70	Conv-BN-LeakyReLU	(512, 13, 13)	(1024, 13, 13)	
71	weightedFeatureFusion	(1024, 13, 13)	(1024, 13, 13)	
72	Conv-BN-LeakyReLU	(1024, 13, 13)	(512, 13, 13)	
73	Conv-BN-LeakyReLU	(512, 13, 13)	(1024, 13, 13)	
74	weightedFeatureFusion	(1024, 13, 13)	(1024, 13, 13)	
75	Conv-BN-LeakyReLU	(1024, 13, 13)	(512, 13, 13)	
76	Conv-BN-LeakyReLU	(512, 13, 13)	(1024, 13, 13)	
77	Conv-BN-LeakyReLU	(1024, 13, 13)	(512, 13, 13)	
78	MaxPool	(512, 13, 13)	(512, 13, 13)	Ceil_mode=false: output크기 내림 dilation=1:stride제어하는 param 예) $H_{out} = \frac{H_{in}+2*p-d*(k-1)-1}{s} + 1$
79	Route		(512, 13, 13)	'layers':[-2] → x = out[77]

80	MaxPool	(512, 13, 13)	(512, 13, 13)	
81	Route		(512, 13, 13)	'layers':[-4] → x = out[77]
82	MaxPool	(512, 13, 13)	(512, 13, 13)	
83	Route		(2048, 13, 13)	'layers':[-1, -3, -5, -6] → x.concat(out[82]+out[80]+out[78]+out[77])
84	Conv-BN-LeakyReLU	(2048, 13, 13)	(512, 13, 13)	
85	Conv-BN-LeakyReLU	(512, 13, 13)	(512, 13, 13)	
86	Conv-BN-LeakyReLU	(512, 13, 13)	(512, 13, 13)	
87	Conv-BN-LeakyReLU	(512, 13, 13)	(1024, 13, 13)	
88	Conv	(1024, 13, 13)	(255, 13, 13)	
89	YOLO Layer	(255, 13, 13)	(3, 13, 13, 85)	Mask: [116, 90] , [156, 198] , [373, 326] Output: [bs, 3, 13, 13, 85]
90	Route		(512, 13, 13)	'layers':[-4] → x = out[86]
91	Conv-BN-LeakyReLU	(512, 13, 13)	(256, 13, 13)	
92	Upsample	(256, 13, 13)	(256, 26, 26)	
93	Route		(768, 26, 26)	'layers':[-1, 61] → x.concat(out[92]+out[61]) = (768,26,26)
94	Conv-BN-LeakyReLU	(768, 26, 26)	(256, 26, 26)	
95	Conv-BN-LeakyReLU	(256, 26, 26)	(512, 26, 26)	
96	Conv-BN-LeakyReLU	(512, 26, 26)	(256, 26, 26)	
97	Conv-BN-LeakyReLU	(256, 26, 26)	(512, 26, 26)	
98	Conv-BN-LeakyReLU	(512, 26, 26)	(256, 26, 26)	
99	Conv-BN-LeakyReLU	(256, 26, 26)	(512, 26, 26)	
100	Conv	(512, 26, 26)	(255, 26, 26)	
101	YOLO Layer	(255, 26, 26)	(3, 26, 26, 85)	Mask: [30, 61], [62, 45], [59, 119] Output: [bs, 3, 26, 26, 85]
102	Route		(256, 26, 26)	'layers':[-4] → x = out[98]
103	Conv-BN-LeakyReLU	(256, 26, 26)	(128, 26, 26)	
104	Upsample	(128, 26, 26)	(128, 52, 52)	
105	Route		(384, 52, 52)	'layers':[-1, 36] → x.concat(out[104]+out[36]) = (384,52,52)
106	Conv-BN-LeakyReLU	(384, 52, 52)	(128, 52, 52)	
107	Conv-BN-LeakyReLU	(128, 52, 52)	(256, 52, 52)	
108	Conv-BN-LeakyReLU	(256, 52, 52)	(128, 52, 52)	
109	Conv-BN-LeakyReLU	(128, 52, 52)	(256, 52, 52)	
110	Conv-BN-LeakyReLU	(256, 52, 52)	(128, 52, 52)	
111	Conv-BN-LeakyReLU	(128, 52, 52)	(256, 52, 52)	
112	Conv	(256, 52, 52)	(255, 52, 52)	
113	YOLO Layer	(255, 52, 52)		Mask: [10, 13], [16, 30], [33, 23] Output: [bs, 3, 52, 52, 85]

총 114개 layer