2021 DLCV
HW1
R09942171 電信所碩一 黃繼綸

Problem1

1. Network architecture(Resnext101_32x8d)

Layer (type:depth-idx)	Output Shape	Param #
├─Conv2d: 1-1	[-1, 64, 112, 112]	9,408
├─BatchNorm2d: 1-2	[-1, 64, 112, 112]	128
├─ReLU: 1-3	[-1, 64, 112, 112]	--
├─MaxPool2d: 1-4	[-1, 64, 56, 56]	--
├─Sequential: 1-5	[-1, 256, 56, 56]	--
│ └─Bottleneck: 2-1	[-1, 256, 56, 56]	--
│ │ └─Conv2d: 3-1	[-1, 256, 56, 56]	16,384
│ │ └─BatchNorm2d: 3-2	[-1, 256, 56, 56]	512
│ │ └─ReLU: 3-3	[-1, 256, 56, 56]	--
│ │ └─Conv2d: 3-4	[-1, 256, 56, 56]	18,432
│ │ └─BatchNorm2d: 3-5	[-1, 256, 56, 56]	512
│ │ └─ReLU: 3-6	[-1, 256, 56, 56]	--
│ │ └─Conv2d: 3-7	[-1, 256, 56, 56]	65,536
│ │ └─BatchNorm2d: 3-8	[-1, 256, 56, 56]	512
│ │ └─Sequential: 3-9	[-1, 256, 56, 56]	16,896
│ │ └─ReLU: 3-10	[-1, 256, 56, 56]	--
│ └─Bottleneck: 2-2	[-1, 256, 56, 56]	--
│ │ └─Conv2d: 3-11	[-1, 256, 56, 56]	65,536
│ │ └─BatchNorm2d: 3-12	[-1, 256, 56, 56]	512
│ │ └─ReLU: 3-13	[-1, 256, 56, 56]	--
│ │ └─Conv2d: 3-14	[-1, 256, 56, 56]	18,432
│ │ └─BatchNorm2d: 3-15	[-1, 256, 56, 56]	512
│ │ └─ReLU: 3-16	[-1, 256, 56, 56]	--
│ │ └─Conv2d: 3-17	[-1, 256, 56, 56]	65,536
│ │ └─BatchNorm2d: 3-18	[-1, 256, 56, 56]	512
│ │ └─ReLU: 3-19	[-1, 256, 56, 56]	--
│ └─Bottleneck: 2-3	[-1, 256, 56, 56]	--
│ │ └─Conv2d: 3-20	[-1, 256, 56, 56]	65,536
│ │ └─BatchNorm2d: 3-21	[-1, 256, 56, 56]	512
│ │ └─ReLU: 3-22	[-1, 256, 56, 56]	--
│ │ └─Conv2d: 3-23	[-1, 256, 56, 56]	18,432
│ │ └─BatchNorm2d: 3-24	[-1, 256, 56, 56]	512
│ │ └─ReLU: 3-25	[-1, 256, 56, 56]	--
│ │ └─Conv2d: 3-26	[-1, 256, 56, 56]	65,536
│ │ └─BatchNorm2d: 3-27	[-1, 256, 56, 56]	512
│ │ └─ReLU: 3-28	[-1, 256, 56, 56]	--
├─Sequential: 1-6	[-1, 512, 28, 28]	--
│ └─Bottleneck: 2-4	[-1, 512, 28, 28]	--
│ │ └─Conv2d: 3-29	[-1, 512, 56, 56]	131,072
│ │ └─BatchNorm2d: 3-30	[-1, 512, 56, 56]	1,024
│ │ └─ReLU: 3-31	[-1, 512, 56, 56]	--
│ │ └─Conv2d: 3-32	[-1, 512, 28, 28]	73,728
│ │ └─BatchNorm2d: 3-33	[-1, 512, 28, 28]	1,024
│ │ └─ReLU: 3-34	[-1, 512, 28, 28]	--
│ │ └─Conv2d: 3-35	[-1, 512, 28, 28]	262,144
│ │ └─BatchNorm2d: 3-36	[-1, 512, 28, 28]	1,024
│ │ └─Sequential: 3-37	[-1, 512, 28, 28]	132,096
│ │ └─ReLU: 3-38	[-1, 512, 28, 28]	--

```
|    └─Bottleneck: 2-5                   [-1, 512, 28, 28]         --
|    |    └─Conv2d: 3-39                 [-1, 512, 28, 28]         262,144
|    |    └─BatchNorm2d: 3-40            [-1, 512, 28, 28]         1,024
|    |    └─ReLU: 3-41                   [-1, 512, 28, 28]         --
|    |    └─Conv2d: 3-42                 [-1, 512, 28, 28]         73,728
|    |    └─BatchNorm2d: 3-43            [-1, 512, 28, 28]         1,024
|    |    └─ReLU: 3-44                   [-1, 512, 28, 28]         --
|    |    └─Conv2d: 3-45                 [-1, 512, 28, 28]         262,144
|    |    └─BatchNorm2d: 3-46            [-1, 512, 28, 28]         1,024
|    |    └─ReLU: 3-47                   [-1, 512, 28, 28]         --
|    └─Bottleneck: 2-6                   [-1, 512, 28, 28]         --
|    |    └─Conv2d: 3-48                 [-1, 512, 28, 28]         262,144
|    |    └─BatchNorm2d: 3-49            [-1, 512, 28, 28]         1,024
|    |    └─ReLU: 3-50                   [-1, 512, 28, 28]         --
|    |    └─Conv2d: 3-51                 [-1, 512, 28, 28]         73,728
|    |    └─BatchNorm2d: 3-52            [-1, 512, 28, 28]         1,024
|    |    └─ReLU: 3-53                   [-1, 512, 28, 28]         --
|    |    └─Conv2d: 3-54                 [-1, 512, 28, 28]         262,144
|    |    └─BatchNorm2d: 3-55            [-1, 512, 28, 28]         1,024
|    |    └─ReLU: 3-56                   [-1, 512, 28, 28]         --
|    └─Bottleneck: 2-7                   [-1, 512, 28, 28]         --
|    |    └─Conv2d: 3-57                 [-1, 512, 28, 28]         262,144
|    |    └─BatchNorm2d: 3-58            [-1, 512, 28, 28]         1,024
|    |    └─ReLU: 3-59                   [-1, 512, 28, 28]         --
|    |    └─Conv2d: 3-60                 [-1, 512, 28, 28]         73,728
|    |    └─BatchNorm2d: 3-61            [-1, 512, 28, 28]         1,024
|    |    └─ReLU: 3-62                   [-1, 512, 28, 28]         --
|    |    └─Conv2d: 3-63                 [-1, 512, 28, 28]         262,144
|    |    └─BatchNorm2d: 3-64            [-1, 512, 28, 28]         1,024
|    |    └─ReLU: 3-65                   [-1, 512, 28, 28]         --
├─Sequential: 1-7                        [-1, 1024, 14, 14]        --
|    └─Bottleneck: 2-8                   [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-66                 [-1, 1024, 28, 28]        524,288
|    |    └─BatchNorm2d: 3-67            [-1, 1024, 28, 28]        2,048
|    |    └─ReLU: 3-68                   [-1, 1024, 28, 28]        --
|    |    └─Conv2d: 3-69                 [-1, 1024, 14, 14]        294,912
|    |    └─BatchNorm2d: 3-70            [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-71                   [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-72                 [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-73            [-1, 1024, 14, 14]        2,048
|    |    └─Sequential: 3-74             [-1, 1024, 14, 14]        526,336
|    |    └─ReLU: 3-75                   [-1, 1024, 14, 14]        --
|    └─Bottleneck: 2-9                   [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-76                 [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-77            [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-78                   [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-79                 [-1, 1024, 14, 14]        294,912
|    |    └─BatchNorm2d: 3-80            [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-81                   [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-82                 [-1, 1024, 14, 14]        1,048,576
```

```
|    |    └─Conv2d: 3-91                 [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-92            [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-93                   [-1, 1024, 14, 14]        --
|    └─Bottleneck: 2-11                  [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-94                 [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-95            [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-96                   [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-97                 [-1, 1024, 14, 14]        294,912
|    |    └─BatchNorm2d: 3-98            [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-99                   [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-100                [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-101           [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-102                  [-1, 1024, 14, 14]        --
|    └─Bottleneck: 2-12                  [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-103                [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-104           [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-105                  [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-106                [-1, 1024, 14, 14]        294,912
|    |    └─BatchNorm2d: 3-107           [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-108                  [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-109                [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-110           [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-111                  [-1, 1024, 14, 14]        --
|    └─Bottleneck: 2-13                  [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-112                [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-113           [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-114                  [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-115                [-1, 1024, 14, 14]        294,912
|    |    └─BatchNorm2d: 3-116           [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-117                  [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-118                [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-119           [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-120                  [-1, 1024, 14, 14]        --
|    └─Bottleneck: 2-14                  [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-121                [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-122           [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-123                  [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-124                [-1, 1024, 14, 14]        294,912
|    |    └─BatchNorm2d: 3-125           [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-126                  [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-127                [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-128           [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-129                  [-1, 1024, 14, 14]        --
|    └─Bottleneck: 2-15                  [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-130                [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-131           [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-132                  [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-133                [-1, 1024, 14, 14]        294,912
|    |    └─BatchNorm2d: 3-134           [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-135                  [-1, 1024, 14, 14]        --
```

		└─Conv2d: 3-136	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-137	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-138	[-1, 1024, 14, 14]	--
	└─Bottleneck: 2-16		[-1, 1024, 14, 14]	--
		└─Conv2d: 3-139	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-140	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-141	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-142	[-1, 1024, 14, 14]	294,912
		└─BatchNorm2d: 3-143	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-144	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-145	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-146	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-147	[-1, 1024, 14, 14]	--
	└─Bottleneck: 2-17		[-1, 1024, 14, 14]	--
		└─Conv2d: 3-148	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-149	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-150	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-151	[-1, 1024, 14, 14]	294,912
		└─BatchNorm2d: 3-152	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-153	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-154	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-155	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-156	[-1, 1024, 14, 14]	--
	└─Bottleneck: 2-18		[-1, 1024, 14, 14]	--
		└─Conv2d: 3-157	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-158	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-159	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-160	[-1, 1024, 14, 14]	294,912
		└─BatchNorm2d: 3-161	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-162	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-163	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-164	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-165	[-1, 1024, 14, 14]	--
	└─Bottleneck: 2-19		[-1, 1024, 14, 14]	--
		└─Conv2d: 3-166	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-167	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-168	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-169	[-1, 1024, 14, 14]	294,912
		└─BatchNorm2d: 3-170	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-171	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-172	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-173	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-174	[-1, 1024, 14, 14]	--
	└─Bottleneck: 2-20		[-1, 1024, 14, 14]	--
		└─Conv2d: 3-175	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-176	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-177	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-178	[-1, 1024, 14, 14]	294,912
		└─BatchNorm2d: 3-179	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-180	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-181	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-182	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-183	[-1, 1024, 14, 14]	--
	└─Bottleneck: 2-21		[-1, 1024, 14, 14]	--
		└─Conv2d: 3-184	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-185	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-186	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-187	[-1, 1024, 14, 14]	294,912
		└─BatchNorm2d: 3-188	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-189	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-190	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-191	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-192	[-1, 1024, 14, 14]	--
	└─Bottleneck: 2-22		[-1, 1024, 14, 14]	--
		└─Conv2d: 3-193	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-194	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-195	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-196	[-1, 1024, 14, 14]	294,912
		└─BatchNorm2d: 3-197	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-198	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-199	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-200	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-201	[-1, 1024, 14, 14]	--
	└─Bottleneck: 2-23		[-1, 1024, 14, 14]	--
		└─Conv2d: 3-202	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-203	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-204	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-205	[-1, 1024, 14, 14]	294,912
		└─BatchNorm2d: 3-206	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-207	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-208	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-209	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-210	[-1, 1024, 14, 14]	--
	└─Bottleneck: 2-24		[-1, 1024, 14, 14]	--
		└─Conv2d: 3-211	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-212	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-213	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-214	[-1, 1024, 14, 14]	294,912
		└─BatchNorm2d: 3-215	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-216	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-217	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-218	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-219	[-1, 1024, 14, 14]	--
	└─Bottleneck: 2-25		[-1, 1024, 14, 14]	--
		└─Conv2d: 3-220	[-1, 1024, 14, 14]	1,048,576
		└─BatchNorm2d: 3-221	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-222	[-1, 1024, 14, 14]	--
		└─Conv2d: 3-223	[-1, 1024, 14, 14]	294,912
		└─BatchNorm2d: 3-224	[-1, 1024, 14, 14]	2,048
		└─ReLU: 3-225	[-1, 1024, 14, 14]	--

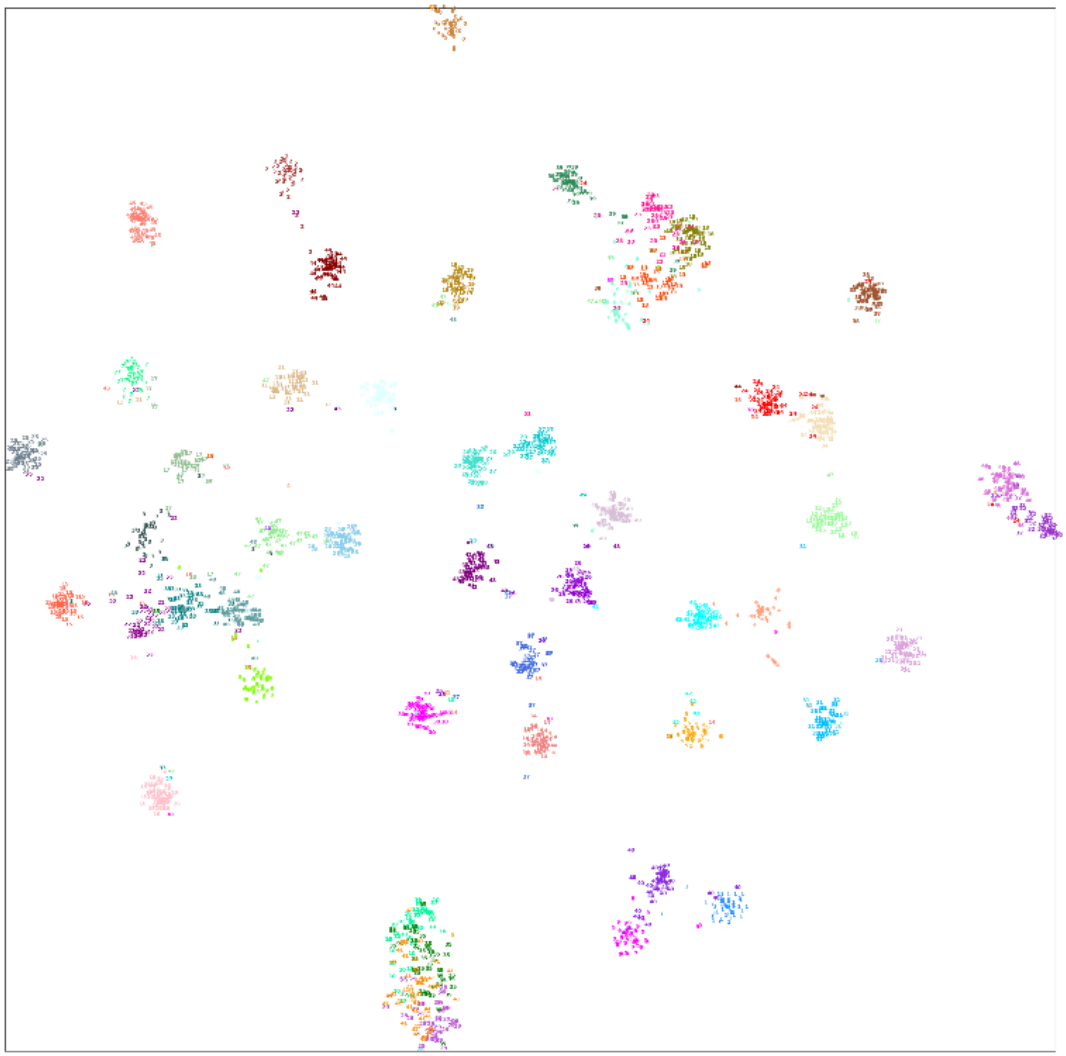
```
|    |    └─Conv2d: 3-226                    [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-227               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-228                      [-1, 1024, 14, 14]        --
|    └─Bottleneck: 2-26                      [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-229                    [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-230               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-231                      [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-232                    [-1, 1024, 14, 14]        294,912
|    |    └─BatchNorm2d: 3-233               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-234                      [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-235                    [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-236               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-237                      [-1, 1024, 14, 14]        --
|    └─Bottleneck: 2-27                      [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-238                    [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-239               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-240                      [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-241                    [-1, 1024, 14, 14]        294,912
|    |    └─BatchNorm2d: 3-242               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-243                      [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-244                    [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-245               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-246                      [-1, 1024, 14, 14]        --
|    └─Bottleneck: 2-28                      [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-247                    [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-248               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-249                      [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-250                    [-1, 1024, 14, 14]        294,912
|    |    └─BatchNorm2d: 3-251               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-252                      [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-253                    [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-254               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-255                      [-1, 1024, 14, 14]        --
|    └─Bottleneck: 2-29                      [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-256                    [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-257               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-258                      [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-259                    [-1, 1024, 14, 14]        294,912
|    |    └─BatchNorm2d: 3-260               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-261                      [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-262                    [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-263               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-264                      [-1, 1024, 14, 14]        --
|    └─Bottleneck: 2-30                      [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-265                    [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-266               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-267                      [-1, 1024, 14, 14]        --
|    |    └─Conv2d: 3-268                    [-1, 1024, 14, 14]        294,912
|    |    └─BatchNorm2d: 3-269               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-270                      [-1, 1024, 14, 14]        --
```

```
|    |    └─Conv2d: 3-271                    [-1, 1024, 14, 14]        1,048,576
|    |    └─BatchNorm2d: 3-272               [-1, 1024, 14, 14]        2,048
|    |    └─ReLU: 3-273                      [-1, 1024, 14, 14]        --
├─Sequential: 1-8                           [-1, 2048, 7, 7]          --
|    └─Bottleneck: 2-31                      [-1, 2048, 7, 7]          --
|    |    └─Conv2d: 3-274                    [-1, 2048, 14, 14]        2,097,152
|    |    └─BatchNorm2d: 3-275               [-1, 2048, 14, 14]        4,096
|    |    └─ReLU: 3-276                      [-1, 2048, 14, 14]        --
|    |    └─Conv2d: 3-277                    [-1, 2048, 7, 7]          1,179,648
|    |    └─BatchNorm2d: 3-278               [-1, 2048, 7, 7]          4,096
|    |    └─ReLU: 3-279                      [-1, 2048, 7, 7]          --
|    |    └─Conv2d: 3-280                    [-1, 2048, 7, 7]          4,194,304
|    |    └─BatchNorm2d: 3-281               [-1, 2048, 7, 7]          4,096
|    |    └─Sequential: 3-282                [-1, 2048, 7, 7]          2,101,248
|    |    └─ReLU: 3-283                      [-1, 2048, 7, 7]          --
|    └─Bottleneck: 2-32                      [-1, 2048, 7, 7]          --
|    |    └─Conv2d: 3-284                    [-1, 2048, 7, 7]          4,194,304
|    |    └─BatchNorm2d: 3-285               [-1, 2048, 7, 7]          4,096
|    |    └─ReLU: 3-286                      [-1, 2048, 7, 7]          --
|    |    └─Conv2d: 3-287                    [-1, 2048, 7, 7]          1,179,648
|    |    └─BatchNorm2d: 3-288               [-1, 2048, 7, 7]          4,096
|    |    └─ReLU: 3-289                      [-1, 2048, 7, 7]          --
|    |    └─Conv2d: 3-290                    [-1, 2048, 7, 7]          4,194,304
|    |    └─BatchNorm2d: 3-291               [-1, 2048, 7, 7]          4,096
|    |    └─ReLU: 3-292                      [-1, 2048, 7, 7]          --
|    └─Bottleneck: 2-33                      [-1, 2048, 7, 7]          --
|    |    └─Conv2d: 3-293                    [-1, 2048, 7, 7]          4,194,304
|    |    └─BatchNorm2d: 3-294               [-1, 2048, 7, 7]          4,096
|    |    └─ReLU: 3-295                      [-1, 2048, 7, 7]          --
|    |    └─Conv2d: 3-296                    [-1, 2048, 7, 7]          1,179,648
|    |    └─BatchNorm2d: 3-297               [-1, 2048, 7, 7]          4,096
|    |    └─ReLU: 3-298                      [-1, 2048, 7, 7]          --
|    |    └─Conv2d: 3-299                    [-1, 2048, 7, 7]          4,194,304
|    |    └─BatchNorm2d: 3-300               [-1, 2048, 7, 7]          4,096
|    |    └─ReLU: 3-301                      [-1, 2048, 7, 7]          --
├─AdaptiveAvgPool2d: 1-9                    [-1, 2048, 1, 1]          --
├─Linear: 1-10                              [-1, 50]                  102,450
==========================================================================================
Total params: 86,844,786
Trainable params: 86,844,786
Non-trainable params: 0
Total mult-adds (G): 16.59
==========================================================================================
Input size (MB): 0.57
Forward/backward pass size (MB): 476.22
Params size (MB): 331.29
Estimated Total Size (MB): 808.08
==========================================================================================
```

2. Validation accuracy：89.60%

3.

On my result of the t-SNE visualization, we can see that some clusters are very close and some testing data is separated into the wrong clusters. However, we can observe that if the two clusters are very close, it means that the two categories are very similar, and vice verse.



▲t-SNE visualization

Problem2

1. VGG16-FCN32s

```
Layer (type:depth-idx)                   Output Shape              Param #
==========================================================================================
├─Sequential: 1-1                        [-1, 512, 21, 21]         --
|    └─Conv2d: 2-1                       [-1, 64, 646, 646]        1,792
|    └─ReLU: 2-2                         [-1, 64, 646, 646]        --
|    └─Conv2d: 2-3                       [-1, 64, 646, 646]        36,928
|    └─ReLU: 2-4                         [-1, 64, 646, 646]        --
|    └─MaxPool2d: 2-5                    [-1, 64, 323, 323]        --
|    └─Conv2d: 2-6                       [-1, 128, 323, 323]       73,856
|    └─ReLU: 2-7                         [-1, 128, 323, 323]       --
|    └─Conv2d: 2-8                       [-1, 128, 323, 323]       147,584
|    └─ReLU: 2-9                         [-1, 128, 323, 323]       --
|    └─MaxPool2d: 2-10                   [-1, 128, 162, 162]       --
|    └─Conv2d: 2-11                      [-1, 256, 162, 162]       295,168
|    └─ReLU: 2-12                        [-1, 256, 162, 162]       --
|    └─Conv2d: 2-13                      [-1, 256, 162, 162]       590,080
|    └─ReLU: 2-14                        [-1, 256, 162, 162]       --
|    └─Conv2d: 2-15                      [-1, 256, 162, 162]       590,080
|    └─ReLU: 2-16                        [-1, 256, 162, 162]       --
|    └─MaxPool2d: 2-17                   [-1, 256, 81, 81]         --
|    └─Conv2d: 2-18                      [-1, 512, 81, 81]         1,180,160
|    └─ReLU: 2-19                        [-1, 512, 81, 81]         --
|    └─Conv2d: 2-20                      [-1, 512, 81, 81]         2,359,808
|    └─ReLU: 2-21                        [-1, 512, 81, 81]         --
|    └─Conv2d: 2-22                      [-1, 512, 81, 81]         2,359,808
|    └─ReLU: 2-23                        [-1, 512, 81, 81]         --
|    └─MaxPool2d: 2-24                   [-1, 512, 41, 41]         --
|    └─Conv2d: 2-25                      [-1, 512, 41, 41]         2,359,808
|    └─ReLU: 2-26                        [-1, 512, 41, 41]         --
|    └─Conv2d: 2-27                      [-1, 512, 41, 41]         2,359,808
|    └─ReLU: 2-28                        [-1, 512, 41, 41]         --
|    └─Conv2d: 2-29                      [-1, 512, 41, 41]         2,359,808
|    └─ReLU: 2-30                        [-1, 512, 41, 41]         --
|    └─MaxPool2d: 2-31                   [-1, 512, 21, 21]         --
├─Sequential: 1-2                        [-1, 7, 15, 15]           --
|    └─Conv2d: 2-32                      [-1, 4096, 15, 15]        102,764,544
|    └─ReLU: 2-33                        [-1, 4096, 15, 15]        --
|    └─Dropout: 2-34                     [-1, 4096, 15, 15]        --
|    └─Conv2d: 2-35                      [-1, 4096, 15, 15]        16,781,312
|    └─ReLU: 2-36                        [-1, 4096, 15, 15]        --
|    └─Dropout: 2-37                     [-1, 4096, 15, 15]        --
|    └─Conv2d: 2-38                      [-1, 7, 15, 15]           28,679
├─ConvTranspose2d: 1-3                   [-1, 7, 512, 512]         200,704
==========================================================================================
Total params: 134,489,927
Trainable params: 134,489,927
Non-trainable params: 0
Total mult-adds (G): 208.13
==========================================================================================
Input size (MB): 2.30
Forward/backward pass size (MB): 889.74
Params size (MB): 513.04
Estimated Total Size (MB): 1405.07
==========================================================================================
```

2. Deeplabv3_resnet50

```
==========================================================================================
Layer (type:depth-idx)                   Output Shape              Param #
==========================================================================================
├─IntermediateLayerGetter: 1-1          [[-1, 2048, 56, 56]]      --
│    └─Conv2d: 2-1                      [-1, 64, 224, 224]        9,408
│    └─BatchNorm2d: 2-2                 [-1, 64, 224, 224]        128
│    └─ReLU: 2-3                        [-1, 64, 224, 224]        --
│    └─MaxPool2d: 2-4                   [-1, 64, 112, 112]        --
│    └─Sequential: 2-5                  [-1, 256, 112, 112]       --
│    │    └─Bottleneck: 3-1             [-1, 256, 112, 112]       75,008
│    │    └─Bottleneck: 3-2             [-1, 256, 112, 112]       70,400
│    │    └─Bottleneck: 3-3             [-1, 256, 112, 112]       70,400
│    └─Sequential: 2-6                  [-1, 512, 56, 56]         --
│    │    └─Bottleneck: 3-4             [-1, 512, 56, 56]         379,392
│    │    └─Bottleneck: 3-5             [-1, 512, 56, 56]         280,064
│    │    └─Bottleneck: 3-6             [-1, 512, 56, 56]         280,064
│    │    └─Bottleneck: 3-7             [-1, 512, 56, 56]         280,064
│    └─Sequential: 2-7                  [-1, 1024, 56, 56]        --
│    │    └─Bottleneck: 3-8             [-1, 1024, 56, 56]        1,512,448
│    │    └─Bottleneck: 3-9             [-1, 1024, 56, 56]        1,117,184
│    │    └─Bottleneck: 3-10            [-1, 1024, 56, 56]        1,117,184
│    │    └─Bottleneck: 3-11            [-1, 1024, 56, 56]        1,117,184
│    │    └─Bottleneck: 3-12            [-1, 1024, 56, 56]        1,117,184
│    │    └─Bottleneck: 3-13            [-1, 1024, 56, 56]        1,117,184
│    └─Sequential: 2-8                  [-1, 2048, 56, 56]        --
│    │    └─Bottleneck: 3-14            [-1, 2048, 56, 56]        6,039,552
│    │    └─Bottleneck: 3-15            [-1, 2048, 56, 56]        4,462,592
│    │    └─Bottleneck: 3-16            [-1, 2048, 56, 56]        4,462,592
├─DeepLabHead: 1-2                      [-1, 7, 56, 56]           --
│    └─ASPP: 2-9                        [-1, 256, 56, 56]         --
│    │    └─Sequential: 3-17            [-1, 256, 56, 56]         328,192
│    └─Conv2d: 2-10                     [-1, 256, 56, 56]         589,824
│    └─BatchNorm2d: 2-11                [-1, 256, 56, 56]         512
│    └─ReLU: 2-12                       [-1, 256, 56, 56]         --
│    └─Conv2d: 2-13                     [-1, 7, 56, 56]           1,799
├─FCNHead: 1-3                          [-1, 7, 56, 56]           --
│    └─Conv2d: 2-14                     [-1, 256, 56, 56]         2,359,296
│    └─BatchNorm2d: 2-15                [-1, 256, 56, 56]         512
│    └─ReLU: 2-16                       [-1, 256, 56, 56]         --
│    └─Dropout: 2-17                    [-1, 256, 56, 56]         --
│    └─Conv2d: 2-18                     [-1, 7, 56, 56]           1,799
==========================================================================================
```

```
Total params: 26,789,966
Trainable params: 26,789,966
Non-trainable params: 0
Total mult-adds (G): 77.88
==========================================================================================
Input size (MB): 2.30
Forward/backward pass size (MB): 1353.96
Params size (MB): 102.20
Estimated Total Size (MB): 1458.45
==========================================================================================
```

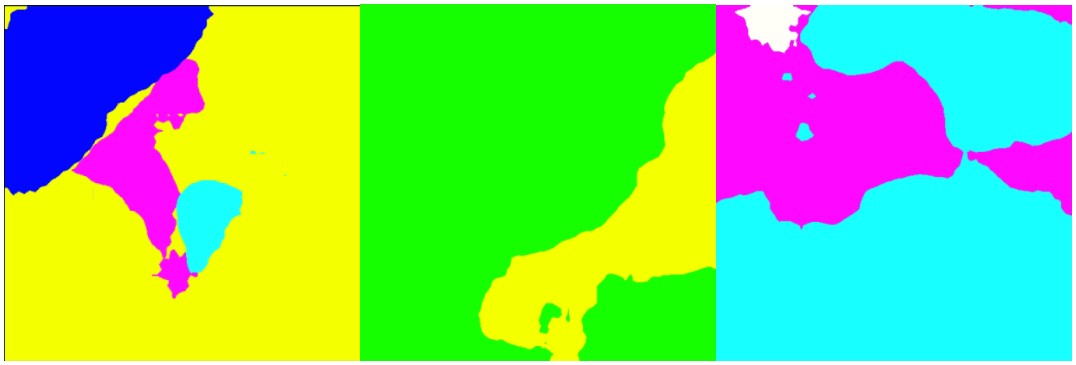
3. mIoU: 71.54%

4.

Early(mIoU=40.37%):

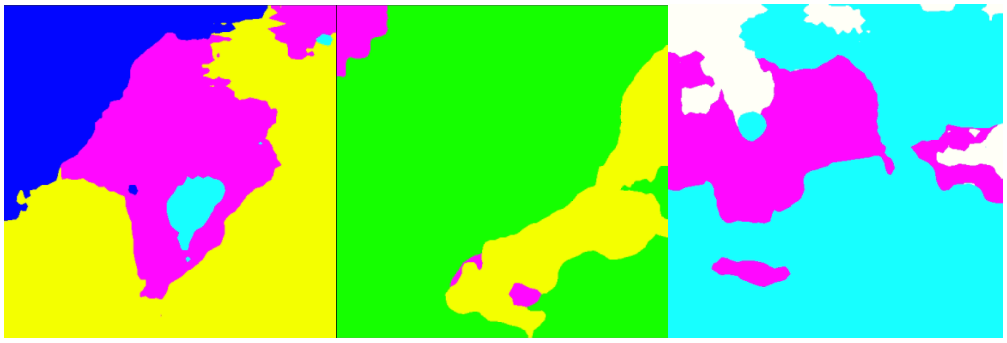
▲0010_mask.png ▲0097_mask.png ▲0107_mask.png

Middle(mIoU=66.81%):



▲0010_mask.png ▲0097_mask.png ▲0107_mask.png

Final stage(mIoU=71.54%):



▲0010_mask.png ▲0097_mask.png ▲0107_mask.png