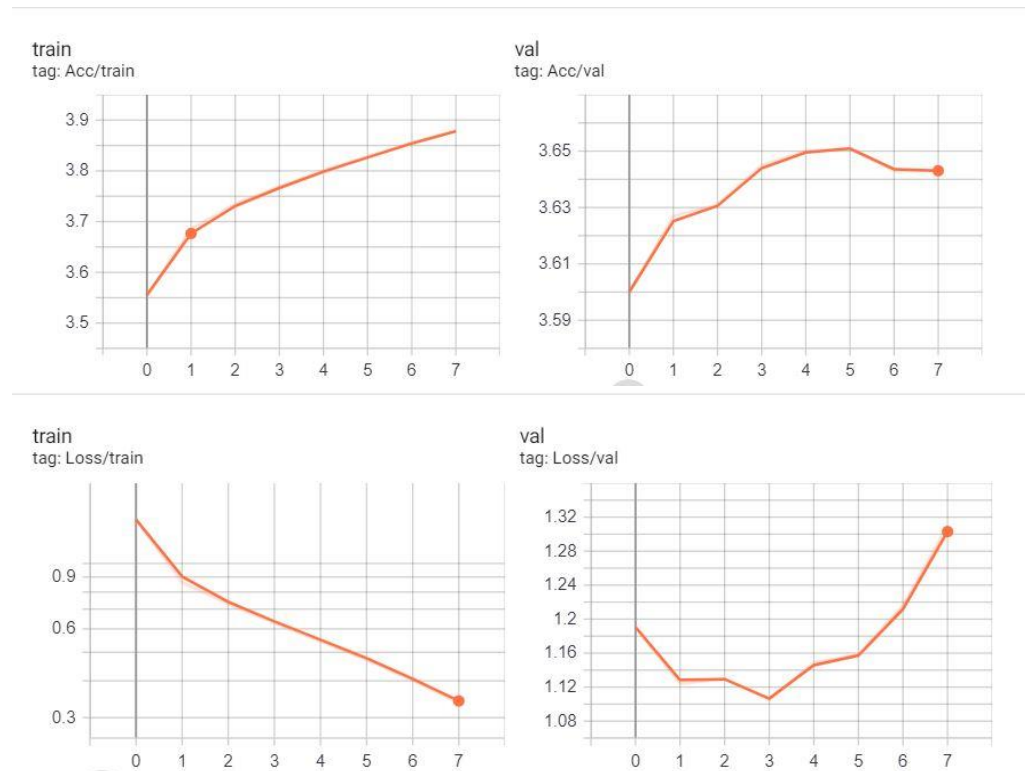


Computer Vision Exercise

Task 1: Semantic Segmentation

Pretraining task:

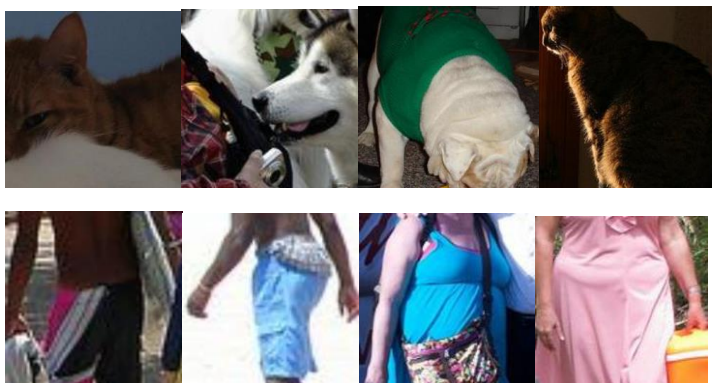
The pretrained Resnet helped to reduce training time. The initial training using 256 x 256 size images was taking too long for a single epoch (~ 3 hours) so I decided to reduce the size of the images to 128 x 128. This speeded up the training by a significant amount. After 3 epochs the validation loss started increasing and the validation accuracy began to flatten, so after waiting for 5 more epochs I stopped the training. The accuracy and loss curves are as follows:



Nearest Neighbour Task

Despite the reduction in size the network seems to have learned some meaningful features as seen from the images. Using a size of 256 and 128, I obtained 2 sets of results.

For 256 x 256



For 128 x 128



Overall observation is that the network gives more consistent results when the input size for the nearest neighbour task matches the input size for which the network was trained.

Binary Segmentation task

For the Binary segmentation task, I used BCELogits loss and SGD optimizer without any scheduler. During the 1st epoch of validation, I realized that my input was being resized by one pixel after the decoder stage in the deeplabV. This was because in the deeplab decoder the forward method input variable took the low-level feature dimension and multiplied it by 4. Since my image had a width of 191 and the output of the decoder was $48 \times 4 = 192$, it caused a discrepancy. I rectified this by force resizing all images to 128 x 128 in the transforms.py file.

```

17: self.encoder = nn.Sequential(
18:     nn.Conv2d(3, 16, 3, padding=1),
19:     nn.ReLU(),
20:     nn.Conv2d(16, 32, 3, padding=1),
21:     nn.ReLU(),
22:     nn.Conv2d(32, 64, 3, padding=1),
23:     nn.ReLU(),
24:     nn.Conv2d(64, 128, 3, padding=1),
25:     nn.ReLU(),
26:     nn.Conv2d(128, 256, 3, padding=1),
27:     nn.ReLU(),
28:     nn.Conv2d(256, 512, 3, padding=1),
29:     nn.ReLU(),
30:     nn.Conv2d(512, 1024, 3, padding=1),
31:     nn.ReLU(),
32:     nn.Conv2d(1024, 2048, 3, padding=1),
33:     nn.ReLU(),
34:     nn.Conv2d(2048, 4096, 3, padding=1),
35:     nn.ReLU(),
36:     nn.Conv2d(4096, 8192, 3, padding=1),
37:     nn.ReLU(),
38:     nn.Conv2d(8192, 16384, 3, padding=1),
39:     nn.ReLU(),
40:     nn.Conv2d(16384, 32768, 3, padding=1),
41:     nn.ReLU(),
42:     nn.Conv2d(32768, 65536, 3, padding=1),
43:     nn.ReLU(),
44:     nn.Conv2d(65536, 131072, 3, padding=1),
45:     nn.ReLU(),
46:     nn.Conv2d(131072, 262144, 3, padding=1),
47:     nn.ReLU(),
48:     nn.Conv2d(262144, 524288, 3, padding=1),
49:     nn.ReLU(),
50:     nn.Conv2d(524288, 1048576, 3, padding=1),
51:     nn.ReLU(),
52:     nn.Conv2d(1048576, 2097152, 3, padding=1),
53:     nn.ReLU(),
54:     nn.Conv2d(2097152, 4194304, 3, padding=1),
55:     nn.ReLU(),
56:     nn.Conv2d(4194304, 8388608, 3, padding=1),
57:     nn.ReLU(),
58:     nn.Conv2d(8388608, 16777216, 3, padding=1),
59:     nn.ReLU(),
60:     nn.Conv2d(16777216, 33554432, 3, padding=1),
61:     nn.ReLU(),
62:     nn.Conv2d(33554432, 67108864, 3, padding=1),
63:     nn.ReLU(),
64:     nn.Conv2d(67108864, 134217728, 3, padding=1),
65:     nn.ReLU(),
66:     nn.Conv2d(134217728, 268435456, 3, padding=1),
67:     nn.ReLU(),
68:     nn.Conv2d(268435456, 536870912, 3, padding=1),
69:     nn.ReLU(),
70:     nn.Conv2d(536870912, 1073741824, 3, padding=1),
71:     nn.ReLU(),
72:     nn.Conv2d(1073741824, 2147483648, 3, padding=1),
73:     nn.ReLU(),
74:     nn.Conv2d(2147483648, 4294967296, 3, padding=1),
75:     nn.ReLU(),
76:     nn.Conv2d(4294967296, 8589934592, 3, padding=1),
77:     nn.ReLU(),
78:     nn.Conv2d(8589934592, 17179869184, 3, padding=1),
79:     nn.ReLU(),
80:     nn.Conv2d(17179869184, 34359738368, 3, padding=1),
81:     nn.ReLU(),
82:     nn.Conv2d(34359738368, 68719476736, 3, padding=1),
83:     nn.ReLU(),
84:     nn.Conv2d(68719476736, 137438953472, 3, padding=1),
85:     nn.ReLU(),
86:     nn.Conv2d(137438953472, 274877906944, 3, padding=1),
87:     nn.ReLU(),
88:     nn.Conv2d(274877906944, 549755813888, 3, padding=1),
89:     nn.ReLU(),
90:     nn.Conv2d(549755813888, 1099511627776, 3, padding=1),
91:     nn.ReLU(),
92:     nn.Conv2d(1099511627776, 2199023255552, 3, padding=1),
93:     nn.ReLU(),
94:     nn.Conv2d(2199023255552, 4398046511104, 3, padding=1),
95:     nn.ReLU(),
96:     nn.Conv2d(4398046511104, 8796093022208, 3, padding=1),
97:     nn.ReLU(),
98:     nn.Conv2d(8796093022208, 17592186044416, 3, padding=1),
99:     nn.ReLU(),
100:    nn.Conv2d(17592186044416, 35184372088832, 3, padding=1),
101:    nn.ReLU(),
102:    nn.Conv2d(35184372088832, 70368744177664, 3, padding=1),
103:    nn.ReLU(),
104:    nn.Conv2d(70368744177664, 140737488355328, 3, padding=1),
105:    nn.ReLU(),
106:    nn.Conv2d(140737488355328, 281474976710656, 3, padding=1),
107:    nn.ReLU(),
108:    nn.Conv2d(281474976710656, 562949953421312, 3, padding=1),
109:    nn.ReLU(),
110:    nn.Conv2d(562949953421312, 1125899906842624, 3, padding=1),
111:    nn.ReLU(),
112:    nn.Conv2d(1125899906842624, 2251799813685248, 3, padding=1),
113:    nn.ReLU(),
114:    nn.Conv2d(2251799813685248, 4503599627370496, 3, padding=1),
115:    nn.ReLU(),
116:    nn.Conv2d(4503599627370496, 9007199254740992, 3, padding=1),
117:    nn.ReLU(),
118:    nn.Conv2d(9007199254740992, 18014398509481984, 3, padding=1),
119:    nn.ReLU(),
120:    nn.Conv2d(18014398509481984, 36028797018963968, 3, padding=1),
121:    nn.ReLU(),
122:    nn.Conv2d(36028797018963968, 72057594037927936, 3, padding=1),
123:    nn.ReLU(),
124:    nn.Conv2d(72057594037927936, 144115188075855872, 3, padding=1),
125:    nn.ReLU(),
126:    nn.Conv2d(144115188075855872, 288230376151711744, 3, padding=1),
127:    nn.ReLU(),
128:    nn.Conv2d(288230376151711744, 576460752303423488, 3, padding=1),
129:    nn.ReLU(),
130:    nn.Conv2d(576460752303423488, 1152921504606846976, 3, padding=1),
131:    nn.ReLU(),
132:    nn.Conv2d(1152921504606846976, 2305843009213693952, 3, padding=1),
133:    nn.ReLU(),
134:    nn.Conv2d(2305843009213693952, 4611686018427387904, 3, padding=1),
135:    nn.ReLU(),
136:    nn.Conv2d(4611686018427387904, 9223372036854775808, 3, padding=1),
137:    nn.ReLU(),
138:    nn.Conv2d(9223372036854775808, 18446744073709551616, 3, padding=1),
139:    nn.ReLU(),
140:    nn.Conv2d(18446744073709551616, 36893488147419103232, 3, padding=1),
141:    nn.ReLU(),
142:    nn.Conv2d(36893488147419103232, 73786976294838206464, 3, padding=1),
143:    nn.ReLU(),
144:    nn.Conv2d(73786976294838206464, 147573952589676412928, 3, padding=1),
145:    nn.ReLU(),
146:    nn.Conv2d(147573952589676412928, 295147905179352825856, 3, padding=1),
147:    nn.ReLU(),
148:    nn.Conv2d(295147905179352825856, 590295810358705651712, 3, padding=1),
149:    nn.ReLU(),
150:    nn.Conv2d(590295810358705651712, 1180591620717411303424, 3, padding=1),
151:    nn.ReLU(),
152:    nn.Conv2d(1180591620717411303424, 2361183241434822606848, 3, padding=1),
153:    nn.ReLU(),
154:    nn.Conv2d(2361183241434822606848, 4722366482869645213696, 3, padding=1),
155:    nn.ReLU(),
156:    nn.Conv2d(4722366482869645213696, 9444732965739290427392, 3, padding=1),
157:    nn.ReLU(),
158:    nn.Conv2d(9444732965739290427392, 18889465931478580854784, 3, padding=1),
159:    nn.ReLU(),
160:    nn.Conv2d(18889465931478580854784, 37778931862957161709568, 3, padding=1),
161:    nn.ReLU(),
162:    nn.Conv2d(37778931862957161709568, 75557863725914323419136, 3, padding=1),
163:    nn.ReLU(),
164:    nn.Conv2d(75557863725914323419136, 151115727451828646838272, 3, padding=1),
165:    nn.ReLU(),
166:    nn.Conv2d(151115727451828646838272, 302231454903657293676544, 3, padding=1),
167:    nn.ReLU(),
168:    nn.Conv2d(302231454903657293676544, 604462909807314587353088, 3, padding=1),
169:    nn.ReLU(),
170:    nn.Conv2d(604462909807314587353088, 1208925819614629174706176, 3, padding=1),
171:    nn.ReLU(),
172:    nn.Conv2d(1208925819614629174706176, 2417851639229258349412352, 3, padding=1),
173:    nn.ReLU(),
174:    nn.Conv2d(2417851639229258349412352, 4835703278458516698824704, 3, padding=1),
175:    nn.ReLU(),
176:    nn.Conv2d(4835703278458516698824704, 9671406556917033397649408, 3, padding=1),
177:    nn.ReLU(),
178:    nn.Conv2d(9671406556917033397649408, 19342813113834066795298816, 3, padding=1),
179:    nn.ReLU(),
180:    nn.Conv2d(19342813113834066795298816, 38685626227668133590597632, 3, padding=1),
181:    nn.ReLU(),
182:    nn.Conv2d(38685626227668133590597632, 77371252455336267181195264, 3, padding=1),
183:    nn.ReLU(),
184:    nn.Conv2d(77371252455336267181195264, 154742504910672534362390528, 3, padding=1),
185:    nn.ReLU(),
186:    nn.Conv2d(154742504910672534362390528, 309485009821345068724781056, 3, padding=1),
187:    nn.ReLU(),
188:    nn.Conv2d(309485009821345068724781056, 618970019642690137449562112, 3, padding=1),
189:    nn.ReLU(),
190:    nn.Conv2d(618970019642690137449562112, 1237940039285380274899124224, 3, padding=1),
191:    nn.ReLU(),
192:    nn.Conv2d(1237940039285380274899124224, 2475880078570760549798248448, 3, padding=1),
193:    nn.ReLU(),
194:    nn.Conv2d(2475880078570760549798248448, 4951760157141521099596496896, 3, padding=1),
195:    nn.ReLU(),
196:    nn.Conv2d(4951760157141521099596496896, 9903520314283042199192993792, 3, padding=1),
197:    nn.ReLU(),
198:    nn.Conv2d(9903520314283042199192993792, 19807040628566084398385987584, 3, padding=1),
199:    nn.ReLU(),
200:    nn.Conv2d(19807040628566084398385987584, 39614081257132168796771975168, 3, padding=1),
201:    nn.ReLU(),
202:    nn.Conv2d(39614081257132168796771975168, 79228162514264337593543950336, 3, padding=1),
203:    nn.ReLU(),
204:    nn.Conv2d(79228162514264337593543950336, 158456325028528675187087900672, 3, padding=1),
205:    nn.ReLU(),
206:    nn.Conv2d(158456325028528675187087900672, 316912650057057350374175801344, 3, padding=1),
207:    nn.ReLU(),
208:    nn.Conv2d(316912650057057350374175801344, 633825300114114700748351602688, 3, padding=1),
209:    nn.ReLU(),
210:    nn.Conv2d(633825300114114700748351602688, 1267650600228229401496703205376, 3, padding=1),
211:    nn.ReLU(),
212:    nn.Conv2d(1267650600228229401496703205376, 2535301200456458802993406410752, 3, padding=1),
213:    nn.ReLU(),
214:    nn.Conv2d(2535301200456458802993406410752, 5070602400912917605986812821504, 3, padding=1),
215:    nn.ReLU(),
216:    nn.Conv2d(5070602400912917605986812821504, 10141204801825835211973625643008, 3, padding=1),
217:    nn.ReLU(),
218:    nn.Conv2d(10141204801825835211973625643008, 20282409603651670423947251286016, 3, padding=1),
219:    nn.ReLU(),
220:    nn.Conv2d(20282409603651670423947251286016, 40564819207303340847894502572032, 3, padding=1),
221:    nn.ReLU(),
222:    nn.Conv2d(40564819207303340847894502572032, 81129638414606681695789005144064, 3, padding=1),
223:    nn.ReLU(),
224:    nn.Conv2d(81129638414606681695789005144064, 162259276829213363391578010288128, 3, padding=1),
225:    nn.ReLU(),
226:    nn.Conv2d(162259276829213363391578010288128, 324518553658426726783156020576256, 3, padding=1),
227:    nn.ReLU(),
228:    nn.Conv2d(324518553658426726783156020576256, 649037107316853453566312041152512, 3, padding=1),
229:    nn.ReLU(),
230:    nn.Conv2d(649037107316853453566312041152512, 1298074214633706907132624082305024, 3, padding=1),
231:    nn.ReLU(),
232:    nn.Conv2d(1298074214633706907132624082305024, 2596148429267413814265248164610048, 3, padding=1),
233:    nn.ReLU(),
234:    nn.Conv2d(2596148429267413814265248164610048, 5192296858534827628530496329220096, 3, padding=1),
235:    nn.ReLU(),
236:    nn.Conv2d(5192296858534827628530496329220096, 10384593717069655257060992658440192, 3, padding=1),
237:    nn.ReLU(),
238:    nn.Conv2d(10384593717069655257060992658440192, 20769187434139310514121985316880384, 3, padding=1),
239:    nn.ReLU(),
240:    nn.Conv2d(20769187434139310514121985316880384, 41538374868278621028243970633760768, 3, padding=1),
241:    nn.ReLU(),
242:    nn.Conv2d(41538374868278621028243970633760768, 83076749736557242056487941267521536, 3, padding=1),
243:    nn.ReLU(),
244:    nn.Conv2d(83076749736557242056487941267521536, 166153499473114484112975882535043072, 3, padding=1),
245:    nn.ReLU(),
246:    nn.Conv2d(166153499473114484112975882535043072, 332306998946228968225951765070086144, 3, padding=1),
247:    nn.ReLU(),
248:    nn.Conv2d(332306998946228968225951765070086144, 664613997892457936451903530140172288, 3, padding=1),
249:    nn.ReLU(),
250:    nn.Conv2d(664613997892457936451903530140172288, 1329227995784915872903807060280344576, 3, padding=1),
251:    nn.ReLU(),
252:    nn.Conv2d(1329227995784915872903807060280344576, 2658455991569831745807614120560689152, 3, padding=1),
253:    nn.ReLU(),
254:    nn.Conv2d(2658455991569831745807614120560689152, 5316911983139663491615228241121378304, 3, padding=1),
255:    nn.ReLU(),
256:    nn.Conv2d(5316911983139663491615228241121378304, 10633823966279326983230456482242756608, 3, padding=1),
257:    nn.ReLU(),
258:    nn.Conv2d(10633823966279326983230456482242756608, 21267647932558653966460912964485513216, 3, padding=1),
259:    nn.ReLU(),
260:    nn.Conv2d(21267647932558653966460912964485513216, 42535295865117307932921825928971026432, 3, padding=1),
261:    nn.ReLU(),
262:    nn.Conv2d(42535295865117307932921825928971026432, 85070591730234615865843651857942052864, 3, padding=1),
263:    nn.ReLU(),
264:    nn.Conv2d(85070591730234615865843651857942052864, 170141183460469231731687303715884105728, 3, padding=1),
265:    nn.ReLU(),
266:    nn.Conv2d(170141183460469231731687303715884105728, 340282366920938463463374607431768211456, 3, padding=1),
267:    nn.ReLU(),
268:    nn.Conv2d(340282366920938463463374607431768211456, 680564733841876926926749214863536422912, 3, padding=1),
269:    nn.ReLU(),
270:    nn.Conv2d(680564733841876926926749214863536422912, 1361129467683753853853498429727072845824, 3, padding=1),
271:    nn.ReLU(),
272:    nn.Conv2d(1361129467683753853853498429727072845824, 2722258935367507707706996859454145691648, 3, padding=1),
273:    nn.ReLU(),
274:    nn.Conv2d(2722258935367507707706996859454145691648, 5444517870735015415413993718908291383296, 3, padding=1),
275:    nn.ReLU(),
276:    nn.Conv2d(5444517870735015415413993718908291383296, 10889035741470030830827987437816582766592, 3, padding=1),
277:    nn.ReLU(),
278:    nn.Conv2d(10889035741470030830827987437816582766592, 21778071482940061661655974875633165533184, 3, padding=1),
279:    nn.ReLU(),
280:    nn.Conv2d(21778071482940061661655974875633165533184, 43556142965880123323311949751266331066368, 3, padding=1),
281:    nn.ReLU(),
282:    nn.Conv2d(43556142965880123323311949751266331066368, 87112285931760246646623899502532662132736, 3, padding=1),
283:    nn.ReLU(),
284:    nn.Conv2d(87112285931760246646623899502532662132736, 174224571863520493293247799005065324265472, 3, padding=1),
285:    nn.ReLU(),
286:    nn.Conv2d(174224571863520493293247799005065324265472, 348449143727040986586495598010130648530944, 3, padding=1),
287:    nn.ReLU(),
288:    nn.Conv2d(348449143727040986586495598010130648530944, 696898287454081973172991196020261297061888, 3, padding=1),
289:    nn.ReLU(),
290:    nn.Conv2d(696898287454081973172991196020261297061888, 1393796574908163946345982392040522594123776, 3, padding=1),
291:    nn.ReLU(),
292:    nn.Conv2d(1393796574908163946345982392040522594123776, 2787593149816327892691964784081045188247552, 3, padding=1),
293:    nn.ReLU(),
294:    nn.Conv2d(2787593149816327892691964784081045188247552, 5575186299632655785383929568162090376495104, 3, padding=1),
295:    nn.ReLU(),
296:    nn.Conv2d(5575186299632655785383929568162090376495104, 11150372599265311570767859136324180752990208, 3, padding=1),
297:    nn.ReLU(),
298:    nn.Conv2d(11150372599265311570767859136324180752990208, 22300745198530623141535718272648361505980416, 3, padding=1),
299:    nn.ReLU(),
300:    nn.Conv2d(22300745198530623141535718272648361505980416, 44601490397061246283071436545296723011960832, 3, padding=1),
301:    nn.ReLU(),
302:    nn.Conv2d(44601490397061246283071436545296723011960832, 89202980794122492566142873090593446023921664, 3, padding=1),
303:    nn.ReLU(),
304:    nn.Conv2d(89202980794122492566142873090593446023921664, 178405961588244985132285746181186892047843328, 3, padding=1),
305:    nn.ReLU(),
306:    nn.Conv2d(178405961588244985132285746181186892047843328, 356811923176489970264571492362373784095686656, 3, padding=1),
307:    nn.ReLU(),
308:    nn.Conv2d(356811923176489970264571492362373784095686656, 713623846352979940529142984724747568191373312, 3, padding=1),
309:    nn.ReLU(),
310:    nn.Conv2d(713623846352979940529142984724747568191373312, 1427247692705959881058285969449495136382746624, 3, padding=1),
311:    nn.ReLU(),
312:    nn.Conv2d(1427247692705959881058285969449495136382746624, 2854495385411919762116571938898990272765493248, 3, padding=1),
313:    nn.ReLU(),
314:    nn.Conv2d(2854495385411919762116571938898990272765493248, 5708990770823839524233143877797980545530986496, 3, padding=1),
315:    nn.ReLU(),
316:    nn.Conv2d(5708990770823839524233143877797980545530986496, 11417981541647679048466287755595961091061972992, 3, padding=1),
317:    nn.ReLU(),
318:    nn.Conv2d(11417981541647679048466287755595961091061972992, 22835963083295358096932575511191922182123945984, 3, padding=1),
319:    nn.ReLU(),
320:    nn.Conv2d(22835963083295358096932575511191922182123945984, 45671926166590716193865151022383844364247891968, 3, padding=1),
321:    nn.ReLU(),
322:    nn.Conv2d(45671926166590716193865151022383844364247891
```

