

□□ - backbone

```
1 D:\ProgramData\Anaconda3\envs\mmpretrain\python.exe E:\Code\mmpretrain\TEST_file\
backbone.py
2 torch.Size([1, 3, 224, 224])
3 -----
```

4	Layer (type)	Output Shape	Param #
5	=====	=====	=====
6	Conv2d-1	[-1, 16, 112, 112]	432
7	BatchNorm2d-2	[-1, 16, 112, 112]	32
8	Swish-3	[-1, 16, 112, 112]	0
9	ConvModule-4	[-1, 16, 112, 112]	0
10	Conv2d-5	[-1, 64, 112, 112]	1,024
11	BatchNorm2d-6	[-1, 64, 112, 112]	128
12	Swish-7	[-1, 64, 112, 112]	0
13	ConvModule-8	[-1, 64, 112, 112]	0
14	Conv2d-9	[-1, 64, 112, 112]	576
15	BatchNorm2d-10	[-1, 64, 112, 112]	128
16	Swish-11	[-1, 64, 112, 112]	0
17	ConvModule-12	[-1, 64, 112, 112]	0
18	Conv2d-13	[-1, 32, 112, 112]	2,048
19	BatchNorm2d-14	[-1, 32, 112, 112]	64
20	ConvModule-15	[-1, 32, 112, 112]	0
21	AdaptiveAvgPool2d-16	[-1, 32, 1, 1]	0
22	Conv2d-17	[-1, 2, 1, 1]	66
23	ReLU-18	[-1, 2, 1, 1]	0
24	Conv2d-19	[-1, 32, 1, 1]	96
25	Sigmoid-20	[-1, 32, 1, 1]	0
26	SEBlock-21	[-1, 32, 112, 112]	0

27	InvertedResidual-22	[-1, 32, 112, 112]	0
28	Conv2d-23	[-1, 128, 112, 112]	4,096
29	BatchNorm2d-24	[-1, 128, 112, 112]	256
30	Swish-25	[-1, 128, 112, 112]	0
31	ConvModule-26	[-1, 128, 112, 112]	0
32	Conv2d-27	[-1, 128, 56, 56]	1,152
33	BatchNorm2d-28	[-1, 128, 56, 56]	256
34	Swish-29	[-1, 128, 56, 56]	0
35	ConvModule-30	[-1, 128, 56, 56]	0
36	Conv2d-31	[-1, 48, 56, 56]	6,144
37	BatchNorm2d-32	[-1, 48, 56, 56]	96
38	ConvModule-33	[-1, 48, 56, 56]	0
39	AdaptiveAvgPool2d-34	[-1, 48, 1, 1]	0
40	Conv2d-35	[-1, 3, 1, 1]	147
41	ReLU-36	[-1, 3, 1, 1]	0
42	Conv2d-37	[-1, 48, 1, 1]	192
43	Sigmoid-38	[-1, 48, 1, 1]	0
44	SEBlock-39	[-1, 48, 1, 1]	0
45	InvertedResidual-40	[-1, 48, 56, 56]	0
46	Conv2d-41	[-1, 48, 56, 56]	0
47	BatchNorm2d-42	[-1, 192, 56, 56]	9,216
48	Swish-43	[-1, 192, 56, 56]	384
49	ConvModule-44	[-1, 192, 56, 56]	0
50	Conv2d-45	[-1, 192, 56, 56]	0
51	BatchNorm2d-46	[-1, 192, 56, 56]	1,728
52	Swish-47	[-1, 192, 56, 56]	384
53	ConvModule-48	[-1, 192, 56, 56]	0

54	Conv2d-49	[-1, 48, 56, 56]	9,216
55	BatchNorm2d-50	[-1, 48, 56, 56]	96
56	ConvModule-51	[-1, 48, 56, 56]	0
57	AdaptiveAvgPool2d-52	[-1, 48, 1, 1]	0
58	Conv2d-53	[-1, 3, 1, 1]	147
59	ReLU-54	[-1, 3, 1, 1]	0
60	Conv2d-55	[-1, 48, 1, 1]	192
61	Sigmoid-56	[-1, 48, 1, 1]	0
62	SEBlock-57	[-1, 48, 56, 56]	0
63	InvertedResidual-58	[-1, 48, 56, 56]	0
64	Conv2d-59	[-1, 192, 56, 56]	9,216
65	BatchNorm2d-60	[-1, 192, 56, 56]	384
66	Swish-61	[-1, 192, 56, 56]	0
67	ConvModule-62	[-1, 192, 56, 56]	0
68	Conv2d-63	[-1, 192, 56, 56]	1,728
69	BatchNorm2d-64	[-1, 192, 56, 56]	384
70	Swish-65	[-1, 192, 56, 56]	0
71	ConvModule-66	[-1, 192, 56, 56]	0
72	Conv2d-67	[-1, 48, 56, 56]	9,216
73	BatchNorm2d-68	[-1, 48, 56, 56]	96
74	ConvModule-69	[-1, 48, 56, 56]	0
75	AdaptiveAvgPool2d-70	[-1, 48, 1, 1]	0
76	Conv2d-71	[-1, 3, 1, 1]	147
77	ReLU-72	[-1, 3, 1, 1]	0
78	Conv2d-73	[-1, 48, 1, 1]	192
79	Sigmoid-74	[-1, 48, 1, 1]	0
80	SEBlock-75	[-1, 48, 56, 56]	0

81	InvertedResidual-76	[-1, 48, 56, 56]	0
82	Conv2d-77	[-1, 192, 56, 56]	9,216
83	BatchNorm2d-78	[-1, 192, 56, 56]	384
84	Swish-79	[-1, 192, 56, 56]	0
85	ConvModule-80	[-1, 192, 56, 56]	0
86	Conv2d-81	[-1, 192, 28, 28]	1,728
87	BatchNorm2d-82	[-1, 192, 28, 28]	384
88	Swish-83	[-1, 192, 28, 28]	0
89	ConvModule-84	[-1, 192, 28, 28]	0
90	Conv2d-85	[-1, 64, 28, 28]	12,288
91	BatchNorm2d-86	[-1, 64, 28, 28]	128
92	ConvModule-87	[-1, 64, 28, 28]	0
93	AdaptiveAvgPool2d-88	[-1, 64, 1, 1]	0
94	Conv2d-89	[-1, 4, 1, 1]	260
95	ReLU-90	[-1, 4, 1, 1]	0
96	Conv2d-91	[-1, 64, 1, 1]	320
97	Sigmoid-92	[-1, 64, 1, 1]	0
98	SEBlock-93	[-1, 64, 28, 28]	0
99	InvertedResidual-94	[-1, 64, 28, 28]	0
100	Conv2d-95	[-1, 22, 28, 28]	4,356
101	Conv2d-96	[-1, 22, 28, 28]	4,356
102	ConvConcatModule-97	[-1, 64, 28, 28]	0
103	Conv2d-98	[-1, 96, 28, 28]	6,144
104	ConvModule-99	[-1, 96, 28, 28]	0
105	LayerNorm-100	[-1, 196, 96]	192
106	Linear-101	[-1, 196, 288]	27,936
107	Linear-102	[-1, 196, 96]	9,312

□□ - backbone

108	Dropout-103	[-1, 196, 96]	0
109	Identity-104	[-1, 196, 96]	0
110	DropPath-105	[-1, 196, 96]	0
111	MultiheadAttention-106	[-1, 196, 96]	0
112	LayerNorm-107	[-1, 196, 96]	192
113	Linear-108	[-1, 196, 192]	18,624
114	Swish-109	[-1, 196, 192]	0
115	Dropout-110	[-1, 196, 192]	0
116	Linear-111	[-1, 196, 96]	18,528
117	Dropout-112	[-1, 196, 96]	0
118	Identity-113	[-1, 196, 96]	0
119	DropPath-114	[-1, 196, 96]	0
120	FFN-115	[-1, 196, 96]	0
121	TransformerEncoderLayer-116	[-1, 196, 96]	0
122	LayerNorm-117	[-1, 196, 96]	192
123	Linear-118	[-1, 196, 288]	27,936
124	Linear-119	[-1, 196, 96]	9,312
125	Dropout-120	[-1, 196, 96]	0
126	Identity-121	[-1, 196, 96]	0
127	DropPath-122	[-1, 196, 96]	0
128	MultiheadAttention-123	[-1, 196, 96]	0
129	LayerNorm-124	[-1, 196, 96]	192
130	Linear-125	[-1, 196, 192]	18,624
131	Swish-126	[-1, 196, 192]	0
132	Dropout-127	[-1, 196, 192]	0
133	Linear-128	[-1, 196, 96]	18,528
134	Dropout-129	[-1, 196, 96]	0

□□ - backbone

135	Identity-130	[-1, 196, 96]	0
136	DropPath-131	[-1, 196, 96]	0
137	FFN-132	[-1, 196, 96]	0
138	TransformerEncoderLayer-133	[-1, 196, 96]	0
139	LayerNorm-134	[-1, 196, 96]	192
140	Conv2d-135	[-1, 64, 28, 28]	6,144
141	BatchNorm2d-136	[-1, 64, 28, 28]	128
142	Swish-137	[-1, 64, 28, 28]	0
143	ConvModule-138	[-1, 64, 28, 28]	0
144	Conv2d-139	[-1, 64, 28, 28]	92,160
145	BatchNorm2d-140	[-1, 64, 28, 28]	128
146	Swish-141	[-1, 64, 28, 28]	0
147	ConvModule-142	[-1, 64, 28, 28]	0
148	MobileVitBlock-143	[-1, 64, 28, 28]	0
149	Conv2d-144	[-1, 256, 28, 28]	16,384
150	BatchNorm2d-145	[-1, 256, 28, 28]	512
151	Swish-146	[-1, 256, 28, 28]	0
152	ConvModule-147	[-1, 256, 28, 28]	0
153	Conv2d-148	[-1, 256, 14, 14]	2,304
154	BatchNorm2d-149	[-1, 256, 14, 14]	512
155	Swish-150	[-1, 256, 14, 14]	0
156	ConvModule-151	[-1, 256, 14, 14]	0
157	Conv2d-152	[-1, 80, 14, 14]	20,480
158	BatchNorm2d-153	[-1, 80, 14, 14]	160
159	ConvModule-154	[-1, 80, 14, 14]	0
160	AdaptiveAvgPool2d-155	[-1, 80, 1, 1]	0
161	Conv2d-156	[-1, 5, 1, 1]	405

162	ReLU-157	[-1, 5, 1, 1]	0
163	Conv2d-158	[-1, 80, 1, 1]	480
164	Sigmoid-159	[-1, 80, 1, 1]	0
165	SEBlock-160	[-1, 80, 14, 14]	0
166	InvertedResidual-161	[-1, 80, 14, 14]	0
167	Conv2d-162	[-1, 20, 14, 14]	3,600
168	Conv2d-163	[-1, 20, 14, 14]	3,600
169	Conv2d-164	[-1, 20, 14, 14]	3,600
170	ConvConcatModule-165	[-1, 80, 14, 14]	0
171	Conv2d-166	[-1, 120, 14, 14]	9,600
172	ConvModule-167	[-1, 120, 14, 14]	0
173	LayerNorm-168	[-1, 49, 120]	240
174	Linear-169	[-1, 49, 360]	43,560
175	Linear-170	[-1, 49, 120]	14,520
176	Dropout-171	[-1, 49, 120]	0
177	Identity-172	[-1, 49, 120]	0
178	DropPath-173	[-1, 49, 120]	0
179	MultiheadAttention-174	[-1, 49, 120]	0
180	LayerNorm-175	[-1, 49, 120]	240
181	Linear-176	[-1, 49, 240]	29,040
182	Swish-177	[-1, 49, 240]	0
183	Dropout-178	[-1, 49, 240]	0
184	Linear-179	[-1, 49, 120]	28,920
185	Dropout-180	[-1, 49, 120]	0
186	Identity-181	[-1, 49, 120]	0
187	DropPath-182	[-1, 49, 120]	0
188	FFN-183	[-1, 49, 120]	0

□□ - backbone

189	TransformerEncoderLayer-184	[-1, 49, 120]	0
190	LayerNorm-185	[-1, 49, 120]	240
191	Linear-186	[-1, 49, 360]	43,560
192	Linear-187	[-1, 49, 120]	14,520
193	Dropout-188	[-1, 49, 120]	0
194	Identity-189	[-1, 49, 120]	0
195	DropPath-190	[-1, 49, 120]	0
196	MultiheadAttention-191	[-1, 49, 120]	0
197	LayerNorm-192	[-1, 49, 120]	240
198	Linear-193	[-1, 49, 240]	29,040
199	Swish-194	[-1, 49, 240]	0
200	Dropout-195	[-1, 49, 240]	0
201	Linear-196	[-1, 49, 120]	28,920
202	Dropout-197	[-1, 49, 120]	0
203	Identity-198	[-1, 49, 120]	0
204	DropPath-199	[-1, 49, 120]	0
205	FFN-200	[-1, 49, 120]	0
206	TransformerEncoderLayer-201	[-1, 49, 120]	0
207	LayerNorm-202	[-1, 49, 120]	240
208	Linear-203	[-1, 49, 360]	43,560
209	Linear-204	[-1, 49, 120]	14,520
210	Dropout-205	[-1, 49, 120]	0
211	Identity-206	[-1, 49, 120]	0
212	DropPath-207	[-1, 49, 120]	0
213	MultiheadAttention-208	[-1, 49, 120]	0
214	LayerNorm-209	[-1, 49, 120]	240
215	Linear-210	[-1, 49, 240]	29,040

□□ - backbone

216	Swish-211	[-1, 49, 240]	0
217	Dropout-212	[-1, 49, 240]	0
218	Linear-213	[-1, 49, 120]	28,920
219	Dropout-214	[-1, 49, 120]	0
220	Identity-215	[-1, 49, 120]	0
221	DropPath-216	[-1, 49, 120]	0
222	FFN-217	[-1, 49, 120]	0
223	TransformerEncoderLayer-218	[-1, 49, 120]	0
224	LayerNorm-219	[-1, 49, 120]	240
225	Linear-220	[-1, 49, 360]	43,560
226	Linear-221	[-1, 49, 120]	14,520
227	Dropout-222	[-1, 49, 120]	0
228	Identity-223	[-1, 49, 120]	0
229	DropPath-224	[-1, 49, 120]	0
230	MultiheadAttention-225	[-1, 49, 120]	0
231	LayerNorm-226	[-1, 49, 120]	240
232	Linear-227	[-1, 49, 240]	29,040
233	Swish-228	[-1, 49, 240]	0
234	Dropout-229	[-1, 49, 240]	0
235	Linear-230	[-1, 49, 120]	28,920
236	Dropout-231	[-1, 49, 120]	0
237	Identity-232	[-1, 49, 120]	0
238	DropPath-233	[-1, 49, 120]	0
239	FFN-234	[-1, 49, 120]	0
240	TransformerEncoderLayer-235	[-1, 49, 120]	0
241	LayerNorm-236	[-1, 49, 120]	240
242	Conv2d-237	[-1, 80, 14, 14]	9,600

□□ - backbone

243	BatchNorm2d-238	[-1, 80, 14, 14]	160
244	Swish-239	[-1, 80, 14, 14]	0
245	ConvModule-240	[-1, 80, 14, 14]	0
246	Conv2d-241	[-1, 80, 14, 14]	144,000
247	BatchNorm2d-242	[-1, 80, 14, 14]	160
248	Swish-243	[-1, 80, 14, 14]	0
249	ConvModule-244	[-1, 80, 14, 14]	0
250	MobileVitBlock-245	[-1, 80, 14, 14]	0
251	Conv2d-246	[-1, 320, 14, 14]	25,600
252	BatchNorm2d-247	[-1, 320, 14, 14]	640
253	Swish-248	[-1, 320, 14, 14]	0
254	ConvModule-249	[-1, 320, 14, 14]	0
255	Conv2d-250	[-1, 320, 7, 7]	2,880
256	BatchNorm2d-251	[-1, 320, 7, 7]	640
257	Swish-252	[-1, 320, 7, 7]	0
258	ConvModule-253	[-1, 320, 7, 7]	0
259	Conv2d-254	[-1, 96, 7, 7]	30,720
260	BatchNorm2d-255	[-1, 96, 7, 7]	192
261	ConvModule-256	[-1, 96, 7, 7]	0
262	AdaptiveAvgPool2d-257	[-1, 96, 1, 1]	0
263	Conv2d-258	[-1, 6, 1, 1]	582
264	ReLU-259	[-1, 6, 1, 1]	0
265	Conv2d-260	[-1, 96, 1, 1]	672
266	Sigmoid-261	[-1, 96, 1, 1]	0
267	SEBlock-262	[-1, 96, 7, 7]	0
268	InvertedResidual-263	[-1, 96, 7, 7]	0
269	Conv2d-264	[-1, 20, 7, 7]	3,600

□□ - backbone

270	Conv2d-265	[-1, 20, 7, 7]	3,600
271	Conv2d-266	[-1, 20, 7, 7]	3,600
272	Conv2d-267	[-1, 20, 7, 7]	3,600
273	ConvConcatModule-268	[-1, 96, 7, 7]	0
274	Conv2d-269	[-1, 144, 7, 7]	13,824
275	ConvModule-270	[-1, 144, 7, 7]	0
276	LayerNorm-271	[-1, 16, 144]	288
277	Linear-272	[-1, 16, 432]	62,640
278	Linear-273	[-1, 16, 144]	20,880
279	Dropout-274	[-1, 16, 144]	0
280	Identity-275	[-1, 16, 144]	0
281	DropPath-276	[-1, 16, 144]	0
282	MultiheadAttention-277	[-1, 16, 144]	0
283	LayerNorm-278	[-1, 16, 144]	288
284	Linear-279	[-1, 16, 288]	41,760
285	Swish-280	[-1, 16, 288]	0
286	Dropout-281	[-1, 16, 288]	0
287	Linear-282	[-1, 16, 144]	41,616
288	Dropout-283	[-1, 16, 144]	0
289	Identity-284	[-1, 16, 144]	0
290	DropPath-285	[-1, 16, 144]	0
291	FFN-286	[-1, 16, 144]	0
292	TransformerEncoderLayer-287	[-1, 16, 144]	0
293	LayerNorm-288	[-1, 16, 144]	288
294	Linear-289	[-1, 16, 432]	62,640
295	Linear-290	[-1, 16, 144]	20,880
296	Dropout-291	[-1, 16, 144]	0

□□ - backbone

297	Identity-292	[-1, 16, 144]	0
298	DropPath-293	[-1, 16, 144]	0
299	MultiheadAttention-294	[-1, 16, 144]	0
300	LayerNorm-295	[-1, 16, 144]	288
301	Linear-296	[-1, 16, 288]	41,760
302	Swish-297	[-1, 16, 288]	0
303	Dropout-298	[-1, 16, 288]	0
304	Linear-299	[-1, 16, 144]	41,616
305	Dropout-300	[-1, 16, 144]	0
306	Identity-301	[-1, 16, 144]	0
307	DropPath-302	[-1, 16, 144]	0
308	FFN-303	[-1, 16, 144]	0
309	TransformerEncoderLayer-304	[-1, 16, 144]	0
310	LayerNorm-305	[-1, 16, 144]	288
311	Linear-306	[-1, 16, 432]	62,640
312	Linear-307	[-1, 16, 144]	20,880
313	Dropout-308	[-1, 16, 144]	0
314	Identity-309	[-1, 16, 144]	0
315	DropPath-310	[-1, 16, 144]	0
316	MultiheadAttention-311	[-1, 16, 144]	0
317	LayerNorm-312	[-1, 16, 144]	288
318	Linear-313	[-1, 16, 288]	41,760
319	Swish-314	[-1, 16, 288]	0
320	Dropout-315	[-1, 16, 288]	0
321	Linear-316	[-1, 16, 144]	41,616
322	Dropout-317	[-1, 16, 144]	0
323	Identity-318	[-1, 16, 144]	0

□□ - backbone

324	DropPath-319	[-1, 16, 144]	0
325	FFN-320	[-1, 16, 144]	0
326	TransformerEncoderLayer-321	[-1, 16, 144]	0
327	LayerNorm-322	[-1, 16, 144]	288
328	Conv2d-323	[-1, 96, 7, 7]	13,824
329	BatchNorm2d-324	[-1, 96, 7, 7]	192
330	Swish-325	[-1, 96, 7, 7]	0
331	ConvModule-326	[-1, 96, 7, 7]	0
332	Conv2d-327	[-1, 96, 7, 7]	207,360
333	BatchNorm2d-328	[-1, 96, 7, 7]	192
334	Swish-329	[-1, 96, 7, 7]	0
335	ConvModule-330	[-1, 96, 7, 7]	0
336	MobileVitBlock-331	[-1, 96, 7, 7]	0
337	Conv2d-332	[-1, 384, 7, 7]	36,864
338	BatchNorm2d-333	[-1, 384, 7, 7]	768
339	Swish-334	[-1, 384, 7, 7]	0
340	ConvModule-335	[-1, 384, 7, 7]	0
341	=====		
342	Total params: 1,881,474		
343	Trainable params: 1,881,474		
344	Non-trainable params: 0		
345	-----		
346	Input size (MB): 0.57		
347	Forward/backward pass size (MB): 277.72		
348	Params size (MB): 7.18		
349	Estimated Total Size (MB): 285.47		
350	-----		

□□ - backbone

351

352 □□□□□□□□□□□□□□□□ 0

353