Animal Detection in Man-made Environments
Supplementary

Abhineet Singh 1, Marcin Pietrasik 2, Gabriell Natha 2, Nehla Ghouaiel 2, Ken Brizel 2, and Nilanjan Ray 1

1Department of Computing Science, University of Alberta
2Alberta Centre for Advanced MNT Products (ACAMP)

I. INTRODUCTION

This supplementary provides details of the tracking (Sec. II) and multi-model pooling experiments (Sec. III) along with the corresponding results. Class-agnostic Recall-Precision (cRP) for training configurations #1 - #5 (Table II) along with class-specific results for #1 and #2 (Tables III - VI) are also included. Finally, links are provided for more synthetic data samples generated using the four mask generation methods detailed in the paper (Table I).

A video demonstrating the end-to-end real and synthetic data generation pipeline created using the labeling tool accompanies this document. A higher resolution version of the same is also available on YouTube [2] along with mask generation results produced by several instance and semantic segmentation methods [1].

II. YOLO+DASiamRPN

Algorithm 1 details the process used for combining YOLO with DASiamRPN tracker to reduce false negatives using temporal information in videos. The cumulative_score in line 46 is a composite measure of overall tracker fitness defined as:

\[
\text{cumulative\_score} = \text{cumulative\_confidence} \times \frac{\text{assoc\_count} + 1}{\text{unassoc\_count} + 1}
\]

where assoc_count is the number of frames that the tracker has been successfully associated in since its initialization, unassoc_count is the number of frames since it was last associated and cumulative_confidence is the product of tracker confidence in all tracked frames so far as well as the confidence of the detection on which it was initialized.

Removing trackers that have remained unassociated for too long (lines 30-34) can help to increase precision by reducing tracked boxes corresponding to false positive detections. However, this also leads to a significant drop in recall since false negatives often occur in several consecutive frames which causes trackers corresponding to real objects to be removed. Thus, the overall performance, shown in Fig. 1, is nearly identical to not using tracking at all. Experiments were done with max_unassoc ∈ {2, 3, 4, 5} and best results were obtained with max_unassoc = 2, which is the only one included here.

III. MUTI-MODEL POOLING

Muti-model pooling can be an effective way to reduce false negatives only if the pooled models exhibit distinct patterns of missing objects. As shown in Sec. 4.2 of the main paper, model recall depends strongly on the range of backgrounds in the training set. It would thus seem that model separation can best be achieved by training them on images with significantly different backgrounds, e.g. with snowy winter and green summer scenes or under sunny and cloudy / rainy weather conditions. However, the existing labeled data does not contain sufficient

Algorithm 1 YOLO+DASiamRPN

1: \text{min\_conf, min\_iou, max\_unassoc, max\_trackers, min\_init\_gap} \leftarrow \text{params}
2: \text{trackers} \leftarrow \emptyset
3: \text{for frame in video sequence do}
4: \quad \text{detections} \leftarrow \text{YOLO (frame)}
5: \quad \text{associated\_detections} \leftarrow \emptyset
6: \quad \text{associated\_trackers} \leftarrow \emptyset
7: \quad \text{for} t \text{ in trackers do}
8: \quad \quad t.\text{update(frame)}
9: \quad \quad \text{if} t.\text{confidence} < \text{min\_conf} \text{ then}
10: \quad \quad \quad \text{trackers} \leftarrow \text{trackers} \setminus t
11: \quad \quad \text{continue}
12: \quad \quad \text{end if}
13: \quad \quad \text{if} t \text{ in associated\_trackers} \text{ then}
14: \quad \quad \quad \text{continue}
15: \quad \quad \text{end if}
16: \quad \quad \text{for} d \text{ in detections do}
17: \quad \quad \quad \text{if} d \text{ in associated\_detections} \text{ then}
18: \quad \quad \quad \quad \text{continue}
19: \quad \quad \quad \text{end if}
20: \quad \quad \quad \text{if} d.\text{class} = t.\text{class} \text{ and iou(d, t) > min\_iou} \text{ then}
21: \quad \quad \quad \quad \text{associated\_detections} \leftarrow \text{associated\_detections} \cup d
22: \quad \quad \quad \quad \text{associated\_trackers} \leftarrow \text{associated\_trackers} \cup t
23: \quad \quad \quad \quad \text{t\_assoc\_count} \leftarrow 0
24: \quad \quad \quad \quad \text{t\_assoc\_count} \leftarrow t\_assoc\_count + 1
25: \quad \quad \text{continue}
26: \quad \quad \text{end if}
27: \quad \quad \text{end for}
28: \quad \quad \text{end for}
29: \quad \text{unassociated\_trackers} \leftarrow \text{trackers} \setminus \text{associated\_trackers}
30: \quad \text{for} t \text{ in unassociated\_trackers do}
31: \quad \quad \text{t\_unassoc\_count} \leftarrow \text{t\_unassoc\_count} + 1
32: \quad \quad \text{if} \text{t\_unassoc\_count} > \text{max\_unassoc} \text{ then}
33: \quad \quad \quad \text{trackers} \leftarrow \text{trackers} \setminus t
34: \quad \quad \text{continue}
35: \quad \quad \text{end if}
36: \quad \text{end for}
37: \quad \text{unassociated\_detections} \leftarrow \text{detections} \setminus \text{associated\_detections}
38: \quad \text{tracker\_init\_gap} \leftarrow \text{tracker\_init\_gap} + 1
39: \quad \text{if} \text{tracker\_init\_gap} > \text{min\_init\_gap} \text{ then}
40: \quad \quad \text{tracker\_init\_gap} \leftarrow 0
41: \quad \quad \text{for} d \text{ in unassociated\_detections do}
42: \quad \quad \quad t \leftarrow \text{initialize\_new\_tracker(frame, d)}
43: \quad \quad \quad \text{t\_unassoc\_count} \leftarrow 0
44: \quad \quad \quad \text{t\_assoc\_count} \leftarrow 1
45: \quad \quad \quad \text{trackers} \leftarrow \text{trackers} \cup t
46: \quad \quad \text{end for}
47: \quad \text{end for}
48: \quad \text{trackers} \leftarrow \text{sort trackers by cumulative\_score} \text{ (Eq. 1)}
49: \quad \text{trackers} \leftarrow \{t | t_i \in \text{trackers}, 1 \leq i \leq \text{max\_trackers}\}
50: \quad \text{end if}
51: \quad \text{end for}
variety to divide it into class-balanced subsets that each have relatively homogeneous backgrounds differing significantly from those in other subsets. Model separation was instead achieved by training different architectures on the same data.

Two pooling strategies were employed:

1) Model Switching: All detections from the model with the highest confidence detection are retained in each frame while all others models’ detections are discarded.

2) Model Aggregation: Detections from all models are pooled followed by class-agnostic non-maximum suppression applied to boxes from different models. Both strategies require inter-model confidence comparison which is not straightforward to do since, as shown in Fig. 3 of the main paper (and Tables III - VI here), the meaning of confidence magnitude in terms of detection accuracy varies significantly between models. As a result, raw confidence values from different models need to be normalized before they can be compared. Several ways of doing this were explored but the one that worked best was to scale the confidence values of each model so that its mRP threshold becomes 0.5 while the range of values remains [0, 1].

Experiments were done by pooling YOLO and RETINA as well as YOLO, SSDIN and SSDMO but the former combination gave much better results so only these are included here. Similarly, configurations #1, #3 and #5 were all tested but only #5 is included as it produced the best pooling results relative to the individual detectors. Fig. 2 gives results for the two strategies in #5, both with and without confidence normalization. Somewhat surprisingly, none of the normalization techniques managed to outperform unnormalized confidence based pooling. Also, none of the pooling strategies managed to increase recall over the better of the pooled detectors while always leading to a significant drop in precision. This shows that pooling is not a practical way to improve recall, at least not with the relatively simple methods employed here so as not to compromise speed.

IV. CLASS-AGNOSTIC RECALL-PRECISION (cRP)

Table II shows the cRPs for all models in configurations #1 - #5. It is interesting to note that the large relative increase in cRP exhibited by YOLO in #1 and #3 is not present in the other easier configurations. This might indicate that YOLO manifests its tendency to overfit to specific background-foreground combinations most strongly with very limited training data.

V. CLASS-SPECIFIC RESULTS

Results for configuration #1 are in tables III and IV while those for #2 are in tables V and VI. The second column gives the recall – precision (RP) value corresponding to the score threshold where the two are equal while the threshold itself is given in the third column. The next three columns provide the recall, precision and their average for the score threshold corresponding to the overall mRP, i.e. where the mean recall and mean precision are equal. The last column gives the total number of ground truth boxes available for each class. Note that many images had multiple objects so that this count is more than the number of images in the corresponding test set.

It can be seen that the large inter-class variation of RP thresholds holds for all the models as well as for both low and high performance scenarios, represented by #1 and #2 respectively. AP and RP show considerable variation too, though only in #1. This is probably because animals with the larger ranges of backgrounds in the training set are often falsely detected in test images of remaining animals, thus leading to significant biases in their recall and precision.

VI. SYNTHETIC DATA SAMPLES

Table I provides links to Google Photos albums showing synthetic data samples generated using all 4 methods - (left to right in each row) manual masks, Mask RCNN, SiamMask and no mask/Gaussian blending.

---

**References**


Fig. 2: Mean Recall and Precision for RETINA and YOLO pooling in #5 with and without confidence normalization

<table>
<thead>
<tr>
<th>Model</th>
<th>1: 1K/Seq/8</th>
<th>2: 1K/Even/8</th>
<th>3: 10K/Seq/6</th>
<th>4: 5%/Start/6</th>
<th>5: 500/Static/3</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAS</td>
<td>87.17</td>
<td>96.42</td>
<td>93.8</td>
<td>92.98</td>
<td>95.83</td>
</tr>
<tr>
<td>INRES</td>
<td>84.64</td>
<td>96.13</td>
<td>95.08</td>
<td>94.32</td>
<td>92.73</td>
</tr>
<tr>
<td>RES101</td>
<td>89.71</td>
<td>99.65</td>
<td>95.16</td>
<td>97.15</td>
<td>95.93</td>
</tr>
<tr>
<td>RFCN</td>
<td>88.93</td>
<td>99.54</td>
<td>94.64</td>
<td>97.27</td>
<td>93.88</td>
</tr>
<tr>
<td>RETINA</td>
<td>74.62</td>
<td>99.62</td>
<td>90.18</td>
<td>95.27</td>
<td>92.01</td>
</tr>
<tr>
<td>SSDIN</td>
<td>69.56</td>
<td>98.62</td>
<td>84.75</td>
<td>92</td>
<td>89.56</td>
</tr>
<tr>
<td>SSDMO</td>
<td>64.89</td>
<td>98.42</td>
<td>86.28</td>
<td>92.84</td>
<td>90.62</td>
</tr>
<tr>
<td>YOLO</td>
<td>80.76</td>
<td>96.67</td>
<td>91.7</td>
<td>94.16</td>
<td>90.36</td>
</tr>
</tbody>
</table>

TABLE II: Class agnostic Recall-Precision (cRP) for training configurations #1 - #5
<table>
<thead>
<tr>
<th>Training Configuration #1: 1K/Seq/8</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAS</td>
</tr>
<tr>
<td><strong>Class Specific</strong></td>
</tr>
<tr>
<td>class</td>
</tr>
<tr>
<td>bear</td>
</tr>
<tr>
<td>bison</td>
</tr>
<tr>
<td>cow</td>
</tr>
<tr>
<td>coyote</td>
</tr>
<tr>
<td>deer</td>
</tr>
<tr>
<td>elk</td>
</tr>
<tr>
<td>horse</td>
</tr>
<tr>
<td>moose</td>
</tr>
<tr>
<td>average</td>
</tr>
<tr>
<td><strong>INRES</strong></td>
</tr>
<tr>
<td><strong>Class Specific</strong></td>
</tr>
<tr>
<td>class</td>
</tr>
<tr>
<td>bear</td>
</tr>
<tr>
<td>bison</td>
</tr>
<tr>
<td>cow</td>
</tr>
<tr>
<td>coyote</td>
</tr>
<tr>
<td>deer</td>
</tr>
<tr>
<td>elk</td>
</tr>
<tr>
<td>horse</td>
</tr>
<tr>
<td>moose</td>
</tr>
<tr>
<td>average</td>
</tr>
<tr>
<td><strong>RES101</strong></td>
</tr>
<tr>
<td><strong>Class Specific</strong></td>
</tr>
<tr>
<td>class</td>
</tr>
<tr>
<td>bear</td>
</tr>
<tr>
<td>bison</td>
</tr>
<tr>
<td>cow</td>
</tr>
<tr>
<td>coyote</td>
</tr>
<tr>
<td>deer</td>
</tr>
<tr>
<td>elk</td>
</tr>
<tr>
<td>horse</td>
</tr>
<tr>
<td>moose</td>
</tr>
<tr>
<td>average</td>
</tr>
<tr>
<td><strong>RFCN</strong></td>
</tr>
<tr>
<td><strong>Class Specific</strong></td>
</tr>
<tr>
<td>class</td>
</tr>
<tr>
<td>bear</td>
</tr>
<tr>
<td>bison</td>
</tr>
<tr>
<td>cow</td>
</tr>
<tr>
<td>coyote</td>
</tr>
<tr>
<td>deer</td>
</tr>
<tr>
<td>elk</td>
</tr>
<tr>
<td>horse</td>
</tr>
<tr>
<td>moose</td>
</tr>
<tr>
<td>average</td>
</tr>
</tbody>
</table>

**TABLE III**: Training configuration #1 class-specific results for NAS, INRES, RES101 and RFCN
### Training Configuration #1: 1K/Seq/8

<table>
<thead>
<tr>
<th>Class</th>
<th>AP(%)</th>
<th>RP(%)</th>
<th>Score(%)</th>
<th>Recall(%)</th>
<th>Precision(%)</th>
<th>Average(%)</th>
<th>GT</th>
</tr>
</thead>
<tbody>
<tr>
<td>bear</td>
<td>50.16</td>
<td>50.86</td>
<td>36.15</td>
<td>47.53</td>
<td>52.61</td>
<td>50.07</td>
<td>25320</td>
</tr>
<tr>
<td>bison</td>
<td>53.61</td>
<td>51.01</td>
<td>28.55</td>
<td>45.93</td>
<td>59.69</td>
<td>52.81</td>
<td>31612</td>
</tr>
<tr>
<td>cow</td>
<td>14.22</td>
<td>18.95</td>
<td>97.88</td>
<td>46.31</td>
<td>6.66</td>
<td>26.48</td>
<td>4222</td>
</tr>
<tr>
<td>coyote</td>
<td>43.94</td>
<td>42.57</td>
<td>15.13</td>
<td>32.99</td>
<td>71.96</td>
<td>52.47</td>
<td>22337</td>
</tr>
<tr>
<td>deer</td>
<td>47.7</td>
<td>47.82</td>
<td>53.13</td>
<td>53.83</td>
<td>42.61</td>
<td>48.22</td>
<td>24465</td>
</tr>
<tr>
<td>elk</td>
<td>47.46</td>
<td>46.71</td>
<td>17.82</td>
<td>36.56</td>
<td>63.72</td>
<td>50.14</td>
<td>25353</td>
</tr>
<tr>
<td>horse</td>
<td>36.16</td>
<td>38.24</td>
<td>84.68</td>
<td>51.01</td>
<td>14.08</td>
<td>32.55</td>
<td>4840</td>
</tr>
<tr>
<td>moose</td>
<td>25.73</td>
<td>27</td>
<td>36.93</td>
<td>24.89</td>
<td>27.53</td>
<td>26.21</td>
<td>25033</td>
</tr>
<tr>
<td>average</td>
<td>39.87</td>
<td>40.4</td>
<td>46.28</td>
<td>42.38</td>
<td>42.36</td>
<td>42.37</td>
<td>163182</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class Specific</th>
<th>mRP threshold 41.20%</th>
<th>GT</th>
</tr>
</thead>
<tbody>
<tr>
<td>bear</td>
<td>50.16</td>
<td>25320</td>
</tr>
<tr>
<td>bison</td>
<td>53.61</td>
<td>31612</td>
</tr>
<tr>
<td>cow</td>
<td>14.22</td>
<td>4222</td>
</tr>
<tr>
<td>coyote</td>
<td>43.94</td>
<td>22337</td>
</tr>
<tr>
<td>deer</td>
<td>47.7</td>
<td>24465</td>
</tr>
<tr>
<td>elk</td>
<td>47.46</td>
<td>25353</td>
</tr>
<tr>
<td>horse</td>
<td>36.16</td>
<td>4840</td>
</tr>
<tr>
<td>moose</td>
<td>25.73</td>
<td>25033</td>
</tr>
<tr>
<td>average</td>
<td>39.87</td>
<td>163182</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class Specific</th>
<th>mRP threshold 11.30%</th>
<th>GT</th>
</tr>
</thead>
<tbody>
<tr>
<td>bear</td>
<td>48.2</td>
<td>25320</td>
</tr>
<tr>
<td>bison</td>
<td>39.95</td>
<td>31612</td>
</tr>
<tr>
<td>cow</td>
<td>28.56</td>
<td>4222</td>
</tr>
<tr>
<td>coyote</td>
<td>40.68</td>
<td>22337</td>
</tr>
<tr>
<td>deer</td>
<td>37.36</td>
<td>24465</td>
</tr>
<tr>
<td>elk</td>
<td>38.71</td>
<td>25353</td>
</tr>
<tr>
<td>horse</td>
<td>46.31</td>
<td>4840</td>
</tr>
<tr>
<td>moose</td>
<td>23.46</td>
<td>25033</td>
</tr>
<tr>
<td>average</td>
<td>37.9</td>
<td>163182</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class Specific</th>
<th>mRP threshold 27.30%</th>
<th>GT</th>
</tr>
</thead>
<tbody>
<tr>
<td>bear</td>
<td>54.58</td>
<td>25320</td>
</tr>
<tr>
<td>bison</td>
<td>39.8</td>
<td>31612</td>
</tr>
<tr>
<td>cow</td>
<td>28.22</td>
<td>4222</td>
</tr>
<tr>
<td>coyote</td>
<td>40.14</td>
<td>22337</td>
</tr>
<tr>
<td>deer</td>
<td>40.22</td>
<td>24465</td>
</tr>
<tr>
<td>elk</td>
<td>38.47</td>
<td>25353</td>
</tr>
<tr>
<td>horse</td>
<td>26.45</td>
<td>4840</td>
</tr>
<tr>
<td>moose</td>
<td>23.38</td>
<td>25033</td>
</tr>
<tr>
<td>average</td>
<td>38.52</td>
<td>163182</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class Specific</th>
<th>mRP threshold 0.50%</th>
<th>GT</th>
</tr>
</thead>
<tbody>
<tr>
<td>bear</td>
<td>19.3</td>
<td>25320</td>
</tr>
<tr>
<td>bison</td>
<td>18.07</td>
<td>31612</td>
</tr>
<tr>
<td>cow</td>
<td>13.39</td>
<td>4222</td>
</tr>
<tr>
<td>coyote</td>
<td>17.57</td>
<td>22337</td>
</tr>
<tr>
<td>deer</td>
<td>28.71</td>
<td>24465</td>
</tr>
<tr>
<td>elk</td>
<td>27.98</td>
<td>25353</td>
</tr>
<tr>
<td>horse</td>
<td>17.16</td>
<td>4840</td>
</tr>
<tr>
<td>moose</td>
<td>13.65</td>
<td>25033</td>
</tr>
<tr>
<td>average</td>
<td>19.48</td>
<td>163182</td>
</tr>
</tbody>
</table>

**TABLE IV:** Training configuration #1 class-specific results for RETINA, SSDIN, SSDMO and YOLO.
### Training Configuration #2: 1K/Even/8

#### NAS

<table>
<thead>
<tr>
<th>Class</th>
<th>AP (%)</th>
<th>RP (%)</th>
<th>Score (%)</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
<th>Average (%)</th>
<th>GT</th>
</tr>
</thead>
<tbody>
<tr>
<td>bear</td>
<td>97.78</td>
<td>93.94</td>
<td>93.93</td>
<td>95.12</td>
<td>90.44</td>
<td>92.78</td>
<td>25268</td>
</tr>
<tr>
<td>bison</td>
<td>97.53</td>
<td>94.44</td>
<td>59.38</td>
<td>92.85</td>
<td>97.57</td>
<td>95.21</td>
<td>31294</td>
</tr>
<tr>
<td>cow</td>
<td>97.03</td>
<td>94.93</td>
<td>45.74</td>
<td>92.76</td>
<td>98.7</td>
<td>95.73</td>
<td>4340</td>
</tr>
<tr>
<td>coyote</td>
<td>96.96</td>
<td>91.61</td>
<td>90.7</td>
<td>93.11</td>
<td>87.62</td>
<td>90.36</td>
<td>22300</td>
</tr>
<tr>
<td>deer</td>
<td>97.89</td>
<td>93.54</td>
<td>66.8</td>
<td>91.72</td>
<td>95.9</td>
<td>93.81</td>
<td>24374</td>
</tr>
<tr>
<td>elk</td>
<td>98.5</td>
<td>95.17</td>
<td>86.24</td>
<td>95.51</td>
<td>94.66</td>
<td>95.08</td>
<td>25292</td>
</tr>
<tr>
<td>horse</td>
<td>97.82</td>
<td>95.71</td>
<td>91.63</td>
<td>96.45</td>
<td>92.95</td>
<td>94.7</td>
<td>4759</td>
</tr>
<tr>
<td>moose</td>
<td>98.36</td>
<td>95.15</td>
<td>84.48</td>
<td>95.25</td>
<td>94.91</td>
<td>95.08</td>
<td>24943</td>
</tr>
<tr>
<td>average</td>
<td>97.73</td>
<td>94.31</td>
<td>77.36</td>
<td>94.1</td>
<td>94.09</td>
<td>94.09</td>
<td>162570</td>
</tr>
</tbody>
</table>

#### INRES

<table>
<thead>
<tr>
<th>Class</th>
<th>AP (%)</th>
<th>RP (%)</th>
<th>Score (%)</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
<th>Average (%)</th>
<th>GT</th>
</tr>
</thead>
<tbody>
<tr>
<td>bear</td>
<td>97.37</td>
<td>96.37</td>
<td>3.03</td>
<td>95.7</td>
<td>99.05</td>
<td>97.37</td>
<td>25268</td>
</tr>
<tr>
<td>bison</td>
<td>92.13</td>
<td>90.73</td>
<td>3.37</td>
<td>90.31</td>
<td>96.07</td>
<td>93.19</td>
<td>31294</td>
</tr>
<tr>
<td>cow</td>
<td>98.62</td>
<td>98.23</td>
<td>20.62</td>
<td>98.32</td>
<td>96.3</td>
<td>97.31</td>
<td>4340</td>
</tr>
<tr>
<td>coyote</td>
<td>98.67</td>
<td>96.75</td>
<td>24.83</td>
<td>97.47</td>
<td>94.66</td>
<td>96.07</td>
<td>22300</td>
</tr>
<tr>
<td>deer</td>
<td>98.53</td>
<td>97.03</td>
<td>38.24</td>
<td>97.92</td>
<td>93.09</td>
<td>95.5</td>
<td>24374</td>
</tr>
<tr>
<td>elk</td>
<td>98.13</td>
<td>97.2</td>
<td>10.21</td>
<td>97.13</td>
<td>97.69</td>
<td>97.41</td>
<td>25292</td>
</tr>
<tr>
<td>horse</td>
<td>92.16</td>
<td>90.1</td>
<td>3.77</td>
<td>90</td>
<td>97.45</td>
<td>93.72</td>
<td>4759</td>
</tr>
<tr>
<td>moose</td>
<td>99.5</td>
<td>98.27</td>
<td>69.91</td>
<td>99.21</td>
<td>91.52</td>
<td>95.36</td>
<td>24943</td>
</tr>
<tr>
<td>average</td>
<td>96.64</td>
<td>95.59</td>
<td>21.75</td>
<td>95.76</td>
<td>95.73</td>
<td>95.74</td>
<td>162570</td>
</tr>
</tbody>
</table>

#### RES101

<table>
<thead>
<tr>
<th>Class</th>
<th>AP (%)</th>
<th>RP (%)</th>
<th>Score (%)</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
<th>Average (%)</th>
<th>GT</th>
</tr>
</thead>
<tbody>
<tr>
<td>bear</td>
<td>99.72</td>
<td>99.53</td>
<td>88.64</td>
<td>99.52</td>
<td>99.59</td>
<td>99.55</td>
<td>25268</td>
</tr>
<tr>
<td>bison</td>
<td>99.79</td>
<td>99.42</td>
<td>86.91</td>
<td>99.35</td>
<td>99.53</td>
<td>99.44</td>
<td>31294</td>
</tr>
<tr>
<td>cow</td>
<td>99.86</td>
<td>99.65</td>
<td>89.18</td>
<td>99.65</td>
<td>99.72</td>
<td>99.69</td>
<td>4340</td>
</tr>
<tr>
<td>coyote</td>
<td>99.92</td>
<td>99.34</td>
<td>92.52</td>
<td>99.39</td>
<td>99.31</td>
<td>99.35</td>
<td>22300</td>
</tr>
<tr>
<td>deer</td>
<td>99.87</td>
<td>99.5</td>
<td>96.2</td>
<td>99.57</td>
<td>99.39</td>
<td>99.48</td>
<td>24374</td>
</tr>
<tr>
<td>elk</td>
<td>99.94</td>
<td>99.84</td>
<td>75.47</td>
<td>99.81</td>
<td>99.92</td>
<td>99.86</td>
<td>25292</td>
</tr>
<tr>
<td>horse</td>
<td>99.77</td>
<td>99.62</td>
<td>98.57</td>
<td>99.66</td>
<td>99.23</td>
<td>99.44</td>
<td>4759</td>
</tr>
<tr>
<td>moose</td>
<td>99.91</td>
<td>99.6</td>
<td>75.28</td>
<td>99.49</td>
<td>99.75</td>
<td>99.62</td>
<td>24943</td>
</tr>
<tr>
<td>average</td>
<td>99.85</td>
<td>99.56</td>
<td>87.85</td>
<td>99.56</td>
<td>99.55</td>
<td>99.56</td>
<td>162570</td>
</tr>
</tbody>
</table>

#### RFCN

<table>
<thead>
<tr>
<th>Class</th>
<th>AP (%)</th>
<th>RP (%)</th>
<th>Score (%)</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
<th>Average (%)</th>
<th>GT</th>
</tr>
</thead>
<tbody>
<tr>
<td>bison</td>
<td>99.63</td>
<td>99.09</td>
<td>79.33</td>
<td>98.99</td>
<td>99.23</td>
<td>99.11</td>
<td>31294</td>
</tr>
<tr>
<td>cow</td>
<td>99.88</td>
<td>99.72</td>
<td>49.54</td>
<td>99.65</td>
<td>99.86</td>
<td>99.76</td>
<td>4340</td>
</tr>
<tr>
<td>coyote</td>
<td>99.95</td>
<td>99.48</td>
<td>89.86</td>
<td>99.57</td>
<td>99.31</td>
<td>99.44</td>
<td>22300</td>
</tr>
<tr>
<td>deer</td>
<td>99.86</td>
<td>99.61</td>
<td>73.09</td>
<td>99.48</td>
<td>99.71</td>
<td>99.6</td>
<td>24374</td>
</tr>
<tr>
<td>elk</td>
<td>99.74</td>
<td>99.58</td>
<td>95.96</td>
<td>99.66</td>
<td>99.23</td>
<td>99.44</td>
<td>25292</td>
</tr>
<tr>
<td>horse</td>
<td>99.86</td>
<td>99.56</td>
<td>81.89</td>
<td>99.52</td>
<td>99.59</td>
<td>99.56</td>
<td>4759</td>
</tr>
<tr>
<td>moose</td>
<td>99.86</td>
<td>99.56</td>
<td>81.89</td>
<td>99.52</td>
<td>99.59</td>
<td>99.56</td>
<td>24943</td>
</tr>
<tr>
<td>average</td>
<td>99.8</td>
<td>99.53</td>
<td>80.09</td>
<td>99.51</td>
<td>99.5</td>
<td>99.51</td>
<td>162570</td>
</tr>
</tbody>
</table>

TABLE V: Training configuration #2 class-specific results for NAS, INRES, RES101 and RFCN
### Table VI: Training configuration #2 class-specific results for RETINA, SSDIN, SSDMO and YOLO

<table>
<thead>
<tr>
<th>Class</th>
<th>RETINA</th>
<th>SSDIN</th>
<th>SSDMO</th>
<th>YOLO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Class Specific</td>
<td>mRP threshold 28.30 %</td>
<td>mRP threshold 22.30 %</td>
<td>mRP threshold 34.10 %</td>
</tr>
<tr>
<td></td>
<td>AP(%)</td>
<td>RP(%)</td>
<td>Score(%)</td>
<td>Recall(%)</td>
</tr>
<tr>
<td>bear</td>
<td>99.88</td>
<td>99.66</td>
<td>29.67</td>
<td>99.68</td>
</tr>
<tr>
<td>bison</td>
<td>99.91</td>
<td>99.34</td>
<td>25.21</td>
<td>99.13</td>
</tr>
<tr>
<td>cow</td>
<td>100</td>
<td>99.82</td>
<td>52.52</td>
<td>99.95</td>
</tr>
<tr>
<td>coyote</td>
<td>99.99</td>
<td>99.64</td>
<td>42.03</td>
<td>99.96</td>
</tr>
<tr>
<td>deer</td>
<td>99.98</td>
<td>99.67</td>
<td>30.56</td>
<td>99.72</td>
</tr>
<tr>
<td>elk</td>
<td>99.99</td>
<td>99.8</td>
<td>38.48</td>
<td>99.89</td>
</tr>
<tr>
<td>horse</td>
<td>99.89</td>
<td>99.41</td>
<td>26.15</td>
<td>98.76</td>
</tr>
<tr>
<td>moose</td>
<td>99.96</td>
<td>99.72</td>
<td>27.8</td>
<td>99.7</td>
</tr>
<tr>
<td>average</td>
<td>99.95</td>
<td>99.63</td>
<td>34.05</td>
<td>99.6</td>
</tr>
<tr>
<td>bear</td>
<td>99.63</td>
<td>98.36</td>
<td>34.83</td>
<td>98.97</td>
</tr>
<tr>
<td>bison</td>
<td>99.79</td>
<td>99.18</td>
<td>13.43</td>
<td>98.87</td>
</tr>
<tr>
<td>cow</td>
<td>99.82</td>
<td>98.43</td>
<td>27.94</td>
<td>98.71</td>
</tr>
<tr>
<td>coyote</td>
<td>99.65</td>
<td>97.54</td>
<td>32.92</td>
<td>98.37</td>
</tr>
<tr>
<td>deer</td>
<td>99.88</td>
<td>98.89</td>
<td>13.25</td>
<td>98.23</td>
</tr>
<tr>
<td>elk</td>
<td>99.96</td>
<td>99.36</td>
<td>13.82</td>
<td>99.1</td>
</tr>
<tr>
<td>horse</td>
<td>99.78</td>
<td>98.72</td>
<td>19.4</td>
<td>98.59</td>
</tr>
<tr>
<td>moose</td>
<td>99.76</td>
<td>98.55</td>
<td>13.25</td>
<td>97.67</td>
</tr>
<tr>
<td>average</td>
<td>99.78</td>
<td>98.63</td>
<td>21.1</td>
<td>98.56</td>
</tr>
<tr>
<td>bear</td>
<td>99.64</td>
<td>98.56</td>
<td>19.39</td>
<td>97.1</td>
</tr>
<tr>
<td>bison</td>
<td>99.72</td>
<td>98.37</td>
<td>49.9</td>
<td>98.83</td>
</tr>
<tr>
<td>cow</td>
<td>99.88</td>
<td>98.55</td>
<td>46.9</td>
<td>99.24</td>
</tr>
<tr>
<td>coyote</td>
<td>99.87</td>
<td>98.52</td>
<td>35.81</td>
<td>98.67</td>
</tr>
<tr>
<td>deer</td>
<td>99.9</td>
<td>98.81</td>
<td>27.96</td>
<td>98.35</td>
</tr>
<tr>
<td>elk</td>
<td>99.93</td>
<td>99.09</td>
<td>19.37</td>
<td>98.49</td>
</tr>
<tr>
<td>horse</td>
<td>99.69</td>
<td>99.01</td>
<td>23.16</td>
<td>97.48</td>
</tr>
<tr>
<td>moose</td>
<td>99.73</td>
<td>97.47</td>
<td>37.34</td>
<td>97.85</td>
</tr>
<tr>
<td>average</td>
<td>99.79</td>
<td>98.55</td>
<td>32.48</td>
<td>98.33</td>
</tr>
<tr>
<td>bear</td>
<td>99.92</td>
<td>96.74</td>
<td>11.85</td>
<td>97.33</td>
</tr>
<tr>
<td>bison</td>
<td>98.18</td>
<td>96.17</td>
<td>5.81</td>
<td>95.98</td>
</tr>
<tr>
<td>cow</td>
<td>98.56</td>
<td>95.18</td>
<td>7.31</td>
<td>95.35</td>
</tr>
<tr>
<td>coyote</td>
<td>96.97</td>
<td>94.57</td>
<td>5.5</td>
<td>94.24</td>
</tr>
<tr>
<td>deer</td>
<td>99.04</td>
<td>96.04</td>
<td>14.66</td>
<td>96.94</td>
</tr>
<tr>
<td>elk</td>
<td>97.97</td>
<td>97.75</td>
<td>0.38</td>
<td>97.22</td>
</tr>
<tr>
<td>horse</td>
<td>97.5</td>
<td>95.15</td>
<td>5.73</td>
<td>94.94</td>
</tr>
<tr>
<td>moose</td>
<td>99.85</td>
<td>99.22</td>
<td>3.63</td>
<td>98.96</td>
</tr>
<tr>
<td>average</td>
<td>98.37</td>
<td>96.35</td>
<td>6.86</td>
<td>96.37</td>
</tr>
</tbody>
</table>