

Loaded pretrained weights for efficientnet-b3

Layer (type)	Output Shape	Param #
=====		
ZeroPad2d-1	[-1, 3, 225, 225]	0
Conv2dStaticSamePadding-2	[-1, 40, 112, 112]	1,080
BatchNorm2d-3	[-1, 40, 112, 112]	80
MemoryEfficientSwish-4	[-1, 40, 112, 112]	0
ZeroPad2d-5	[-1, 40, 114, 114]	0
Conv2dStaticSamePadding-6	[-1, 40, 112, 112]	360
BatchNorm2d-7	[-1, 40, 112, 112]	80
MemoryEfficientSwish-8	[-1, 40, 112, 112]	0
Identity-9	[-1, 40, 1, 1]	0
Conv2dStaticSamePadding-10	[-1, 10, 1, 1]	410
MemoryEfficientSwish-11	[-1, 10, 1, 1]	0
Identity-12	[-1, 10, 1, 1]	0
Conv2dStaticSamePadding-13	[-1, 40, 1, 1]	440
Identity-14	[-1, 40, 112, 112]	0
Conv2dStaticSamePadding-15	[-1, 24, 112, 112]	960
BatchNorm2d-16	[-1, 24, 112, 112]	48
MBConvBlock-17	[-1, 24, 112, 112]	0
ZeroPad2d-18	[-1, 24, 114, 114]	0
Conv2dStaticSamePadding-19	[-1, 24, 112, 112]	216
BatchNorm2d-20	[-1, 24, 112, 112]	48
MemoryEfficientSwish-21	[-1, 24, 112, 112]	0
Identity-22	[-1, 24, 1, 1]	0
Conv2dStaticSamePadding-23	[-1, 6, 1, 1]	150
MemoryEfficientSwish-24	[-1, 6, 1, 1]	0
Identity-25	[-1, 6, 1, 1]	0
Conv2dStaticSamePadding-26	[-1, 24, 1, 1]	168
Identity-27	[-1, 24, 112, 112]	0
Conv2dStaticSamePadding-28	[-1, 24, 112, 112]	576
BatchNorm2d-29	[-1, 24, 112, 112]	48
MBConvBlock-30	[-1, 24, 112, 112]	0
Identity-31	[-1, 24, 112, 112]	0
Conv2dStaticSamePadding-32	[-1, 144, 112, 112]	3,456
BatchNorm2d-33	[-1, 144, 112, 112]	288
MemoryEfficientSwish-34	[-1, 144, 112, 112]	0
ZeroPad2d-35	[-1, 144, 113, 113]	0
Conv2dStaticSamePadding-36	[-1, 144, 56, 56]	1,296
BatchNorm2d-37	[-1, 144, 56, 56]	288
MemoryEfficientSwish-38	[-1, 144, 56, 56]	0
Identity-39	[-1, 144, 1, 1]	0
Conv2dStaticSamePadding-40	[-1, 6, 1, 1]	870

MemoryEfficientSwish-41	[-1, 6, 1, 1]	0
Identity-42	[-1, 6, 1, 1]	0
Conv2dStaticSamePadding-43	[-1, 144, 1, 1]	1,008
Identity-44	[-1, 144, 56, 56]	0
Conv2dStaticSamePadding-45	[-1, 32, 56, 56]	4,608
BatchNorm2d-46	[-1, 32, 56, 56]	64
MBConvBlock-47	[-1, 32, 56, 56]	0
Identity-48	[-1, 32, 56, 56]	0
Conv2dStaticSamePadding-49	[-1, 192, 56, 56]	6,144
BatchNorm2d-50	[-1, 192, 56, 56]	384
MemoryEfficientSwish-51	[-1, 192, 56, 56]	0
ZeroPad2d-52	[-1, 192, 58, 58]	0
Conv2dStaticSamePadding-53	[-1, 192, 56, 56]	1,728
BatchNorm2d-54	[-1, 192, 56, 56]	384
MemoryEfficientSwish-55	[-1, 192, 56, 56]	0
Identity-56	[-1, 192, 1, 1]	0
Conv2dStaticSamePadding-57	[-1, 8, 1, 1]	1,544
MemoryEfficientSwish-58	[-1, 8, 1, 1]	0
Identity-59	[-1, 8, 1, 1]	0
Conv2dStaticSamePadding-60	[-1, 192, 1, 1]	1,728
Identity-61	[-1, 192, 56, 56]	0
Conv2dStaticSamePadding-62	[-1, 32, 56, 56]	6,144
BatchNorm2d-63	[-1, 32, 56, 56]	64
MBConvBlock-64	[-1, 32, 56, 56]	0
Identity-65	[-1, 32, 56, 56]	0
Conv2dStaticSamePadding-66	[-1, 192, 56, 56]	6,144
BatchNorm2d-67	[-1, 192, 56, 56]	384
MemoryEfficientSwish-68	[-1, 192, 56, 56]	0
ZeroPad2d-69	[-1, 192, 58, 58]	0
Conv2dStaticSamePadding-70	[-1, 192, 56, 56]	1,728
BatchNorm2d-71	[-1, 192, 56, 56]	384
MemoryEfficientSwish-72	[-1, 192, 56, 56]	0
Identity-73	[-1, 192, 1, 1]	0
Conv2dStaticSamePadding-74	[-1, 8, 1, 1]	1,544
MemoryEfficientSwish-75	[-1, 8, 1, 1]	0
Identity-76	[-1, 8, 1, 1]	0
Conv2dStaticSamePadding-77	[-1, 192, 1, 1]	1,728
Identity-78	[-1, 192, 56, 56]	0
Conv2dStaticSamePadding-79	[-1, 32, 56, 56]	6,144
BatchNorm2d-80	[-1, 32, 56, 56]	64
MBConvBlock-81	[-1, 32, 56, 56]	0
Identity-82	[-1, 32, 56, 56]	0
Conv2dStaticSamePadding-83	[-1, 192, 56, 56]	6,144
BatchNorm2d-84	[-1, 192, 56, 56]	384

MemoryEfficientSwish-85	[-1, 192, 56, 56]	0
ZeroPad2d-86	[-1, 192, 60, 60]	0
Conv2dStaticSamePadding-87	[-1, 192, 28, 28]	4,800
BatchNorm2d-88	[-1, 192, 28, 28]	384
MemoryEfficientSwish-89	[-1, 192, 28, 28]	0
Identity-90	[-1, 192, 1, 1]	0
Conv2dStaticSamePadding-91	[-1, 8, 1, 1]	1,544
MemoryEfficientSwish-92	[-1, 8, 1, 1]	0
Identity-93	[-1, 8, 1, 1]	0
Conv2dStaticSamePadding-94	[-1, 192, 1, 1]	1,728
Identity-95	[-1, 192, 28, 28]	0
Conv2dStaticSamePadding-96	[-1, 48, 28, 28]	9,216
BatchNorm2d-97	[-1, 48, 28, 28]	96
MBConvBlock-98	[-1, 48, 28, 28]	0
Identity-99	[-1, 48, 28, 28]	0
Conv2dStaticSamePadding-100	[-1, 288, 28, 28]	13,824
BatchNorm2d-101	[-1, 288, 28, 28]	576
MemoryEfficientSwish-102	[-1, 288, 28, 28]	0
ZeroPad2d-103	[-1, 288, 32, 32]	0
Conv2dStaticSamePadding-104	[-1, 288, 28, 28]	7,200
BatchNorm2d-105	[-1, 288, 28, 28]	576
MemoryEfficientSwish-106	[-1, 288, 28, 28]	0
Identity-107	[-1, 288, 1, 1]	0
Conv2dStaticSamePadding-108	[-1, 12, 1, 1]	3,468
MemoryEfficientSwish-109	[-1, 12, 1, 1]	0
Identity-110	[-1, 12, 1, 1]	0
Conv2dStaticSamePadding-111	[-1, 288, 1, 1]	3,744
Identity-112	[-1, 288, 28, 28]	0
Conv2dStaticSamePadding-113	[-1, 48, 28, 28]	13,824
BatchNorm2d-114	[-1, 48, 28, 28]	96
MBConvBlock-115	[-1, 48, 28, 28]	0
Identity-116	[-1, 48, 28, 28]	0
Conv2dStaticSamePadding-117	[-1, 288, 28, 28]	13,824
BatchNorm2d-118	[-1, 288, 28, 28]	576
MemoryEfficientSwish-119	[-1, 288, 28, 28]	0
ZeroPad2d-120	[-1, 288, 32, 32]	0
Conv2dStaticSamePadding-121	[-1, 288, 28, 28]	7,200
BatchNorm2d-122	[-1, 288, 28, 28]	576
MemoryEfficientSwish-123	[-1, 288, 28, 28]	0
Identity-124	[-1, 288, 1, 1]	0
Conv2dStaticSamePadding-125	[-1, 12, 1, 1]	3,468
MemoryEfficientSwish-126	[-1, 12, 1, 1]	0
Identity-127	[-1, 12, 1, 1]	0
Conv2dStaticSamePadding-128	[-1, 288, 1, 1]	3,744

Identity-129	[-1, 288, 28, 28]	0
Conv2dStaticSamePadding-130	[-1, 48, 28, 28]	13,824
BatchNorm2d-131	[-1, 48, 28, 28]	96
MBConvBlock-132	[-1, 48, 28, 28]	0
Identity-133	[-1, 48, 28, 28]	0
Conv2dStaticSamePadding-134	[-1, 288, 28, 28]	13,824
BatchNorm2d-135	[-1, 288, 28, 28]	576
MemoryEfficientSwish-136	[-1, 288, 28, 28]	0
ZeroPad2d-137	[-1, 288, 29, 29]	0
Conv2dStaticSamePadding-138	[-1, 288, 14, 14]	2,592
BatchNorm2d-139	[-1, 288, 14, 14]	576
MemoryEfficientSwish-140	[-1, 288, 14, 14]	0
Identity-141	[-1, 288, 1, 1]	0
Conv2dStaticSamePadding-142	[-1, 12, 1, 1]	3,468
MemoryEfficientSwish-143	[-1, 12, 1, 1]	0
Identity-144	[-1, 12, 1, 1]	0
Conv2dStaticSamePadding-145	[-1, 288, 1, 1]	3,744
Identity-146	[-1, 288, 14, 14]	0
Conv2dStaticSamePadding-147	[-1, 96, 14, 14]	27,648
BatchNorm2d-148	[-1, 96, 14, 14]	192
MBConvBlock-149	[-1, 96, 14, 14]	0
Identity-150	[-1, 96, 14, 14]	0
Conv2dStaticSamePadding-151	[-1, 576, 14, 14]	55,296
BatchNorm2d-152	[-1, 576, 14, 14]	1,152
MemoryEfficientSwish-153	[-1, 576, 14, 14]	0
ZeroPad2d-154	[-1, 576, 16, 16]	0
Conv2dStaticSamePadding-155	[-1, 576, 14, 14]	5,184
BatchNorm2d-156	[-1, 576, 14, 14]	1,152
MemoryEfficientSwish-157	[-1, 576, 14, 14]	0
Identity-158	[-1, 576, 1, 1]	0
Conv2dStaticSamePadding-159	[-1, 24, 1, 1]	13,848
MemoryEfficientSwish-160	[-1, 24, 1, 1]	0
Identity-161	[-1, 24, 1, 1]	0
Conv2dStaticSamePadding-162	[-1, 576, 1, 1]	14,400
Identity-163	[-1, 576, 14, 14]	0
Conv2dStaticSamePadding-164	[-1, 96, 14, 14]	55,296
BatchNorm2d-165	[-1, 96, 14, 14]	192
MBConvBlock-166	[-1, 96, 14, 14]	0
Identity-167	[-1, 96, 14, 14]	0
Conv2dStaticSamePadding-168	[-1, 576, 14, 14]	55,296
BatchNorm2d-169	[-1, 576, 14, 14]	1,152
MemoryEfficientSwish-170	[-1, 576, 14, 14]	0
ZeroPad2d-171	[-1, 576, 16, 16]	0
Conv2dStaticSamePadding-172	[-1, 576, 14, 14]	5,184

BatchNorm2d-173	[-1, 576, 14, 14]	1,152
MemoryEfficientSwish-174	[-1, 576, 14, 14]	0
Identity-175	[-1, 576, 1, 1]	0
Conv2dStaticSamePadding-176	[-1, 24, 1, 1]	13,848
MemoryEfficientSwish-177	[-1, 24, 1, 1]	0
Identity-178	[-1, 24, 1, 1]	0
Conv2dStaticSamePadding-179	[-1, 576, 1, 1]	14,400
Identity-180	[-1, 576, 14, 14]	0
Conv2dStaticSamePadding-181	[-1, 96, 14, 14]	55,296
BatchNorm2d-182	[-1, 96, 14, 14]	192
MBConvBlock-183	[-1, 96, 14, 14]	0
Identity-184	[-1, 96, 14, 14]	0
Conv2dStaticSamePadding-185	[-1, 576, 14, 14]	55,296
BatchNorm2d-186	[-1, 576, 14, 14]	1,152
MemoryEfficientSwish-187	[-1, 576, 14, 14]	0
ZeroPad2d-188	[-1, 576, 16, 16]	0
Conv2dStaticSamePadding-189	[-1, 576, 14, 14]	5,184
BatchNorm2d-190	[-1, 576, 14, 14]	1,152
MemoryEfficientSwish-191	[-1, 576, 14, 14]	0
Identity-192	[-1, 576, 1, 1]	0
Conv2dStaticSamePadding-193	[-1, 24, 1, 1]	13,848
MemoryEfficientSwish-194	[-1, 24, 1, 1]	0
Identity-195	[-1, 24, 1, 1]	0
Conv2dStaticSamePadding-196	[-1, 576, 1, 1]	14,400
Identity-197	[-1, 576, 14, 14]	0
Conv2dStaticSamePadding-198	[-1, 96, 14, 14]	55,296
BatchNorm2d-199	[-1, 96, 14, 14]	192
MBConvBlock-200	[-1, 96, 14, 14]	0
Identity-201	[-1, 96, 14, 14]	0
Conv2dStaticSamePadding-202	[-1, 576, 14, 14]	55,296
BatchNorm2d-203	[-1, 576, 14, 14]	1,152
MemoryEfficientSwish-204	[-1, 576, 14, 14]	0
ZeroPad2d-205	[-1, 576, 16, 16]	0
Conv2dStaticSamePadding-206	[-1, 576, 14, 14]	5,184
BatchNorm2d-207	[-1, 576, 14, 14]	1,152
MemoryEfficientSwish-208	[-1, 576, 14, 14]	0
Identity-209	[-1, 576, 1, 1]	0
Conv2dStaticSamePadding-210	[-1, 24, 1, 1]	13,848
MemoryEfficientSwish-211	[-1, 24, 1, 1]	0
Identity-212	[-1, 24, 1, 1]	0
Conv2dStaticSamePadding-213	[-1, 576, 1, 1]	14,400
Identity-214	[-1, 576, 14, 14]	0
Conv2dStaticSamePadding-215	[-1, 96, 14, 14]	55,296
BatchNorm2d-216	[-1, 96, 14, 14]	192

MBConvBlock-217	[-1, 96, 14, 14]	0
Identity-218	[-1, 96, 14, 14]	0
Conv2dStaticSamePadding-219	[-1, 576, 14, 14]	55,296
BatchNorm2d-220	[-1, 576, 14, 14]	1,152
MemoryEfficientSwish-221	[-1, 576, 14, 14]	0
ZeroPad2d-222	[-1, 576, 18, 18]	0
Conv2dStaticSamePadding-223	[-1, 576, 14, 14]	14,400
BatchNorm2d-224	[-1, 576, 14, 14]	1,152
MemoryEfficientSwish-225	[-1, 576, 14, 14]	0
Identity-226	[-1, 576, 1, 1]	0
Conv2dStaticSamePadding-227	[-1, 24, 1, 1]	13,848
MemoryEfficientSwish-228	[-1, 24, 1, 1]	0
Identity-229	[-1, 24, 1, 1]	0
Conv2dStaticSamePadding-230	[-1, 576, 1, 1]	14,400
Identity-231	[-1, 576, 14, 14]	0
Conv2dStaticSamePadding-232	[-1, 136, 14, 14]	78,336
BatchNorm2d-233	[-1, 136, 14, 14]	272
MBConvBlock-234	[-1, 136, 14, 14]	0
Identity-235	[-1, 136, 14, 14]	0
Conv2dStaticSamePadding-236	[-1, 816, 14, 14]	110,976
BatchNorm2d-237	[-1, 816, 14, 14]	1,632
MemoryEfficientSwish-238	[-1, 816, 14, 14]	0
ZeroPad2d-239	[-1, 816, 18, 18]	0
Conv2dStaticSamePadding-240	[-1, 816, 14, 14]	20,400
BatchNorm2d-241	[-1, 816, 14, 14]	1,632
MemoryEfficientSwish-242	[-1, 816, 14, 14]	0
Identity-243	[-1, 816, 1, 1]	0
Conv2dStaticSamePadding-244	[-1, 34, 1, 1]	27,778
MemoryEfficientSwish-245	[-1, 34, 1, 1]	0
Identity-246	[-1, 34, 1, 1]	0
Conv2dStaticSamePadding-247	[-1, 816, 1, 1]	28,560
Identity-248	[-1, 816, 14, 14]	0
Conv2dStaticSamePadding-249	[-1, 136, 14, 14]	110,976
BatchNorm2d-250	[-1, 136, 14, 14]	272
MBConvBlock-251	[-1, 136, 14, 14]	0
Identity-252	[-1, 136, 14, 14]	0
Conv2dStaticSamePadding-253	[-1, 816, 14, 14]	110,976
BatchNorm2d-254	[-1, 816, 14, 14]	1,632
MemoryEfficientSwish-255	[-1, 816, 14, 14]	0
ZeroPad2d-256	[-1, 816, 18, 18]	0
Conv2dStaticSamePadding-257	[-1, 816, 14, 14]	20,400
BatchNorm2d-258	[-1, 816, 14, 14]	1,632
MemoryEfficientSwish-259	[-1, 816, 14, 14]	0
Identity-260	[-1, 816, 1, 1]	0

Conv2dStaticSamePadding-261	[-1, 34, 1, 1]	27,778
MemoryEfficientSwish-262	[-1, 34, 1, 1]	0
Identity-263	[-1, 34, 1, 1]	0
Conv2dStaticSamePadding-264	[-1, 816, 1, 1]	28,560
Identity-265	[-1, 816, 14, 14]	0
Conv2dStaticSamePadding-266	[-1, 136, 14, 14]	110,976
BatchNorm2d-267	[-1, 136, 14, 14]	272
MBConvBlock-268	[-1, 136, 14, 14]	0
Identity-269	[-1, 136, 14, 14]	0
Conv2dStaticSamePadding-270	[-1, 816, 14, 14]	110,976
BatchNorm2d-271	[-1, 816, 14, 14]	1,632
MemoryEfficientSwish-272	[-1, 816, 14, 14]	0
ZeroPad2d-273	[-1, 816, 18, 18]	0
Conv2dStaticSamePadding-274	[-1, 816, 14, 14]	20,400
BatchNorm2d-275	[-1, 816, 14, 14]	1,632
MemoryEfficientSwish-276	[-1, 816, 14, 14]	0
Identity-277	[-1, 816, 1, 1]	0
Conv2dStaticSamePadding-278	[-1, 34, 1, 1]	27,778
MemoryEfficientSwish-279	[-1, 34, 1, 1]	0
Identity-280	[-1, 34, 1, 1]	0
Conv2dStaticSamePadding-281	[-1, 816, 1, 1]	28,560
Identity-282	[-1, 816, 14, 14]	0
Conv2dStaticSamePadding-283	[-1, 136, 14, 14]	110,976
BatchNorm2d-284	[-1, 136, 14, 14]	272
MBConvBlock-285	[-1, 136, 14, 14]	0
Identity-286	[-1, 136, 14, 14]	0
Conv2dStaticSamePadding-287	[-1, 816, 14, 14]	110,976
BatchNorm2d-288	[-1, 816, 14, 14]	1,632
MemoryEfficientSwish-289	[-1, 816, 14, 14]	0
ZeroPad2d-290	[-1, 816, 18, 18]	0
Conv2dStaticSamePadding-291	[-1, 816, 14, 14]	20,400
BatchNorm2d-292	[-1, 816, 14, 14]	1,632
MemoryEfficientSwish-293	[-1, 816, 14, 14]	0
Identity-294	[-1, 816, 1, 1]	0
Conv2dStaticSamePadding-295	[-1, 34, 1, 1]	27,778
MemoryEfficientSwish-296	[-1, 34, 1, 1]	0
Identity-297	[-1, 34, 1, 1]	0
Conv2dStaticSamePadding-298	[-1, 816, 1, 1]	28,560
Identity-299	[-1, 816, 14, 14]	0
Conv2dStaticSamePadding-300	[-1, 136, 14, 14]	110,976
BatchNorm2d-301	[-1, 136, 14, 14]	272
MBConvBlock-302	[-1, 136, 14, 14]	0
Identity-303	[-1, 136, 14, 14]	0
Conv2dStaticSamePadding-304	[-1, 816, 14, 14]	110,976

BatchNorm2d-305	[-1, 816, 14, 14]	1,632
MemoryEfficientSwish-306	[-1, 816, 14, 14]	0
ZeroPad2d-307	[-1, 816, 18, 18]	0
Conv2dStaticSamePadding-308	[-1, 816, 7, 7]	20,400
BatchNorm2d-309	[-1, 816, 7, 7]	1,632
MemoryEfficientSwish-310	[-1, 816, 7, 7]	0
Identity-311	[-1, 816, 1, 1]	0
Conv2dStaticSamePadding-312	[-1, 34, 1, 1]	27,778
MemoryEfficientSwish-313	[-1, 34, 1, 1]	0
Identity-314	[-1, 34, 1, 1]	0
Conv2dStaticSamePadding-315	[-1, 816, 1, 1]	28,560
Identity-316	[-1, 816, 7, 7]	0
Conv2dStaticSamePadding-317	[-1, 232, 7, 7]	189,312
BatchNorm2d-318	[-1, 232, 7, 7]	464
MBConvBlock-319	[-1, 232, 7, 7]	0
Identity-320	[-1, 232, 7, 7]	0
Conv2dStaticSamePadding-321	[-1, 1392, 7, 7]	322,944
BatchNorm2d-322	[-1, 1392, 7, 7]	2,784
MemoryEfficientSwish-323	[-1, 1392, 7, 7]	0
ZeroPad2d-324	[-1, 1392, 11, 11]	0
Conv2dStaticSamePadding-325	[-1, 1392, 7, 7]	34,800
BatchNorm2d-326	[-1, 1392, 7, 7]	2,784
MemoryEfficientSwish-327	[-1, 1392, 7, 7]	0
Identity-328	[-1, 1392, 1, 1]	0
Conv2dStaticSamePadding-329	[-1, 58, 1, 1]	80,794
MemoryEfficientSwish-330	[-1, 58, 1, 1]	0
Identity-331	[-1, 58, 1, 1]	0
Conv2dStaticSamePadding-332	[-1, 1392, 1, 1]	82,128
Identity-333	[-1, 1392, 7, 7]	0
Conv2dStaticSamePadding-334	[-1, 232, 7, 7]	322,944
BatchNorm2d-335	[-1, 232, 7, 7]	464
MBConvBlock-336	[-1, 232, 7, 7]	0
Identity-337	[-1, 232, 7, 7]	0
Conv2dStaticSamePadding-338	[-1, 1392, 7, 7]	322,944
BatchNorm2d-339	[-1, 1392, 7, 7]	2,784
MemoryEfficientSwish-340	[-1, 1392, 7, 7]	0
ZeroPad2d-341	[-1, 1392, 11, 11]	0
Conv2dStaticSamePadding-342	[-1, 1392, 7, 7]	34,800
BatchNorm2d-343	[-1, 1392, 7, 7]	2,784
MemoryEfficientSwish-344	[-1, 1392, 7, 7]	0
Identity-345	[-1, 1392, 1, 1]	0
Conv2dStaticSamePadding-346	[-1, 58, 1, 1]	80,794
MemoryEfficientSwish-347	[-1, 58, 1, 1]	0
Identity-348	[-1, 58, 1, 1]	0

Conv2dStaticSamePadding-349	[-1, 1392, 1, 1]	82,128
Identity-350	[-1, 1392, 7, 7]	0
Conv2dStaticSamePadding-351	[-1, 232, 7, 7]	322,944
BatchNorm2d-352	[-1, 232, 7, 7]	464
MBConvBlock-353	[-1, 232, 7, 7]	0
Identity-354	[-1, 232, 7, 7]	0
Conv2dStaticSamePadding-355	[-1, 1392, 7, 7]	322,944
BatchNorm2d-356	[-1, 1392, 7, 7]	2,784
MemoryEfficientSwish-357	[-1, 1392, 7, 7]	0
ZeroPad2d-358	[-1, 1392, 11, 11]	0
Conv2dStaticSamePadding-359	[-1, 1392, 7, 7]	34,800
BatchNorm2d-360	[-1, 1392, 7, 7]	2,784
MemoryEfficientSwish-361	[-1, 1392, 7, 7]	0
Identity-362	[-1, 1392, 1, 1]	0
Conv2dStaticSamePadding-363	[-1, 58, 1, 1]	80,794
MemoryEfficientSwish-364	[-1, 58, 1, 1]	0
Identity-365	[-1, 58, 1, 1]	0
Conv2dStaticSamePadding-366	[-1, 1392, 1, 1]	82,128
Identity-367	[-1, 1392, 7, 7]	0
Conv2dStaticSamePadding-368	[-1, 232, 7, 7]	322,944
BatchNorm2d-369	[-1, 232, 7, 7]	464
MBConvBlock-370	[-1, 232, 7, 7]	0
Identity-371	[-1, 232, 7, 7]	0
Conv2dStaticSamePadding-372	[-1, 1392, 7, 7]	322,944
BatchNorm2d-373	[-1, 1392, 7, 7]	2,784
MemoryEfficientSwish-374	[-1, 1392, 7, 7]	0
ZeroPad2d-375	[-1, 1392, 11, 11]	0
Conv2dStaticSamePadding-376	[-1, 1392, 7, 7]	34,800
BatchNorm2d-377	[-1, 1392, 7, 7]	2,784
MemoryEfficientSwish-378	[-1, 1392, 7, 7]	0
Identity-379	[-1, 1392, 1, 1]	0
Conv2dStaticSamePadding-380	[-1, 58, 1, 1]	80,794
MemoryEfficientSwish-381	[-1, 58, 1, 1]	0
Identity-382	[-1, 58, 1, 1]	0
Conv2dStaticSamePadding-383	[-1, 1392, 1, 1]	82,128
Identity-384	[-1, 1392, 7, 7]	0
Conv2dStaticSamePadding-385	[-1, 232, 7, 7]	322,944
BatchNorm2d-386	[-1, 232, 7, 7]	464
MBConvBlock-387	[-1, 232, 7, 7]	0
Identity-388	[-1, 232, 7, 7]	0
Conv2dStaticSamePadding-389	[-1, 1392, 7, 7]	322,944
BatchNorm2d-390	[-1, 1392, 7, 7]	2,784
MemoryEfficientSwish-391	[-1, 1392, 7, 7]	0
ZeroPad2d-392	[-1, 1392, 11, 11]	0

Conv2dStaticSamePadding-393	[-1, 1392, 7, 7]	34,800
BatchNorm2d-394	[-1, 1392, 7, 7]	2,784
MemoryEfficientSwish-395	[-1, 1392, 7, 7]	0
Identity-396	[-1, 1392, 1, 1]	0
Conv2dStaticSamePadding-397	[-1, 58, 1, 1]	80,794
MemoryEfficientSwish-398	[-1, 58, 1, 1]	0
Identity-399	[-1, 58, 1, 1]	0
Conv2dStaticSamePadding-400	[-1, 1392, 1, 1]	82,128
Identity-401	[-1, 1392, 7, 7]	0
Conv2dStaticSamePadding-402	[-1, 232, 7, 7]	322,944
BatchNorm2d-403	[-1, 232, 7, 7]	464
MBConvBlock-404	[-1, 232, 7, 7]	0
Identity-405	[-1, 232, 7, 7]	0
Conv2dStaticSamePadding-406	[-1, 1392, 7, 7]	322,944
BatchNorm2d-407	[-1, 1392, 7, 7]	2,784
MemoryEfficientSwish-408	[-1, 1392, 7, 7]	0
ZeroPad2d-409	[-1, 1392, 9, 9]	0
Conv2dStaticSamePadding-410	[-1, 1392, 7, 7]	12,528
BatchNorm2d-411	[-1, 1392, 7, 7]	2,784
MemoryEfficientSwish-412	[-1, 1392, 7, 7]	0
Identity-413	[-1, 1392, 1, 1]	0
Conv2dStaticSamePadding-414	[-1, 58, 1, 1]	80,794
MemoryEfficientSwish-415	[-1, 58, 1, 1]	0
Identity-416	[-1, 58, 1, 1]	0
Conv2dStaticSamePadding-417	[-1, 1392, 1, 1]	82,128
Identity-418	[-1, 1392, 7, 7]	0
Conv2dStaticSamePadding-419	[-1, 384, 7, 7]	534,528
BatchNorm2d-420	[-1, 384, 7, 7]	768
MBConvBlock-421	[-1, 384, 7, 7]	0
Identity-422	[-1, 384, 7, 7]	0
Conv2dStaticSamePadding-423	[-1, 2304, 7, 7]	884,736
BatchNorm2d-424	[-1, 2304, 7, 7]	4,608
MemoryEfficientSwish-425	[-1, 2304, 7, 7]	0
ZeroPad2d-426	[-1, 2304, 9, 9]	0
Conv2dStaticSamePadding-427	[-1, 2304, 7, 7]	20,736
BatchNorm2d-428	[-1, 2304, 7, 7]	4,608
MemoryEfficientSwish-429	[-1, 2304, 7, 7]	0
Identity-430	[-1, 2304, 1, 1]	0
Conv2dStaticSamePadding-431	[-1, 96, 1, 1]	221,280
MemoryEfficientSwish-432	[-1, 96, 1, 1]	0
Identity-433	[-1, 96, 1, 1]	0
Conv2dStaticSamePadding-434	[-1, 2304, 1, 1]	223,488
Identity-435	[-1, 2304, 7, 7]	0
Conv2dStaticSamePadding-436	[-1, 384, 7, 7]	884,736

BatchNorm2d-437	[-1, 384, 7, 7]	768
MBConvBlock-438	[-1, 384, 7, 7]	0
Identity-439	[-1, 384, 7, 7]	0
Conv2dStaticSamePadding-440	[-1, 1536, 7, 7]	589,824
BatchNorm2d-441	[-1, 1536, 7, 7]	3,072
MemoryEfficientSwish-442	[-1, 1536, 7, 7]	0
AdaptiveAvgPool2d-443	[-1, 1536, 1, 1]	0
Dropout-444	[-1, 1536]	0
Linear-445	[-1, 2]	3,074

=====

=====

Total params: 10,699,306

Trainable params: 10,699,306

Non-trainable params: 0

Input size (MB): 0.57

Forward/backward pass size (MB): 410.55

Params size (MB): 40.81

Estimated Total Size (MB): 451.94

Loaded pretrained weights for efficientnet-b4

Layer (type)	Output Shape	Param #
ZeroPad2d-1	[-1, 3, 225, 225]	0
Conv2dStaticSamePadding-2	[-1, 48, 112, 112]	1,296
BatchNorm2d-3	[-1, 48, 112, 112]	96
MemoryEfficientSwish-4	[-1, 48, 112, 112]	0
ZeroPad2d-5	[-1, 48, 114, 114]	0
Conv2dStaticSamePadding-6	[-1, 48, 112, 112]	432
BatchNorm2d-7	[-1, 48, 112, 112]	96
MemoryEfficientSwish-8	[-1, 48, 112, 112]	0
Identity-9	[-1, 48, 1, 1]	0
Conv2dStaticSamePadding-10	[-1, 12, 1, 1]	588
MemoryEfficientSwish-11	[-1, 12, 1, 1]	0
Identity-12	[-1, 12, 1, 1]	0
Conv2dStaticSamePadding-13	[-1, 48, 1, 1]	624
Identity-14	[-1, 48, 112, 112]	0
Conv2dStaticSamePadding-15	[-1, 24, 112, 112]	1,152
BatchNorm2d-16	[-1, 24, 112, 112]	48
MBConvBlock-17	[-1, 24, 112, 112]	0
ZeroPad2d-18	[-1, 24, 114, 114]	0
Conv2dStaticSamePadding-19	[-1, 24, 112, 112]	216
BatchNorm2d-20	[-1, 24, 112, 112]	48
MemoryEfficientSwish-21	[-1, 24, 112, 112]	0
Identity-22	[-1, 24, 1, 1]	0
Conv2dStaticSamePadding-23	[-1, 6, 1, 1]	150
MemoryEfficientSwish-24	[-1, 6, 1, 1]	0
Identity-25	[-1, 6, 1, 1]	0
Conv2dStaticSamePadding-26	[-1, 24, 1, 1]	168
Identity-27	[-1, 24, 112, 112]	0
Conv2dStaticSamePadding-28	[-1, 24, 112, 112]	576
BatchNorm2d-29	[-1, 24, 112, 112]	48
MBConvBlock-30	[-1, 24, 112, 112]	0
Identity-31	[-1, 24, 112, 112]	0
Conv2dStaticSamePadding-32	[-1, 144, 112, 112]	3,456
BatchNorm2d-33	[-1, 144, 112, 112]	288
MemoryEfficientSwish-34	[-1, 144, 112, 112]	0
ZeroPad2d-35	[-1, 144, 113, 113]	0
Conv2dStaticSamePadding-36	[-1, 144, 56, 56]	1,296
BatchNorm2d-37	[-1, 144, 56, 56]	288
MemoryEfficientSwish-38	[-1, 144, 56, 56]	0
Identity-39	[-1, 144, 1, 1]	0
Conv2dStaticSamePadding-40	[-1, 6, 1, 1]	870

MemoryEfficientSwish-41	[-1, 6, 1, 1]	0
Identity-42	[-1, 6, 1, 1]	0
Conv2dStaticSamePadding-43	[-1, 144, 1, 1]	1,008
Identity-44	[-1, 144, 56, 56]	0
Conv2dStaticSamePadding-45	[-1, 32, 56, 56]	4,608
BatchNorm2d-46	[-1, 32, 56, 56]	64
MBConvBlock-47	[-1, 32, 56, 56]	0
Identity-48	[-1, 32, 56, 56]	0
Conv2dStaticSamePadding-49	[-1, 192, 56, 56]	6,144
BatchNorm2d-50	[-1, 192, 56, 56]	384
MemoryEfficientSwish-51	[-1, 192, 56, 56]	0
ZeroPad2d-52	[-1, 192, 58, 58]	0
Conv2dStaticSamePadding-53	[-1, 192, 56, 56]	1,728
BatchNorm2d-54	[-1, 192, 56, 56]	384
MemoryEfficientSwish-55	[-1, 192, 56, 56]	0
Identity-56	[-1, 192, 1, 1]	0
Conv2dStaticSamePadding-57	[-1, 8, 1, 1]	1,544
MemoryEfficientSwish-58	[-1, 8, 1, 1]	0
Identity-59	[-1, 8, 1, 1]	0
Conv2dStaticSamePadding-60	[-1, 192, 1, 1]	1,728
Identity-61	[-1, 192, 56, 56]	0
Conv2dStaticSamePadding-62	[-1, 32, 56, 56]	6,144
BatchNorm2d-63	[-1, 32, 56, 56]	64
MBConvBlock-64	[-1, 32, 56, 56]	0
Identity-65	[-1, 32, 56, 56]	0
Conv2dStaticSamePadding-66	[-1, 192, 56, 56]	6,144
BatchNorm2d-67	[-1, 192, 56, 56]	384
MemoryEfficientSwish-68	[-1, 192, 56, 56]	0
ZeroPad2d-69	[-1, 192, 58, 58]	0
Conv2dStaticSamePadding-70	[-1, 192, 56, 56]	1,728
BatchNorm2d-71	[-1, 192, 56, 56]	384
MemoryEfficientSwish-72	[-1, 192, 56, 56]	0
Identity-73	[-1, 192, 1, 1]	0
Conv2dStaticSamePadding-74	[-1, 8, 1, 1]	1,544
MemoryEfficientSwish-75	[-1, 8, 1, 1]	0
Identity-76	[-1, 8, 1, 1]	0
Conv2dStaticSamePadding-77	[-1, 192, 1, 1]	1,728
Identity-78	[-1, 192, 56, 56]	0
Conv2dStaticSamePadding-79	[-1, 32, 56, 56]	6,144
BatchNorm2d-80	[-1, 32, 56, 56]	64
MBConvBlock-81	[-1, 32, 56, 56]	0
Identity-82	[-1, 32, 56, 56]	0
Conv2dStaticSamePadding-83	[-1, 192, 56, 56]	6,144
BatchNorm2d-84	[-1, 192, 56, 56]	384

MemoryEfficientSwish-85	[-1, 192, 56, 56]	0
ZeroPad2d-86	[-1, 192, 58, 58]	0
Conv2dStaticSamePadding-87	[-1, 192, 56, 56]	1,728
BatchNorm2d-88	[-1, 192, 56, 56]	384
MemoryEfficientSwish-89	[-1, 192, 56, 56]	0
Identity-90	[-1, 192, 1, 1]	0
Conv2dStaticSamePadding-91	[-1, 8, 1, 1]	1,544
MemoryEfficientSwish-92	[-1, 8, 1, 1]	0
Identity-93	[-1, 8, 1, 1]	0
Conv2dStaticSamePadding-94	[-1, 192, 1, 1]	1,728
Identity-95	[-1, 192, 56, 56]	0
Conv2dStaticSamePadding-96	[-1, 32, 56, 56]	6,144
BatchNorm2d-97	[-1, 32, 56, 56]	64
MBConvBlock-98	[-1, 32, 56, 56]	0
Identity-99	[-1, 32, 56, 56]	0
Conv2dStaticSamePadding-100	[-1, 192, 56, 56]	6,144
BatchNorm2d-101	[-1, 192, 56, 56]	384
MemoryEfficientSwish-102	[-1, 192, 56, 56]	0
ZeroPad2d-103	[-1, 192, 60, 60]	0
Conv2dStaticSamePadding-104	[-1, 192, 28, 28]	4,800
BatchNorm2d-105	[-1, 192, 28, 28]	384
MemoryEfficientSwish-106	[-1, 192, 28, 28]	0
Identity-107	[-1, 192, 1, 1]	0
Conv2dStaticSamePadding-108	[-1, 8, 1, 1]	1,544
MemoryEfficientSwish-109	[-1, 8, 1, 1]	0
Identity-110	[-1, 8, 1, 1]	0
Conv2dStaticSamePadding-111	[-1, 192, 1, 1]	1,728
Identity-112	[-1, 192, 28, 28]	0
Conv2dStaticSamePadding-113	[-1, 56, 28, 28]	10,752
BatchNorm2d-114	[-1, 56, 28, 28]	112
MBConvBlock-115	[-1, 56, 28, 28]	0
Identity-116	[-1, 56, 28, 28]	0
Conv2dStaticSamePadding-117	[-1, 336, 28, 28]	18,816
BatchNorm2d-118	[-1, 336, 28, 28]	672
MemoryEfficientSwish-119	[-1, 336, 28, 28]	0
ZeroPad2d-120	[-1, 336, 32, 32]	0
Conv2dStaticSamePadding-121	[-1, 336, 28, 28]	8,400
BatchNorm2d-122	[-1, 336, 28, 28]	672
MemoryEfficientSwish-123	[-1, 336, 28, 28]	0
Identity-124	[-1, 336, 1, 1]	0
Conv2dStaticSamePadding-125	[-1, 14, 1, 1]	4,718
MemoryEfficientSwish-126	[-1, 14, 1, 1]	0
Identity-127	[-1, 14, 1, 1]	0
Conv2dStaticSamePadding-128	[-1, 336, 1, 1]	5,040

Identity-129	[-1, 336, 28, 28]	0
Conv2dStaticSamePadding-130	[-1, 56, 28, 28]	18,816
BatchNorm2d-131	[-1, 56, 28, 28]	112
MBConvBlock-132	[-1, 56, 28, 28]	0
Identity-133	[-1, 56, 28, 28]	0
Conv2dStaticSamePadding-134	[-1, 336, 28, 28]	18,816
BatchNorm2d-135	[-1, 336, 28, 28]	672
MemoryEfficientSwish-136	[-1, 336, 28, 28]	0
ZeroPad2d-137	[-1, 336, 32, 32]	0
Conv2dStaticSamePadding-138	[-1, 336, 28, 28]	8,400
BatchNorm2d-139	[-1, 336, 28, 28]	672
MemoryEfficientSwish-140	[-1, 336, 28, 28]	0
Identity-141	[-1, 336, 1, 1]	0
Conv2dStaticSamePadding-142	[-1, 14, 1, 1]	4,718
MemoryEfficientSwish-143	[-1, 14, 1, 1]	0
Identity-144	[-1, 14, 1, 1]	0
Conv2dStaticSamePadding-145	[-1, 336, 1, 1]	5,040
Identity-146	[-1, 336, 28, 28]	0
Conv2dStaticSamePadding-147	[-1, 56, 28, 28]	18,816
BatchNorm2d-148	[-1, 56, 28, 28]	112
MBConvBlock-149	[-1, 56, 28, 28]	0
Identity-150	[-1, 56, 28, 28]	0
Conv2dStaticSamePadding-151	[-1, 336, 28, 28]	18,816
BatchNorm2d-152	[-1, 336, 28, 28]	672
MemoryEfficientSwish-153	[-1, 336, 28, 28]	0
ZeroPad2d-154	[-1, 336, 32, 32]	0
Conv2dStaticSamePadding-155	[-1, 336, 28, 28]	8,400
BatchNorm2d-156	[-1, 336, 28, 28]	672
MemoryEfficientSwish-157	[-1, 336, 28, 28]	0
Identity-158	[-1, 336, 1, 1]	0
Conv2dStaticSamePadding-159	[-1, 14, 1, 1]	4,718
MemoryEfficientSwish-160	[-1, 14, 1, 1]	0
Identity-161	[-1, 14, 1, 1]	0
Conv2dStaticSamePadding-162	[-1, 336, 1, 1]	5,040
Identity-163	[-1, 336, 28, 28]	0
Conv2dStaticSamePadding-164	[-1, 56, 28, 28]	18,816
BatchNorm2d-165	[-1, 56, 28, 28]	112
MBConvBlock-166	[-1, 56, 28, 28]	0
Identity-167	[-1, 56, 28, 28]	0
Conv2dStaticSamePadding-168	[-1, 336, 28, 28]	18,816
BatchNorm2d-169	[-1, 336, 28, 28]	672
MemoryEfficientSwish-170	[-1, 336, 28, 28]	0
ZeroPad2d-171	[-1, 336, 29, 29]	0
Conv2dStaticSamePadding-172	[-1, 336, 14, 14]	3,024

BatchNorm2d-173	[-1, 336, 14, 14]	672
MemoryEfficientSwish-174	[-1, 336, 14, 14]	0
Identity-175	[-1, 336, 1, 1]	0
Conv2dStaticSamePadding-176	[-1, 14, 1, 1]	4,718
MemoryEfficientSwish-177	[-1, 14, 1, 1]	0
Identity-178	[-1, 14, 1, 1]	0
Conv2dStaticSamePadding-179	[-1, 336, 1, 1]	5,040
Identity-180	[-1, 336, 14, 14]	0
Conv2dStaticSamePadding-181	[-1, 112, 14, 14]	37,632
BatchNorm2d-182	[-1, 112, 14, 14]	224
MBConvBlock-183	[-1, 112, 14, 14]	0
Identity-184	[-1, 112, 14, 14]	0
Conv2dStaticSamePadding-185	[-1, 672, 14, 14]	75,264
BatchNorm2d-186	[-1, 672, 14, 14]	1,344
MemoryEfficientSwish-187	[-1, 672, 14, 14]	0
ZeroPad2d-188	[-1, 672, 16, 16]	0
Conv2dStaticSamePadding-189	[-1, 672, 14, 14]	6,048
BatchNorm2d-190	[-1, 672, 14, 14]	1,344
MemoryEfficientSwish-191	[-1, 672, 14, 14]	0
Identity-192	[-1, 672, 1, 1]	0
Conv2dStaticSamePadding-193	[-1, 28, 1, 1]	18,844
MemoryEfficientSwish-194	[-1, 28, 1, 1]	0
Identity-195	[-1, 28, 1, 1]	0
Conv2dStaticSamePadding-196	[-1, 672, 1, 1]	19,488
Identity-197	[-1, 672, 14, 14]	0
Conv2dStaticSamePadding-198	[-1, 112, 14, 14]	75,264
BatchNorm2d-199	[-1, 112, 14, 14]	224
MBConvBlock-200	[-1, 112, 14, 14]	0
Identity-201	[-1, 112, 14, 14]	0
Conv2dStaticSamePadding-202	[-1, 672, 14, 14]	75,264
BatchNorm2d-203	[-1, 672, 14, 14]	1,344
MemoryEfficientSwish-204	[-1, 672, 14, 14]	0
ZeroPad2d-205	[-1, 672, 16, 16]	0
Conv2dStaticSamePadding-206	[-1, 672, 14, 14]	6,048
BatchNorm2d-207	[-1, 672, 14, 14]	1,344
MemoryEfficientSwish-208	[-1, 672, 14, 14]	0
Identity-209	[-1, 672, 1, 1]	0
Conv2dStaticSamePadding-210	[-1, 28, 1, 1]	18,844
MemoryEfficientSwish-211	[-1, 28, 1, 1]	0
Identity-212	[-1, 28, 1, 1]	0
Conv2dStaticSamePadding-213	[-1, 672, 1, 1]	19,488
Identity-214	[-1, 672, 14, 14]	0
Conv2dStaticSamePadding-215	[-1, 112, 14, 14]	75,264
BatchNorm2d-216	[-1, 112, 14, 14]	224

MBConvBlock-217	[-1, 112, 14, 14]	0
Identity-218	[-1, 112, 14, 14]	0
Conv2dStaticSamePadding-219	[-1, 672, 14, 14]	75,264
BatchNorm2d-220	[-1, 672, 14, 14]	1,344
MemoryEfficientSwish-221	[-1, 672, 14, 14]	0
ZeroPad2d-222	[-1, 672, 16, 16]	0
Conv2dStaticSamePadding-223	[-1, 672, 14, 14]	6,048
BatchNorm2d-224	[-1, 672, 14, 14]	1,344
MemoryEfficientSwish-225	[-1, 672, 14, 14]	0
Identity-226	[-1, 672, 1, 1]	0
Conv2dStaticSamePadding-227	[-1, 28, 1, 1]	18,844
MemoryEfficientSwish-228	[-1, 28, 1, 1]	0
Identity-229	[-1, 28, 1, 1]	0
Conv2dStaticSamePadding-230	[-1, 672, 1, 1]	19,488
Identity-231	[-1, 672, 14, 14]	0
Conv2dStaticSamePadding-232	[-1, 112, 14, 14]	75,264
BatchNorm2d-233	[-1, 112, 14, 14]	224
MBConvBlock-234	[-1, 112, 14, 14]	0
Identity-235	[-1, 112, 14, 14]	0
Conv2dStaticSamePadding-236	[-1, 672, 14, 14]	75,264
BatchNorm2d-237	[-1, 672, 14, 14]	1,344
MemoryEfficientSwish-238	[-1, 672, 14, 14]	0
ZeroPad2d-239	[-1, 672, 16, 16]	0
Conv2dStaticSamePadding-240	[-1, 672, 14, 14]	6,048
BatchNorm2d-241	[-1, 672, 14, 14]	1,344
MemoryEfficientSwish-242	[-1, 672, 14, 14]	0
Identity-243	[-1, 672, 1, 1]	0
Conv2dStaticSamePadding-244	[-1, 28, 1, 1]	18,844
MemoryEfficientSwish-245	[-1, 28, 1, 1]	0
Identity-246	[-1, 28, 1, 1]	0
Conv2dStaticSamePadding-247	[-1, 672, 1, 1]	19,488
Identity-248	[-1, 672, 14, 14]	0
Conv2dStaticSamePadding-249	[-1, 112, 14, 14]	75,264
BatchNorm2d-250	[-1, 112, 14, 14]	224
MBConvBlock-251	[-1, 112, 14, 14]	0
Identity-252	[-1, 112, 14, 14]	0
Conv2dStaticSamePadding-253	[-1, 672, 14, 14]	75,264
BatchNorm2d-254	[-1, 672, 14, 14]	1,344
MemoryEfficientSwish-255	[-1, 672, 14, 14]	0
ZeroPad2d-256	[-1, 672, 16, 16]	0
Conv2dStaticSamePadding-257	[-1, 672, 14, 14]	6,048
BatchNorm2d-258	[-1, 672, 14, 14]	1,344
MemoryEfficientSwish-259	[-1, 672, 14, 14]	0
Identity-260	[-1, 672, 1, 1]	0

Conv2dStaticSamePadding-261	[-1, 28, 1, 1]	18,844
MemoryEfficientSwish-262	[-1, 28, 1, 1]	0
Identity-263	[-1, 28, 1, 1]	0
Conv2dStaticSamePadding-264	[-1, 672, 1, 1]	19,488
Identity-265	[-1, 672, 14, 14]	0
Conv2dStaticSamePadding-266	[-1, 112, 14, 14]	75,264
BatchNorm2d-267	[-1, 112, 14, 14]	224
MBConvBlock-268	[-1, 112, 14, 14]	0
Identity-269	[-1, 112, 14, 14]	0
Conv2dStaticSamePadding-270	[-1, 672, 14, 14]	75,264
BatchNorm2d-271	[-1, 672, 14, 14]	1,344
MemoryEfficientSwish-272	[-1, 672, 14, 14]	0
ZeroPad2d-273	[-1, 672, 18, 18]	0
Conv2dStaticSamePadding-274	[-1, 672, 14, 14]	16,800
BatchNorm2d-275	[-1, 672, 14, 14]	1,344
MemoryEfficientSwish-276	[-1, 672, 14, 14]	0
Identity-277	[-1, 672, 1, 1]	0
Conv2dStaticSamePadding-278	[-1, 28, 1, 1]	18,844
MemoryEfficientSwish-279	[-1, 28, 1, 1]	0
Identity-280	[-1, 28, 1, 1]	0
Conv2dStaticSamePadding-281	[-1, 672, 1, 1]	19,488
Identity-282	[-1, 672, 14, 14]	0
Conv2dStaticSamePadding-283	[-1, 160, 14, 14]	107,520
BatchNorm2d-284	[-1, 160, 14, 14]	320
MBConvBlock-285	[-1, 160, 14, 14]	0
Identity-286	[-1, 160, 14, 14]	0
Conv2dStaticSamePadding-287	[-1, 960, 14, 14]	153,600
BatchNorm2d-288	[-1, 960, 14, 14]	1,920
MemoryEfficientSwish-289	[-1, 960, 14, 14]	0
ZeroPad2d-290	[-1, 960, 18, 18]	0
Conv2dStaticSamePadding-291	[-1, 960, 14, 14]	24,000
BatchNorm2d-292	[-1, 960, 14, 14]	1,920
MemoryEfficientSwish-293	[-1, 960, 14, 14]	0
Identity-294	[-1, 960, 1, 1]	0
Conv2dStaticSamePadding-295	[-1, 40, 1, 1]	38,440
MemoryEfficientSwish-296	[-1, 40, 1, 1]	0
Identity-297	[-1, 40, 1, 1]	0
Conv2dStaticSamePadding-298	[-1, 960, 1, 1]	39,360
Identity-299	[-1, 960, 14, 14]	0
Conv2dStaticSamePadding-300	[-1, 160, 14, 14]	153,600
BatchNorm2d-301	[-1, 160, 14, 14]	320
MBConvBlock-302	[-1, 160, 14, 14]	0
Identity-303	[-1, 160, 14, 14]	0
Conv2dStaticSamePadding-304	[-1, 960, 14, 14]	153,600

BatchNorm2d-305	[-1, 960, 14, 14]	1,920
MemoryEfficientSwish-306	[-1, 960, 14, 14]	0
ZeroPad2d-307	[-1, 960, 18, 18]	0
Conv2dStaticSamePadding-308	[-1, 960, 14, 14]	24,000
BatchNorm2d-309	[-1, 960, 14, 14]	1,920
MemoryEfficientSwish-310	[-1, 960, 14, 14]	0
Identity-311	[-1, 960, 1, 1]	0
Conv2dStaticSamePadding-312	[-1, 40, 1, 1]	38,440
MemoryEfficientSwish-313	[-1, 40, 1, 1]	0
Identity-314	[-1, 40, 1, 1]	0
Conv2dStaticSamePadding-315	[-1, 960, 1, 1]	39,360
Identity-316	[-1, 960, 14, 14]	0
Conv2dStaticSamePadding-317	[-1, 160, 14, 14]	153,600
BatchNorm2d-318	[-1, 160, 14, 14]	320
MBConvBlock-319	[-1, 160, 14, 14]	0
Identity-320	[-1, 160, 14, 14]	0
Conv2dStaticSamePadding-321	[-1, 960, 14, 14]	153,600
BatchNorm2d-322	[-1, 960, 14, 14]	1,920
MemoryEfficientSwish-323	[-1, 960, 14, 14]	0
ZeroPad2d-324	[-1, 960, 18, 18]	0
Conv2dStaticSamePadding-325	[-1, 960, 14, 14]	24,000
BatchNorm2d-326	[-1, 960, 14, 14]	1,920
MemoryEfficientSwish-327	[-1, 960, 14, 14]	0
Identity-328	[-1, 960, 1, 1]	0
Conv2dStaticSamePadding-329	[-1, 40, 1, 1]	38,440
MemoryEfficientSwish-330	[-1, 40, 1, 1]	0
Identity-331	[-1, 40, 1, 1]	0
Conv2dStaticSamePadding-332	[-1, 960, 1, 1]	39,360
Identity-333	[-1, 960, 14, 14]	0
Conv2dStaticSamePadding-334	[-1, 160, 14, 14]	153,600
BatchNorm2d-335	[-1, 160, 14, 14]	320
MBConvBlock-336	[-1, 160, 14, 14]	0
Identity-337	[-1, 160, 14, 14]	0
Conv2dStaticSamePadding-338	[-1, 960, 14, 14]	153,600
BatchNorm2d-339	[-1, 960, 14, 14]	1,920
MemoryEfficientSwish-340	[-1, 960, 14, 14]	0
ZeroPad2d-341	[-1, 960, 18, 18]	0
Conv2dStaticSamePadding-342	[-1, 960, 14, 14]	24,000
BatchNorm2d-343	[-1, 960, 14, 14]	1,920
MemoryEfficientSwish-344	[-1, 960, 14, 14]	0
Identity-345	[-1, 960, 1, 1]	0
Conv2dStaticSamePadding-346	[-1, 40, 1, 1]	38,440
MemoryEfficientSwish-347	[-1, 40, 1, 1]	0
Identity-348	[-1, 40, 1, 1]	0

Conv2dStaticSamePadding-349	[-1, 960, 1, 1]	39,360
Identity-350	[-1, 960, 14, 14]	0
Conv2dStaticSamePadding-351	[-1, 160, 14, 14]	153,600
BatchNorm2d-352	[-1, 160, 14, 14]	320
MBConvBlock-353	[-1, 160, 14, 14]	0
Identity-354	[-1, 160, 14, 14]	0
Conv2dStaticSamePadding-355	[-1, 960, 14, 14]	153,600
BatchNorm2d-356	[-1, 960, 14, 14]	1,920
MemoryEfficientSwish-357	[-1, 960, 14, 14]	0
ZeroPad2d-358	[-1, 960, 18, 18]	0
Conv2dStaticSamePadding-359	[-1, 960, 14, 14]	24,000
BatchNorm2d-360	[-1, 960, 14, 14]	1,920
MemoryEfficientSwish-361	[-1, 960, 14, 14]	0
Identity-362	[-1, 960, 1, 1]	0
Conv2dStaticSamePadding-363	[-1, 40, 1, 1]	38,440
MemoryEfficientSwish-364	[-1, 40, 1, 1]	0
Identity-365	[-1, 40, 1, 1]	0
Conv2dStaticSamePadding-366	[-1, 960, 1, 1]	39,360
Identity-367	[-1, 960, 14, 14]	0
Conv2dStaticSamePadding-368	[-1, 160, 14, 14]	153,600
BatchNorm2d-369	[-1, 160, 14, 14]	320
MBConvBlock-370	[-1, 160, 14, 14]	0
Identity-371	[-1, 160, 14, 14]	0
Conv2dStaticSamePadding-372	[-1, 960, 14, 14]	153,600
BatchNorm2d-373	[-1, 960, 14, 14]	1,920
MemoryEfficientSwish-374	[-1, 960, 14, 14]	0
ZeroPad2d-375	[-1, 960, 17, 17]	0
Conv2dStaticSamePadding-376	[-1, 960, 7, 7]	24,000
BatchNorm2d-377	[-1, 960, 7, 7]	1,920
MemoryEfficientSwish-378	[-1, 960, 7, 7]	0
Identity-379	[-1, 960, 1, 1]	0
Conv2dStaticSamePadding-380	[-1, 40, 1, 1]	38,440
MemoryEfficientSwish-381	[-1, 40, 1, 1]	0
Identity-382	[-1, 40, 1, 1]	0
Conv2dStaticSamePadding-383	[-1, 960, 1, 1]	39,360
Identity-384	[-1, 960, 7, 7]	0
Conv2dStaticSamePadding-385	[-1, 272, 7, 7]	261,120
BatchNorm2d-386	[-1, 272, 7, 7]	544
MBConvBlock-387	[-1, 272, 7, 7]	0
Identity-388	[-1, 272, 7, 7]	0
Conv2dStaticSamePadding-389	[-1, 1632, 7, 7]	443,904
BatchNorm2d-390	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-391	[-1, 1632, 7, 7]	0
ZeroPad2d-392	[-1, 1632, 11, 11]	0

Conv2dStaticSamePadding-393	[-1, 1632, 7, 7]	40,800
BatchNorm2d-394	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-395	[-1, 1632, 7, 7]	0
Identity-396	[-1, 1632, 1, 1]	0
Conv2dStaticSamePadding-397	[-1, 68, 1, 1]	111,044
MemoryEfficientSwish-398	[-1, 68, 1, 1]	0
Identity-399	[-1, 68, 1, 1]	0
Conv2dStaticSamePadding-400	[-1, 1632, 1, 1]	112,608
Identity-401	[-1, 1632, 7, 7]	0
Conv2dStaticSamePadding-402	[-1, 272, 7, 7]	443,904
BatchNorm2d-403	[-1, 272, 7, 7]	544
MBConvBlock-404	[-1, 272, 7, 7]	0
Identity-405	[-1, 272, 7, 7]	0
Conv2dStaticSamePadding-406	[-1, 1632, 7, 7]	443,904
BatchNorm2d-407	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-408	[-1, 1632, 7, 7]	0
ZeroPad2d-409	[-1, 1632, 11, 11]	0
Conv2dStaticSamePadding-410	[-1, 1632, 7, 7]	40,800
BatchNorm2d-411	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-412	[-1, 1632, 7, 7]	0
Identity-413	[-1, 1632, 1, 1]	0
Conv2dStaticSamePadding-414	[-1, 68, 1, 1]	111,044
MemoryEfficientSwish-415	[-1, 68, 1, 1]	0
Identity-416	[-1, 68, 1, 1]	0
Conv2dStaticSamePadding-417	[-1, 1632, 1, 1]	112,608
Identity-418	[-1, 1632, 7, 7]	0
Conv2dStaticSamePadding-419	[-1, 272, 7, 7]	443,904
BatchNorm2d-420	[-1, 272, 7, 7]	544
MBConvBlock-421	[-1, 272, 7, 7]	0
Identity-422	[-1, 272, 7, 7]	0
Conv2dStaticSamePadding-423	[-1, 1632, 7, 7]	443,904
BatchNorm2d-424	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-425	[-1, 1632, 7, 7]	0
ZeroPad2d-426	[-1, 1632, 11, 11]	0
Conv2dStaticSamePadding-427	[-1, 1632, 7, 7]	40,800
BatchNorm2d-428	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-429	[-1, 1632, 7, 7]	0
Identity-430	[-1, 1632, 1, 1]	0
Conv2dStaticSamePadding-431	[-1, 68, 1, 1]	111,044
MemoryEfficientSwish-432	[-1, 68, 1, 1]	0
Identity-433	[-1, 68, 1, 1]	0
Conv2dStaticSamePadding-434	[-1, 1632, 1, 1]	112,608
Identity-435	[-1, 1632, 7, 7]	0
Conv2dStaticSamePadding-436	[-1, 272, 7, 7]	443,904

BatchNorm2d-437	[-1, 272, 7, 7]	544
MBConvBlock-438	[-1, 272, 7, 7]	0
Identity-439	[-1, 272, 7, 7]	0
Conv2dStaticSamePadding-440	[-1, 1632, 7, 7]	443,904
BatchNorm2d-441	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-442	[-1, 1632, 7, 7]	0
ZeroPad2d-443	[-1, 1632, 11, 11]	0
Conv2dStaticSamePadding-444	[-1, 1632, 7, 7]	40,800
BatchNorm2d-445	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-446	[-1, 1632, 7, 7]	0
Identity-447	[-1, 1632, 1, 1]	0
Conv2dStaticSamePadding-448	[-1, 68, 1, 1]	111,044
MemoryEfficientSwish-449	[-1, 68, 1, 1]	0
Identity-450	[-1, 68, 1, 1]	0
Conv2dStaticSamePadding-451	[-1, 1632, 1, 1]	112,608
Identity-452	[-1, 1632, 7, 7]	0
Conv2dStaticSamePadding-453	[-1, 272, 7, 7]	443,904
BatchNorm2d-454	[-1, 272, 7, 7]	544
MBConvBlock-455	[-1, 272, 7, 7]	0
Identity-456	[-1, 272, 7, 7]	0
Conv2dStaticSamePadding-457	[-1, 1632, 7, 7]	443,904
BatchNorm2d-458	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-459	[-1, 1632, 7, 7]	0
ZeroPad2d-460	[-1, 1632, 11, 11]	0
Conv2dStaticSamePadding-461	[-1, 1632, 7, 7]	40,800
BatchNorm2d-462	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-463	[-1, 1632, 7, 7]	0
Identity-464	[-1, 1632, 1, 1]	0
Conv2dStaticSamePadding-465	[-1, 68, 1, 1]	111,044
MemoryEfficientSwish-466	[-1, 68, 1, 1]	0
Identity-467	[-1, 68, 1, 1]	0
Conv2dStaticSamePadding-468	[-1, 1632, 1, 1]	112,608
Identity-469	[-1, 1632, 7, 7]	0
Conv2dStaticSamePadding-470	[-1, 272, 7, 7]	443,904
BatchNorm2d-471	[-1, 272, 7, 7]	544
MBConvBlock-472	[-1, 272, 7, 7]	0
Identity-473	[-1, 272, 7, 7]	0
Conv2dStaticSamePadding-474	[-1, 1632, 7, 7]	443,904
BatchNorm2d-475	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-476	[-1, 1632, 7, 7]	0
ZeroPad2d-477	[-1, 1632, 11, 11]	0
Conv2dStaticSamePadding-478	[-1, 1632, 7, 7]	40,800
BatchNorm2d-479	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-480	[-1, 1632, 7, 7]	0

Identity-481	[-1, 1632, 1, 1]	0
Conv2dStaticSamePadding-482	[-1, 68, 1, 1]	111,044
MemoryEfficientSwish-483	[-1, 68, 1, 1]	0
Identity-484	[-1, 68, 1, 1]	0
Conv2dStaticSamePadding-485	[-1, 1632, 1, 1]	112,608
Identity-486	[-1, 1632, 7, 7]	0
Conv2dStaticSamePadding-487	[-1, 272, 7, 7]	443,904
BatchNorm2d-488	[-1, 272, 7, 7]	544
MBConvBlock-489	[-1, 272, 7, 7]	0
Identity-490	[-1, 272, 7, 7]	0
Conv2dStaticSamePadding-491	[-1, 1632, 7, 7]	443,904
BatchNorm2d-492	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-493	[-1, 1632, 7, 7]	0
ZeroPad2d-494	[-1, 1632, 11, 11]	0
Conv2dStaticSamePadding-495	[-1, 1632, 7, 7]	40,800
BatchNorm2d-496	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-497	[-1, 1632, 7, 7]	0
Identity-498	[-1, 1632, 1, 1]	0
Conv2dStaticSamePadding-499	[-1, 68, 1, 1]	111,044
MemoryEfficientSwish-500	[-1, 68, 1, 1]	0
Identity-501	[-1, 68, 1, 1]	0
Conv2dStaticSamePadding-502	[-1, 1632, 1, 1]	112,608
Identity-503	[-1, 1632, 7, 7]	0
Conv2dStaticSamePadding-504	[-1, 272, 7, 7]	443,904
BatchNorm2d-505	[-1, 272, 7, 7]	544
MBConvBlock-506	[-1, 272, 7, 7]	0
Identity-507	[-1, 272, 7, 7]	0
Conv2dStaticSamePadding-508	[-1, 1632, 7, 7]	443,904
BatchNorm2d-509	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-510	[-1, 1632, 7, 7]	0
ZeroPad2d-511	[-1, 1632, 9, 9]	0
Conv2dStaticSamePadding-512	[-1, 1632, 7, 7]	14,688
BatchNorm2d-513	[-1, 1632, 7, 7]	3,264
MemoryEfficientSwish-514	[-1, 1632, 7, 7]	0
Identity-515	[-1, 1632, 1, 1]	0
Conv2dStaticSamePadding-516	[-1, 68, 1, 1]	111,044
MemoryEfficientSwish-517	[-1, 68, 1, 1]	0
Identity-518	[-1, 68, 1, 1]	0
Conv2dStaticSamePadding-519	[-1, 1632, 1, 1]	112,608
Identity-520	[-1, 1632, 7, 7]	0
Conv2dStaticSamePadding-521	[-1, 448, 7, 7]	731,136
BatchNorm2d-522	[-1, 448, 7, 7]	896
MBConvBlock-523	[-1, 448, 7, 7]	0
Identity-524	[-1, 448, 7, 7]	0

Conv2dStaticSamePadding-525	[-1, 2688, 7, 7]	1,204,224
BatchNorm2d-526	[-1, 2688, 7, 7]	5,376
MemoryEfficientSwish-527	[-1, 2688, 7, 7]	0
ZeroPad2d-528	[-1, 2688, 9, 9]	0
Conv2dStaticSamePadding-529	[-1, 2688, 7, 7]	24,192
BatchNorm2d-530	[-1, 2688, 7, 7]	5,376
MemoryEfficientSwish-531	[-1, 2688, 7, 7]	0
Identity-532	[-1, 2688, 1, 1]	0
Conv2dStaticSamePadding-533	[-1, 112, 1, 1]	301,168
MemoryEfficientSwish-534	[-1, 112, 1, 1]	0
Identity-535	[-1, 112, 1, 1]	0
Conv2dStaticSamePadding-536	[-1, 2688, 1, 1]	303,744
Identity-537	[-1, 2688, 7, 7]	0
Conv2dStaticSamePadding-538	[-1, 448, 7, 7]	1,204,224
BatchNorm2d-539	[-1, 448, 7, 7]	896
MBConvBlock-540	[-1, 448, 7, 7]	0
Identity-541	[-1, 448, 7, 7]	0
Conv2dStaticSamePadding-542	[-1, 1792, 7, 7]	802,816
BatchNorm2d-543	[-1, 1792, 7, 7]	3,584
MemoryEfficientSwish-544	[-1, 1792, 7, 7]	0
AdaptiveAvgPool2d-545	[-1, 1792, 1, 1]	0
Dropout-546	[-1, 1792]	0
Linear-547	[-1, 2]	3,586

=====

=====

Total params: 17,552,202
 Trainable params: 17,552,202
 Non-trainable params: 0

Input size (MB): 0.57
 Forward/backward pass size (MB): 539.32
 Params size (MB): 66.96
 Estimated Total Size (MB): 606.85

Loaded pretrained weights for vision transformer

Layer (type)	Output Shape	Param #
=====		
Conv2d-1	[-1, 768, 14, 14]	590,592
Identity-2	[-1, 196, 768]	0
PatchEmbed-3	[-1, 196, 768]	0
Dropout-4	[-1, 197, 768]	0
Identity-5	[-1, 197, 768]	0
Identity-6	[-1, 197, 768]	0
LayerNorm-7	[-1, 197, 768]	1,536
Linear-8	[-1, 197, 2304]	1,771,776
Identity-9	[-1, 12, 197, 64]	0
Identity-10	[-1, 12, 197, 64]	0
Dropout-11	[-1, 12, 197, 197]	0
Linear-12	[-1, 197, 768]	590,592
Dropout-13	[-1, 197, 768]	0
Attention-14	[-1, 197, 768]	0
Identity-15	[-1, 197, 768]	0
Identity-16	[-1, 197, 768]	0
LayerNorm-17	[-1, 197, 768]	1,536
Linear-18	[-1, 197, 3072]	2,362,368
GELU-19	[-1, 197, 3072]	0
Dropout-20	[-1, 197, 3072]	0
Identity-21	[-1, 197, 3072]	0
Linear-22	[-1, 197, 768]	2,360,064
Dropout-23	[-1, 197, 768]	0
Mlp-24	[-1, 197, 768]	0
Identity-25	[-1, 197, 768]	0
Identity-26	[-1, 197, 768]	0
Block-27	[-1, 197, 768]	0
LayerNorm-28	[-1, 197, 768]	1,536
Linear-29	[-1, 197, 2304]	1,771,776
Identity-30	[-1, 12, 197, 64]	0
Identity-31	[-1, 12, 197, 64]	0
Dropout-32	[-1, 12, 197, 197]	0
Linear-33	[-1, 197, 768]	590,592
Dropout-34	[-1, 197, 768]	0
Attention-35	[-1, 197, 768]	0
Identity-36	[-1, 197, 768]	0
Identity-37	[-1, 197, 768]	0
LayerNorm-38	[-1, 197, 768]	1,536
Linear-39	[-1, 197, 3072]	2,362,368
GELU-40	[-1, 197, 3072]	0

Dropout-41	[-1, 197, 3072]	0
Identity-42	[-1, 197, 3072]	0
Linear-43	[-1, 197, 768]	2,360,064
Dropout-44	[-1, 197, 768]	0
Mlp-45	[-1, 197, 768]	0
Identity-46	[-1, 197, 768]	0
Identity-47	[-1, 197, 768]	0
Block-48	[-1, 197, 768]	0
LayerNorm-49	[-1, 197, 768]	1,536
Linear-50	[-1, 197, 2304]	1,771,776
Identity-51	[-1, 12, 197, 64]	0
Identity-52	[-1, 12, 197, 64]	0
Dropout-53	[-1, 12, 197, 197]	0
Linear-54	[-1, 197, 768]	590,592
Dropout-55	[-1, 197, 768]	0
Attention-56	[-1, 197, 768]	0
Identity-57	[-1, 197, 768]	0
Identity-58	[-1, 197, 768]	0
LayerNorm-59	[-1, 197, 768]	1,536
Linear-60	[-1, 197, 3072]	2,362,368
GELU-61	[-1, 197, 3072]	0
Dropout-62	[-1, 197, 3072]	0
Identity-63	[-1, 197, 3072]	0
Linear-64	[-1, 197, 768]	2,360,064
Dropout-65	[-1, 197, 768]	0
Mlp-66	[-1, 197, 768]	0
Identity-67	[-1, 197, 768]	0
Identity-68	[-1, 197, 768]	0
Block-69	[-1, 197, 768]	0
LayerNorm-70	[-1, 197, 768]	1,536
Linear-71	[-1, 197, 2304]	1,771,776
Identity-72	[-1, 12, 197, 64]	0
Identity-73	[-1, 12, 197, 64]	0
Dropout-74	[-1, 12, 197, 197]	0
Linear-75	[-1, 197, 768]	590,592
Dropout-76	[-1, 197, 768]	0
Attention-77	[-1, 197, 768]	0
Identity-78	[-1, 197, 768]	0
Identity-79	[-1, 197, 768]	0
LayerNorm-80	[-1, 197, 768]	1,536
Linear-81	[-1, 197, 3072]	2,362,368
GELU-82	[-1, 197, 3072]	0
Dropout-83	[-1, 197, 3072]	0
Identity-84	[-1, 197, 3072]	0

Linear-85	[-1, 197, 768]	2,360,064
Dropout-86	[-1, 197, 768]	0
Mlp-87	[-1, 197, 768]	0
Identity-88	[-1, 197, 768]	0
Identity-89	[-1, 197, 768]	0
Block-90	[-1, 197, 768]	0
LayerNorm-91	[-1, 197, 768]	1,536
Linear-92	[-1, 197, 2304]	1,771,776
Identity-93	[-1, 12, 197, 64]	0
Identity-94	[-1, 12, 197, 64]	0
Dropout-95	[-1, 12, 197, 197]	0
Linear-96	[-1, 197, 768]	590,592
Dropout-97	[-1, 197, 768]	0
Attention-98	[-1, 197, 768]	0
Identity-99	[-1, 197, 768]	0
Identity-100	[-1, 197, 768]	0
LayerNorm-101	[-1, 197, 768]	1,536
Linear-102	[-1, 197, 3072]	2,362,368
GELU-103	[-1, 197, 3072]	0
Dropout-104	[-1, 197, 3072]	0
Identity-105	[-1, 197, 3072]	0
Linear-106	[-1, 197, 768]	2,360,064
Dropout-107	[-1, 197, 768]	0
Mlp-108	[-1, 197, 768]	0
Identity-109	[-1, 197, 768]	0
Identity-110	[-1, 197, 768]	0
Block-111	[-1, 197, 768]	0
LayerNorm-112	[-1, 197, 768]	1,536
Linear-113	[-1, 197, 2304]	1,771,776
Identity-114	[-1, 12, 197, 64]	0
Identity-115	[-1, 12, 197, 64]	0
Dropout-116	[-1, 12, 197, 197]	0
Linear-117	[-1, 197, 768]	590,592
Dropout-118	[-1, 197, 768]	0
Attention-119	[-1, 197, 768]	0
Identity-120	[-1, 197, 768]	0
Identity-121	[-1, 197, 768]	0
LayerNorm-122	[-1, 197, 768]	1,536
Linear-123	[-1, 197, 3072]	2,362,368
GELU-124	[-1, 197, 3072]	0
Dropout-125	[-1, 197, 3072]	0
Identity-126	[-1, 197, 3072]	0
Linear-127	[-1, 197, 768]	2,360,064
Dropout-128	[-1, 197, 768]	0

Mlp-129	[-1, 197, 768]	0
Identity-130	[-1, 197, 768]	0
Identity-131	[-1, 197, 768]	0
Block-132	[-1, 197, 768]	0
LayerNorm-133	[-1, 197, 768]	1,536
Linear-134	[-1, 197, 2304]	1,771,776
Identity-135	[-1, 12, 197, 64]	0
Identity-136	[-1, 12, 197, 64]	0
Dropout-137	[-1, 12, 197, 197]	0
Linear-138	[-1, 197, 768]	590,592
Dropout-139	[-1, 197, 768]	0
Attention-140	[-1, 197, 768]	0
Identity-141	[-1, 197, 768]	0
Identity-142	[-1, 197, 768]	0
LayerNorm-143	[-1, 197, 768]	1,536
Linear-144	[-1, 197, 3072]	2,362,368
GELU-145	[-1, 197, 3072]	0
Dropout-146	[-1, 197, 3072]	0
Identity-147	[-1, 197, 3072]	0
Linear-148	[-1, 197, 768]	2,360,064
Dropout-149	[-1, 197, 768]	0
Mlp-150	[-1, 197, 768]	0
Identity-151	[-1, 197, 768]	0
Identity-152	[-1, 197, 768]	0
Block-153	[-1, 197, 768]	0
LayerNorm-154	[-1, 197, 768]	1,536
Linear-155	[-1, 197, 2304]	1,771,776
Identity-156	[-1, 12, 197, 64]	0
Identity-157	[-1, 12, 197, 64]	0
Dropout-158	[-1, 12, 197, 197]	0
Linear-159	[-1, 197, 768]	590,592
Dropout-160	[-1, 197, 768]	0
Attention-161	[-1, 197, 768]	0
Identity-162	[-1, 197, 768]	0
Identity-163	[-1, 197, 768]	0
LayerNorm-164	[-1, 197, 768]	1,536
Linear-165	[-1, 197, 3072]	2,362,368
GELU-166	[-1, 197, 3072]	0
Dropout-167	[-1, 197, 3072]	0
Identity-168	[-1, 197, 3072]	0
Linear-169	[-1, 197, 768]	2,360,064
Dropout-170	[-1, 197, 768]	0
Mlp-171	[-1, 197, 768]	0
Identity-172	[-1, 197, 768]	0

Identity-173	[-1, 197, 768]	0
Block-174	[-1, 197, 768]	0
LayerNorm-175	[-1, 197, 768]	1,536
Linear-176	[-1, 197, 2304]	1,771,776
Identity-177	[-1, 12, 197, 64]	0
Identity-178	[-1, 12, 197, 64]	0
Dropout-179	[-1, 12, 197, 197]	0
Linear-180	[-1, 197, 768]	590,592
Dropout-181	[-1, 197, 768]	0
Attention-182	[-1, 197, 768]	0
Identity-183	[-1, 197, 768]	0
Identity-184	[-1, 197, 768]	0
LayerNorm-185	[-1, 197, 768]	1,536
Linear-186	[-1, 197, 3072]	2,362,368
GELU-187	[-1, 197, 3072]	0
Dropout-188	[-1, 197, 3072]	0
Identity-189	[-1, 197, 3072]	0
Linear-190	[-1, 197, 768]	2,360,064
Dropout-191	[-1, 197, 768]	0
Mlp-192	[-1, 197, 768]	0
Identity-193	[-1, 197, 768]	0
Identity-194	[-1, 197, 768]	0
Block-195	[-1, 197, 768]	0
LayerNorm-196	[-1, 197, 768]	1,536
Linear-197	[-1, 197, 2304]	1,771,776
Identity-198	[-1, 12, 197, 64]	0
Identity-199	[-1, 12, 197, 64]	0
Dropout-200	[-1, 12, 197, 197]	0
Linear-201	[-1, 197, 768]	590,592
Dropout-202	[-1, 197, 768]	0
Attention-203	[-1, 197, 768]	0
Identity-204	[-1, 197, 768]	0
Identity-205	[-1, 197, 768]	0
LayerNorm-206	[-1, 197, 768]	1,536
Linear-207	[-1, 197, 3072]	2,362,368
GELU-208	[-1, 197, 3072]	0
Dropout-209	[-1, 197, 3072]	0
Identity-210	[-1, 197, 3072]	0
Linear-211	[-1, 197, 768]	2,360,064
Dropout-212	[-1, 197, 768]	0
Mlp-213	[-1, 197, 768]	0
Identity-214	[-1, 197, 768]	0
Identity-215	[-1, 197, 768]	0
Block-216	[-1, 197, 768]	0

LayerNorm-217	[-1, 197, 768]	1,536
Linear-218	[-1, 197, 2304]	1,771,776
Identity-219	[-1, 12, 197, 64]	0
Identity-220	[-1, 12, 197, 64]	0
Dropout-221	[-1, 12, 197, 197]	0
Linear-222	[-1, 197, 768]	590,592
Dropout-223	[-1, 197, 768]	0
Attention-224	[-1, 197, 768]	0
Identity-225	[-1, 197, 768]	0
Identity-226	[-1, 197, 768]	0
LayerNorm-227	[-1, 197, 768]	1,536
Linear-228	[-1, 197, 3072]	2,362,368
GELU-229	[-1, 197, 3072]	0
Dropout-230	[-1, 197, 3072]	0
Identity-231	[-1, 197, 3072]	0
Linear-232	[-1, 197, 768]	2,360,064
Dropout-233	[-1, 197, 768]	0
Mlp-234	[-1, 197, 768]	0
Identity-235	[-1, 197, 768]	0
Identity-236	[-1, 197, 768]	0
Block-237	[-1, 197, 768]	0
LayerNorm-238	[-1, 197, 768]	1,536
Linear-239	[-1, 197, 2304]	1,771,776
Identity-240	[-1, 12, 197, 64]	0
Identity-241	[-1, 12, 197, 64]	0
Dropout-242	[-1, 12, 197, 197]	0
Linear-243	[-1, 197, 768]	590,592
Dropout-244	[-1, 197, 768]	0
Attention-245	[-1, 197, 768]	0
Identity-246	[-1, 197, 768]	0
Identity-247	[-1, 197, 768]	0
LayerNorm-248	[-1, 197, 768]	1,536
Linear-249	[-1, 197, 3072]	2,362,368
GELU-250	[-1, 197, 3072]	0
Dropout-251	[-1, 197, 3072]	0
Identity-252	[-1, 197, 3072]	0
Linear-253	[-1, 197, 768]	2,360,064
Dropout-254	[-1, 197, 768]	0
Mlp-255	[-1, 197, 768]	0
Identity-256	[-1, 197, 768]	0
Identity-257	[-1, 197, 768]	0
Block-258	[-1, 197, 768]	0
LayerNorm-259	[-1, 197, 768]	1,536
Identity-260	[-1, 768]	0

Dropout-261	[-1, 768]	0
Linear-262	[-1, 2]	1,538

=====

=====

Total params: 85,648,130

Trainable params: 85,648,130

Non-trainable params: 0

Input size (MB): 0.57

Forward/backward pass size (MB): 521.66

Params size (MB): 326.72

Estimated Total Size (MB): 848.96

Loaded pretrained weights for swin transformers

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 128, 56, 56]	6,272
LayerNorm-2	[-1, 56, 56, 128]	256
PatchEmbed-3	[-1, 56, 56, 128]	0
Identity-4	[-1, 56, 56, 128]	0
LayerNorm-5	[-1, 56, 56, 128]	256
Linear-6	[-1, 49, 384]	49,536
Softmax-7	[-1, 4, 49, 49]	0
Dropout-8	[-1, 4, 49, 49]	0
Linear-9	[-1, 49, 128]	16,512
Dropout-10	[-1, 49, 128]	0
WindowAttention-11	[-1, 49, 128]	0
Identity-12	[-1, 56, 56, 128]	0
LayerNorm-13	[-1, 3136, 128]	256
Linear-14	[-1, 3136, 512]	66,048
GELU-15	[-1, 3136, 512]	0
Dropout-16	[-1, 3136, 512]	0
Identity-17	[-1, 3136, 512]	0
Linear-18	[-1, 3136, 128]	65,664
Dropout-19	[-1, 3136, 128]	0
Mlp-20	[-1, 3136, 128]	0
Identity-21	[-1, 3136, 128]	0
SwinTransformerBlock-22	[-1, 56, 56, 128]	0
LayerNorm-23	[-1, 56, 56, 128]	256
Linear-24	[-1, 49, 384]	49,536
Softmax-25	[-1, 4, 49, 49]	0
Dropout-26	[-1, 4, 49, 49]	0
Linear-27	[-1, 49, 128]	16,512
Dropout-28	[-1, 49, 128]	0
WindowAttention-29	[-1, 49, 128]	0
DropPath-30	[-1, 56, 56, 128]	0
LayerNorm-31	[-1, 3136, 128]	256
Linear-32	[-1, 3136, 512]	66,048
GELU-33	[-1, 3136, 512]	0
Dropout-34	[-1, 3136, 512]	0
Identity-35	[-1, 3136, 512]	0
Linear-36	[-1, 3136, 128]	65,664
Dropout-37	[-1, 3136, 128]	0
Mlp-38	[-1, 3136, 128]	0
DropPath-39	[-1, 3136, 128]	0
SwinTransformerBlock-40	[-1, 56, 56, 128]	0

SwinTransformerStage-41	[-1, 56, 56, 128]	0
LayerNorm-42	[-1, 28, 28, 512]	1,024
Linear-43	[-1, 28, 28, 256]	131,072
PatchMerging-44	[-1, 28, 28, 256]	0
LayerNorm-45	[-1, 28, 28, 256]	512
Linear-46	[-1, 49, 768]	197,376
Softmax-47	[-1, 8, 49, 49]	0
Dropout-48	[-1, 8, 49, 49]	0
Linear-49	[-1, 49, 256]	65,792
Dropout-50	[-1, 49, 256]	0
WindowAttention-51	[-1, 49, 256]	0
DropPath-52	[-1, 28, 28, 256]	0
LayerNorm-53	[-1, 784, 256]	512
Linear-54	[-1, 784, 1024]	263,168
GELU-55	[-1, 784, 1024]	0
Dropout-56	[-1, 784, 1024]	0
Identity-57	[-1, 784, 1024]	0
Linear-58	[-1, 784, 256]	262,400
Dropout-59	[-1, 784, 256]	0
Mlp-60	[-1, 784, 256]	0
DropPath-61	[-1, 784, 256]	0
SwinTransformerBlock-62	[-1, 28, 28, 256]	0
LayerNorm-63	[-1, 28, 28, 256]	512
Linear-64	[-1, 49, 768]	197,376
Softmax-65	[-1, 8, 49, 49]	0
Dropout-66	[-1, 8, 49, 49]	0
Linear-67	[-1, 49, 256]	65,792
Dropout-68	[-1, 49, 256]	0
WindowAttention-69	[-1, 49, 256]	0
DropPath-70	[-1, 28, 28, 256]	0
LayerNorm-71	[-1, 784, 256]	512
Linear-72	[-1, 784, 1024]	263,168
GELU-73	[-1, 784, 1024]	0
Dropout-74	[-1, 784, 1024]	0
Identity-75	[-1, 784, 1024]	0
Linear-76	[-1, 784, 256]	262,400
Dropout-77	[-1, 784, 256]	0
Mlp-78	[-1, 784, 256]	0
DropPath-79	[-1, 784, 256]	0
SwinTransformerBlock-80	[-1, 28, 28, 256]	0
SwinTransformerStage-81	[-1, 28, 28, 256]	0
LayerNorm-82	[-1, 14, 14, 1024]	2,048
Linear-83	[-1, 14, 14, 512]	524,288
PatchMerging-84	[-1, 14, 14, 512]	0

LayerNorm-85	[-1, 14, 14, 512]	1,024
Linear-86	[-1, 49, 1536]	787,968
Softmax-87	[-1, 16, 49, 49]	0
Dropout-88	[-1, 16, 49, 49]	0
Linear-89	[-1, 49, 512]	262,656
Dropout-90	[-1, 49, 512]	0
WindowAttention-91	[-1, 49, 512]	0
DropPath-92	[-1, 14, 14, 512]	0
LayerNorm-93	[-1, 196, 512]	1,024
Linear-94	[-1, 196, 2048]	1,050,624
GELU-95	[-1, 196, 2048]	0
Dropout-96	[-1, 196, 2048]	0
Identity-97	[-1, 196, 2048]	0
Linear-98	[-1, 196, 512]	1,049,088
Dropout-99	[-1, 196, 512]	0
Mlp-100	[-1, 196, 512]	0
DropPath-101	[-1, 196, 512]	0
SwinTransformerBlock-102	[-1, 14, 14, 512]	0
LayerNorm-103	[-1, 14, 14, 512]	1,024
Linear-104	[-1, 49, 1536]	787,968
Softmax-105	[-1, 16, 49, 49]	0
Dropout-106	[-1, 16, 49, 49]	0
Linear-107	[-1, 49, 512]	262,656
Dropout-108	[-1, 49, 512]	0
WindowAttention-109	[-1, 49, 512]	0
DropPath-110	[-1, 14, 14, 512]	0
LayerNorm-111	[-1, 196, 512]	1,024
Linear-112	[-1, 196, 2048]	1,050,624
GELU-113	[-1, 196, 2048]	0
Dropout-114	[-1, 196, 2048]	0
Identity-115	[-1, 196, 2048]	0
Linear-116	[-1, 196, 512]	1,049,088
Dropout-117	[-1, 196, 512]	0
Mlp-118	[-1, 196, 512]	0
DropPath-119	[-1, 196, 512]	0
SwinTransformerBlock-120	[-1, 14, 14, 512]	0
LayerNorm-121	[-1, 14, 14, 512]	1,024
Linear-122	[-1, 49, 1536]	787,968
Softmax-123	[-1, 16, 49, 49]	0
Dropout-124	[-1, 16, 49, 49]	0
Linear-125	[-1, 49, 512]	262,656
Dropout-126	[-1, 49, 512]	0
WindowAttention-127	[-1, 49, 512]	0
DropPath-128	[-1, 14, 14, 512]	0

LayerNorm-129	[-1, 196, 512]	1,024
Linear-130	[-1, 196, 2048]	1,050,624
GELU-131	[-1, 196, 2048]	0
Dropout-132	[-1, 196, 2048]	0
Identity-133	[-1, 196, 2048]	0
Linear-134	[-1, 196, 512]	1,049,088
Dropout-135	[-1, 196, 512]	0
Mlp-136	[-1, 196, 512]	0
DropPath-137	[-1, 196, 512]	0
SwinTransformerBlock-138	[-1, 14, 14, 512]	0
LayerNorm-139	[-1, 14, 14, 512]	1,024
Linear-140	[-1, 49, 1536]	787,968
Softmax-141	[-1, 16, 49, 49]	0
Dropout-142	[-1, 16, 49, 49]	0
Linear-143	[-1, 49, 512]	262,656
Dropout-144	[-1, 49, 512]	0
WindowAttention-145	[-1, 49, 512]	0
DropPath-146	[-1, 14, 14, 512]	0
LayerNorm-147	[-1, 196, 512]	1,024
Linear-148	[-1, 196, 2048]	1,050,624
GELU-149	[-1, 196, 2048]	0
Dropout-150	[-1, 196, 2048]	0
Identity-151	[-1, 196, 2048]	0
Linear-152	[-1, 196, 512]	1,049,088
Dropout-153	[-1, 196, 512]	0
Mlp-154	[-1, 196, 512]	0
DropPath-155	[-1, 196, 512]	0
SwinTransformerBlock-156	[-1, 14, 14, 512]	0
LayerNorm-157	[-1, 14, 14, 512]	1,024
Linear-158	[-1, 49, 1536]	787,968
Softmax-159	[-1, 16, 49, 49]	0
Dropout-160	[-1, 16, 49, 49]	0
Linear-161	[-1, 49, 512]	262,656
Dropout-162	[-1, 49, 512]	0
WindowAttention-163	[-1, 49, 512]	0
DropPath-164	[-1, 14, 14, 512]	0
LayerNorm-165	[-1, 196, 512]	1,024
Linear-166	[-1, 196, 2048]	1,050,624
GELU-167	[-1, 196, 2048]	0
Dropout-168	[-1, 196, 2048]	0
Identity-169	[-1, 196, 2048]	0
Linear-170	[-1, 196, 512]	1,049,088
Dropout-171	[-1, 196, 512]	0
Mlp-172	[-1, 196, 512]	0

DropPath-173	[-1, 196, 512]	0
SwinTransformerBlock-174	[-1, 14, 14, 512]	0
LayerNorm-175	[-1, 14, 14, 512]	1,024
Linear-176	[-1, 49, 1536]	787,968
Softmax-177	[-1, 16, 49, 49]	0
Dropout-178	[-1, 16, 49, 49]	0
Linear-179	[-1, 49, 512]	262,656
Dropout-180	[-1, 49, 512]	0
WindowAttention-181	[-1, 49, 512]	0
DropPath-182	[-1, 14, 14, 512]	0
LayerNorm-183	[-1, 196, 512]	1,024
Linear-184	[-1, 196, 2048]	1,050,624
GELU-185	[-1, 196, 2048]	0
Dropout-186	[-1, 196, 2048]	0
Identity-187	[-1, 196, 2048]	0
Linear-188	[-1, 196, 512]	1,049,088
Dropout-189	[-1, 196, 512]	0
Mlp-190	[-1, 196, 512]	0
DropPath-191	[-1, 196, 512]	0
SwinTransformerBlock-192	[-1, 14, 14, 512]	0
LayerNorm-193	[-1, 14, 14, 512]	1,024
Linear-194	[-1, 49, 1536]	787,968
Softmax-195	[-1, 16, 49, 49]	0
Dropout-196	[-1, 16, 49, 49]	0
Linear-197	[-1, 49, 512]	262,656
Dropout-198	[-1, 49, 512]	0
WindowAttention-199	[-1, 49, 512]	0
DropPath-200	[-1, 14, 14, 512]	0
LayerNorm-201	[-1, 196, 512]	1,024
Linear-202	[-1, 196, 2048]	1,050,624
GELU-203	[-1, 196, 2048]	0
Dropout-204	[-1, 196, 2048]	0
Identity-205	[-1, 196, 2048]	0
Linear-206	[-1, 196, 512]	1,049,088
Dropout-207	[-1, 196, 512]	0
Mlp-208	[-1, 196, 512]	0
DropPath-209	[-1, 196, 512]	0
SwinTransformerBlock-210	[-1, 14, 14, 512]	0
LayerNorm-211	[-1, 14, 14, 512]	1,024
Linear-212	[-1, 49, 1536]	787,968
Softmax-213	[-1, 16, 49, 49]	0
Dropout-214	[-1, 16, 49, 49]	0
Linear-215	[-1, 49, 512]	262,656
Dropout-216	[-1, 49, 512]	0

WindowAttention-217	[-1, 49, 512]	0
DropPath-218	[-1, 14, 14, 512]	0
LayerNorm-219	[-1, 196, 512]	1,024
Linear-220	[-1, 196, 2048]	1,050,624
GELU-221	[-1, 196, 2048]	0
Dropout-222	[-1, 196, 2048]	0
Identity-223	[-1, 196, 2048]	0
Linear-224	[-1, 196, 512]	1,049,088
Dropout-225	[-1, 196, 512]	0
Mlp-226	[-1, 196, 512]	0
DropPath-227	[-1, 196, 512]	0
SwinTransformerBlock-228	[-1, 14, 14, 512]	0
LayerNorm-229	[-1, 14, 14, 512]	1,024
Linear-230	[-1, 49, 1536]	787,968
Softmax-231	[-1, 16, 49, 49]	0
Dropout-232	[-1, 16, 49, 49]	0
Linear-233	[-1, 49, 512]	262,656
Dropout-234	[-1, 49, 512]	0
WindowAttention-235	[-1, 49, 512]	0
DropPath-236	[-1, 14, 14, 512]	0
LayerNorm-237	[-1, 196, 512]	1,024
Linear-238	[-1, 196, 2048]	1,050,624
GELU-239	[-1, 196, 2048]	0
Dropout-240	[-1, 196, 2048]	0
Identity-241	[-1, 196, 2048]	0
Linear-242	[-1, 196, 512]	1,049,088
Dropout-243	[-1, 196, 512]	0
Mlp-244	[-1, 196, 512]	0
DropPath-245	[-1, 196, 512]	0
SwinTransformerBlock-246	[-1, 14, 14, 512]	0
LayerNorm-247	[-1, 14, 14, 512]	1,024
Linear-248	[-1, 49, 1536]	787,968
Softmax-249	[-1, 16, 49, 49]	0
Dropout-250	[-1, 16, 49, 49]	0
Linear-251	[-1, 49, 512]	262,656
Dropout-252	[-1, 49, 512]	0
WindowAttention-253	[-1, 49, 512]	0
DropPath-254	[-1, 14, 14, 512]	0
LayerNorm-255	[-1, 196, 512]	1,024
Linear-256	[-1, 196, 2048]	1,050,624
GELU-257	[-1, 196, 2048]	0
Dropout-258	[-1, 196, 2048]	0
Identity-259	[-1, 196, 2048]	0
Linear-260	[-1, 196, 512]	1,049,088

Dropout-261	[-1, 196, 512]	0
Mlp-262	[-1, 196, 512]	0
DropPath-263	[-1, 196, 512]	0
SwinTransformerBlock-264	[-1, 14, 14, 512]	0
LayerNorm-265	[-1, 14, 14, 512]	1,024
Linear-266	[-1, 49, 1536]	787,968
Softmax-267	[-1, 16, 49, 49]	0
Dropout-268	[-1, 16, 49, 49]	0
Linear-269	[-1, 49, 512]	262,656
Dropout-270	[-1, 49, 512]	0
WindowAttention-271	[-1, 49, 512]	0
DropPath-272	[-1, 14, 14, 512]	0
LayerNorm-273	[-1, 196, 512]	1,024
Linear-274	[-1, 196, 2048]	1,050,624
GELU-275	[-1, 196, 2048]	0
Dropout-276	[-1, 196, 2048]	0
Identity-277	[-1, 196, 2048]	0
Linear-278	[-1, 196, 512]	1,049,088
Dropout-279	[-1, 196, 512]	0
Mlp-280	[-1, 196, 512]	0
DropPath-281	[-1, 196, 512]	0
SwinTransformerBlock-282	[-1, 14, 14, 512]	0
LayerNorm-283	[-1, 14, 14, 512]	1,024
Linear-284	[-1, 49, 1536]	787,968
Softmax-285	[-1, 16, 49, 49]	0
Dropout-286	[-1, 16, 49, 49]	0
Linear-287	[-1, 49, 512]	262,656
Dropout-288	[-1, 49, 512]	0
WindowAttention-289	[-1, 49, 512]	0
DropPath-290	[-1, 14, 14, 512]	0
LayerNorm-291	[-1, 196, 512]	1,024
Linear-292	[-1, 196, 2048]	1,050,624
GELU-293	[-1, 196, 2048]	0
Dropout-294	[-1, 196, 2048]	0
Identity-295	[-1, 196, 2048]	0
Linear-296	[-1, 196, 512]	1,049,088
Dropout-297	[-1, 196, 512]	0
Mlp-298	[-1, 196, 512]	0
DropPath-299	[-1, 196, 512]	0
SwinTransformerBlock-300	[-1, 14, 14, 512]	0
LayerNorm-301	[-1, 14, 14, 512]	1,024
Linear-302	[-1, 49, 1536]	787,968
Softmax-303	[-1, 16, 49, 49]	0
Dropout-304	[-1, 16, 49, 49]	0

Linear-305	[-1, 49, 512]	262,656
Dropout-306	[-1, 49, 512]	0
WindowAttention-307	[-1, 49, 512]	0
DropPath-308	[-1, 14, 14, 512]	0
LayerNorm-309	[-1, 196, 512]	1,024
Linear-310	[-1, 196, 2048]	1,050,624
GELU-311	[-1, 196, 2048]	0
Dropout-312	[-1, 196, 2048]	0
Identity-313	[-1, 196, 2048]	0
Linear-314	[-1, 196, 512]	1,049,088
Dropout-315	[-1, 196, 512]	0
Mlp-316	[-1, 196, 512]	0
DropPath-317	[-1, 196, 512]	0
SwinTransformerBlock-318	[-1, 14, 14, 512]	0
LayerNorm-319	[-1, 14, 14, 512]	1,024
Linear-320	[-1, 49, 1536]	787,968
Softmax-321	[-1, 16, 49, 49]	0
Dropout-322	[-1, 16, 49, 49]	0
Linear-323	[-1, 49, 512]	262,656
Dropout-324	[-1, 49, 512]	0
WindowAttention-325	[-1, 49, 512]	0
DropPath-326	[-1, 14, 14, 512]	0
LayerNorm-327	[-1, 196, 512]	1,024
Linear-328	[-1, 196, 2048]	1,050,624
GELU-329	[-1, 196, 2048]	0
Dropout-330	[-1, 196, 2048]	0
Identity-331	[-1, 196, 2048]	0
Linear-332	[-1, 196, 512]	1,049,088
Dropout-333	[-1, 196, 512]	0
Mlp-334	[-1, 196, 512]	0
DropPath-335	[-1, 196, 512]	0
SwinTransformerBlock-336	[-1, 14, 14, 512]	0
LayerNorm-337	[-1, 14, 14, 512]	1,024
Linear-338	[-1, 49, 1536]	787,968
Softmax-339	[-1, 16, 49, 49]	0
Dropout-340	[-1, 16, 49, 49]	0
Linear-341	[-1, 49, 512]	262,656
Dropout-342	[-1, 49, 512]	0
WindowAttention-343	[-1, 49, 512]	0
DropPath-344	[-1, 14, 14, 512]	0
LayerNorm-345	[-1, 196, 512]	1,024
Linear-346	[-1, 196, 2048]	1,050,624
GELU-347	[-1, 196, 2048]	0
Dropout-348	[-1, 196, 2048]	0

Identity-349	[-1, 196, 2048]	0
Linear-350	[-1, 196, 512]	1,049,088
Dropout-351	[-1, 196, 512]	0
Mlp-352	[-1, 196, 512]	0
DropPath-353	[-1, 196, 512]	0
SwinTransformerBlock-354	[-1, 14, 14, 512]	0
LayerNorm-355	[-1, 14, 14, 512]	1,024
Linear-356	[-1, 49, 1536]	787,968
Softmax-357	[-1, 16, 49, 49]	0
Dropout-358	[-1, 16, 49, 49]	0
Linear-359	[-1, 49, 512]	262,656
Dropout-360	[-1, 49, 512]	0
WindowAttention-361	[-1, 49, 512]	0
DropPath-362	[-1, 14, 14, 512]	0
LayerNorm-363	[-1, 196, 512]	1,024
Linear-364	[-1, 196, 2048]	1,050,624
GELU-365	[-1, 196, 2048]	0
Dropout-366	[-1, 196, 2048]	0
Identity-367	[-1, 196, 2048]	0
Linear-368	[-1, 196, 512]	1,049,088
Dropout-369	[-1, 196, 512]	0
Mlp-370	[-1, 196, 512]	0
DropPath-371	[-1, 196, 512]	0
SwinTransformerBlock-372	[-1, 14, 14, 512]	0
LayerNorm-373	[-1, 14, 14, 512]	1,024
Linear-374	[-1, 49, 1536]	787,968
Softmax-375	[-1, 16, 49, 49]	0
Dropout-376	[-1, 16, 49, 49]	0
Linear-377	[-1, 49, 512]	262,656
Dropout-378	[-1, 49, 512]	0
WindowAttention-379	[-1, 49, 512]	0
DropPath-380	[-1, 14, 14, 512]	0
LayerNorm-381	[-1, 196, 512]	1,024
Linear-382	[-1, 196, 2048]	1,050,624
GELU-383	[-1, 196, 2048]	0
Dropout-384	[-1, 196, 2048]	0
Identity-385	[-1, 196, 2048]	0
Linear-386	[-1, 196, 512]	1,049,088
Dropout-387	[-1, 196, 512]	0
Mlp-388	[-1, 196, 512]	0
DropPath-389	[-1, 196, 512]	0
SwinTransformerBlock-390	[-1, 14, 14, 512]	0
LayerNorm-391	[-1, 14, 14, 512]	1,024
Linear-392	[-1, 49, 1536]	787,968

Softmax-393	[-1, 16, 49, 49]	0
Dropout-394	[-1, 16, 49, 49]	0
Linear-395	[-1, 49, 512]	262,656
Dropout-396	[-1, 49, 512]	0
WindowAttention-397	[-1, 49, 512]	0
DropPath-398	[-1, 14, 14, 512]	0
LayerNorm-399	[-1, 196, 512]	1,024
Linear-400	[-1, 196, 2048]	1,050,624
GELU-401	[-1, 196, 2048]	0
Dropout-402	[-1, 196, 2048]	0
Identity-403	[-1, 196, 2048]	0
Linear-404	[-1, 196, 512]	1,049,088
Dropout-405	[-1, 196, 512]	0
Mlp-406	[-1, 196, 512]	0
DropPath-407	[-1, 196, 512]	0
SwinTransformerBlock-408	[-1, 14, 14, 512]	0
SwinTransformerStage-409	[-1, 14, 14, 512]	0
LayerNorm-410	[-1, 7, 7, 2048]	4,096
Linear-411	[-1, 7, 7, 1024]	2,097,152
PatchMerging-412	[-1, 7, 7, 1024]	0
LayerNorm-413	[-1, 7, 7, 1024]	2,048
Linear-414	[-1, 49, 3072]	3,148,800
Softmax-415	[-1, 32, 49, 49]	0
Dropout-416	[-1, 32, 49, 49]	0
Linear-417	[-1, 49, 1024]	1,049,600
Dropout-418	[-1, 49, 1024]	0
WindowAttention-419	[-1, 49, 1024]	0
DropPath-420	[-1, 7, 7, 1024]	0
LayerNorm-421	[-1, 49, 1024]	2,048
Linear-422	[-1, 49, 4096]	4,198,400
GELU-423	[-1, 49, 4096]	0
Dropout-424	[-1, 49, 4096]	0
Identity-425	[-1, 49, 4096]	0
Linear-426	[-1, 49, 1024]	4,195,328
Dropout-427	[-1, 49, 1024]	0
Mlp-428	[-1, 49, 1024]	0
DropPath-429	[-1, 49, 1024]	0
SwinTransformerBlock-430	[-1, 7, 7, 1024]	0
LayerNorm-431	[-1, 7, 7, 1024]	2,048
Linear-432	[-1, 49, 3072]	3,148,800
Softmax-433	[-1, 32, 49, 49]	0
Dropout-434	[-1, 32, 49, 49]	0
Linear-435	[-1, 49, 1024]	1,049,600
Dropout-436	[-1, 49, 1024]	0

WindowAttention-437	[-1, 49, 1024]	0
DropPath-438	[-1, 7, 7, 1024]	0
LayerNorm-439	[-1, 49, 1024]	2,048
Linear-440	[-1, 49, 4096]	4,198,400
GELU-441	[-1, 49, 4096]	0
Dropout-442	[-1, 49, 4096]	0
Identity-443	[-1, 49, 4096]	0
Linear-444	[-1, 49, 1024]	4,195,328
Dropout-445	[-1, 49, 1024]	0
Mlp-446	[-1, 49, 1024]	0
DropPath-447	[-1, 49, 1024]	0
SwinTransformerBlock-448	[-1, 7, 7, 1024]	0
SwinTransformerStage-449	[-1, 7, 7, 1024]	0
LayerNorm-450	[-1, 7, 7, 1024]	2,048
FastAdaptiveAvgPool-451	[-1, 1024]	0
Identity-452	[-1, 1024]	0
SelectAdaptivePool2d-453	[-1, 1024]	0
Dropout-454	[-1, 1024]	0
Linear-455	[-1, 2]	2,050
Identity-456	[-1, 2]	0
ClassifierHead-457	[-1, 2]	0

=====

=====

Total params: 86,681,730

Trainable params: 86,681,730

Non-trainable params: 0

Input size (MB): 0.57

Forward/backward pass size (MB): 639.51

Params size (MB): 330.66

Estimated Total Size (MB): 970.75

Best hyperparameters for the ensemble model (Random Forest):

```
{'bootstrap': True, 'max_depth': 20, 'min_samples_leaf': 1, 'min_samples_split': 2,  
'n_estimators': 100}
```